Care of extremely premature babies in England, 1995 – Present

Andrei Scott Morgan

Department of Neonatology

Institute for Women's Health

University College London

A thesis submitted for the degree of

Doctor of Philosophy (PhD)

June 2015

I, Andrei Scott Morgan, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date: 17 July 2015

Supervisors

Neil Marlow
Professor of Neonatology

Department of Neonatology
Institute for Women's Health
University College London

Elizabeth S. Draper
Professor of Perinatal and Paediatric Epidemiology

Department of Health Sciences
University of Leicester

Abstract

This thesis, arising from the EPICure studies into extremely preterm births, seeks to explain demographic, clinical and organisational factors associated with a large increase in admissions to neonatal intensive care. Using six pre-existing data sets, along with a repeat Unit Profile Study of English perinatal centres conducted in 2011 as part of this thesis, three studies were performed:

- Validation of the 44% increase in the number of admissions to neonatal intensive care at 22–25 weeks gestation seen between 1995 and 2006 in England was attempted using three methods of probabilistic record linkage with Hospital Episode Statistics (HES) data.
- The effects of antenatal steroid administration, tocolysis and Caesarean delivery on perinatal outcomes in the extremely preterm population were investigated.
- Changes in organisational characteristics – staffing and "activity" (expressed as *throughput* and *intensity*) – that have occurred in England were examined using data from three time points.

An increase was seen in the number of extremely premature babies in HES data. Linkage with EPICure data demonstrated that routine data are insufficiently precise for use in epidemiological investigations at the margins of viability. Tocolysis was associated with improved outcomes. Antenatal steroids were associated with improved outcomes at birth following vaginal delivery. No effect was demonstrated for Caesarean delivery on birth outcomes but there was evidence of case selection at gestations below 26 weeks. Organisational data (from 1997, 2006 and 2011) demonstrated reduced numbers of cots between 1997 and 2011 with increases in both throughput and intensity of workload. Staffing levels increased, but still failed to meet recommended standards. Current knowledge of extremely low gestational age births is inadequate for national policy or health care reorganisation. Suggestions were made for how knowledge could be improved.

To Sylvia O.

Who gave to me my love
of maths and people ...

Acknowledgements

I'd first like to say a huge thanks to my supervisors – Neil Marlow (Professor of Neonatology, Institute for Women's Health, UCL) and Liz Draper (Professor of Perinatal Epidemiology, University of Leicester) – who not only helped plan this research project, but provided much valuable guidance and support throughout the time I have been working on it. I'm also very grateful to the other members of the EPICure Research Group: Professor Kate Costeloe and Ms Enid Hennessy, without whom I would not have been able to understand the data in the way that I do. Thank you, too, to Dr. Laura McCormack for the help with coordinating the Unit Profile Study.

There are many people who provided me with help and support along the way. I'd particularly like to give my thanks to those who expressed their faith in me over the last few years: Dan Arthur, Pat Doyle, Phil Edwards, Branwen Hennig, Craig Higgins, Emma Nabavian and Kate Walker of the London School of Hygiene and Tropical Medicine; and Angela Huertas-Ceballos of University College London Hospitals NHS Foundation Trust. Also, the many colleagues I have shared desks or working arrangements with – Tauseef Khan, Shahrul Kamaruzzaman, Amira Shaheen, Eleonora Staines-Urias, Snehal Pinto Pereira, James Woodcock, Beth Howden, Giles Kendall, Cally Tann, Cristina Uria-Avellano and Kate Bennett.

To my many friends, who offered me routes of escape and wonderful meals when the pressures and writing all became too much: without you, I would never have survived! To each and every one of you, big hugs: Rob and Jenny Hunningher, Ben McAlister, Helen Dale, Dave Hrycyszyn, Pennie Quinton, William Lee, Emily Dook, Christopher Grollman, Catriona Towriss and Sam Moon. To the people who read through my thesis – Katie Harron, Chris, Helen, Cat and my mother: you now have an insight into how my head works. Thank you all for your time and constructive criticism. Thanks, too, to the inspirations for my clinical and epidemiological career - Henry Kahn and Mickey Gillmor, Robert and Marilyn Kessler. Finally, to my family: I could not have done this without you – you have my thanks and my love.

Contents

CONTENTS

List of Figures

LIST OF FIGURES

List of Tables

Part I

Background information

Chapter 1

Introduction

Rates of preterm birth have been increasing in many countries over the last two decades.[1, 2] During this time, improvements in newborn care resulting from evidence-based interventions and improved technology have led to increased survival rates of those born extremely prematurely - but not to improvements in intact (morbidity-free) survival.[2] In England, coincident with these changes, neonatal care has undergone a reorganisation into managed clinical networks.[3] The effects of this, and the influences these changes may have exerted on both individuals and society, have not been extensively studied to date.

This thesis arises from the larger EPICure research studies into extremely preterm births. The investigations in this thesis seek to explain the large increase seen in England between 1995 and 2006 in the number of extremely premature babies, from 22 to 25 completed weeks (that is, up to and including 25 weeks and 6 days) gestational age, admitted to neonatal intensive care and, where possible, to extend that knowledge to the present time. Several different sources of data are used for this. Data from the UK Neonatal Staffing Study (conducted in 1997) and from the Unit Profile Study (conducted in 2006) are analysed to examine the organisation of hospital neonatal services. Both data sets refer to the provision of care in the preceding calendar year. These data are supplemented by a revised version of the Unit Profile Survey that was conducted as part of this thesis and which took place in November 2011.

Detailed data about births within the specified gestational age range are obtained from two pre-existing cohort studies that took place in England: the first (known as EPICure) in 1995 and the second (EPICure 2) in 2006. The available data are

supplemented with information obtained from Hospital Episode Statistics (HES), a routinely collected, national administrative database primarily collected for non-clinical purposes.

1.1 Thesis overview

The thesis commences in Part I with a critical appraisal of the currently available literature on the organisation of neonatal care and the changes in practice that have occurred in England since the early 1990s; these topics are presented in Chapter 2, alongside an overview of methodological techniques that will be used. Chapter 3 develops the aims, before moving onto a general introduction to the methods and data sets – including the second Unit Profile Study carried out in November 2011 – used for the different studies that make up this thesis in Chapter 4.

The main body of work is presented in linked chapters (for methods, results and discussion) in the three following parts: Part II contains detail on the record linkage between the EPICure and HES data sets for each of the two years for which data are available (1995 and 2006); Part III describes an investigation into obstetric antecedents of extremely preterm delivery (< 27 completed weeks of gestation); and Part IV reports on the changes in staffing, throughput and activity in English neonatal units that occurred between 1997 and 2011.

The final part of the thesis, Part V, contains a detailed discussion of the findings along with the conclusions. This is followed by the bibliography and appendices – which contain substantial extra information supplemental to the rest of the work.

Chapter 2

Organisation of neonatal care

This chapter details the knowledge base from which the aims are derived and on which the overall thesis is constructed. I begin by explaining some of the key terminology that will be used in the thesis. Then, there is a description of how neonatal care was provided in England during the 1990s, as well as discussion of some of the investigations into extreme prematurity that were performed at that time. These investigations helped form the evidence base for the subsequent reorganisation of neonatal care in England into managed clinical networks (MCNs), now known as operational delivery networks (ODNs), detail of which is presented in section 2.4.

The EPICure 2 study was conducted in 2006 and is discussed in some detail in this chapter. This is accompanied by a summary of other data sources available in England and discussion around techniques for linking data sets together. Further statistical and epidemiological considerations are also presented – specifically, around the three key areas of chance, bias and confounding. The last part of the chapter (section 2.10) focuses on workload measurement and assessment, and uses some examples from the neonatal setting.

2.1 Definitions

Since the introduction of neonatology as a unique discipline in the 1960s,[4] care of the newborn has improved dramatically, with corresponding increases in survival of both term and preterm babies.[5–8] There are, however, a wide range of classifications used in discussions relating to neonatal care, which can make it difficult to ensure that like

is compared with like. It is therefore necessary to establish precise definitions of key terminology.

Traditionally, pregnancy has been determined as starting at the beginning of the mother's last menstrual period (LMP). With the introduction of ultrasound, it became possible to estimate gestational age with greater precision, even when a woman was uncertain of her dates. Thus, ultrasound has superseded use of LMP for establishing the estimated date of delivery (EDD) and, consequently, gestational age.

Until 20-24 weeks of gestational age, a fetus is considered *pre-viable* as the lungs have yet to develop beyond the cannilicular stage in the embryological development of the respiratory system.[9] The saccular period, considered to begin at 24 weeks of gestation (although it may commence earlier), heralds the beginning of the development of the (future) terminal airways from "saccules" into the alveoli of the full-term baby.[9] Endogenous surfactant from type II alveolar cells may be detected at this stage, but production does not mature sufficiently to protect against respiratory distress syndrome (RDS) until between 29 and 32 weeks gestational age.[9, 10]

These biological borders to the limits of viability, in conjunction with current social perceptions as codified in law,[11] help frame numerous definitions – terms such as "live birth", "miscarriage", "abortion", "termination of pregnancy", "still birth (or late fetal death)", "preterm" or "premature" , and so on. These and many others are listed in the glossary, beginning on page 295. Nevertheless, there is a "grey zone" that begins somewhere around 22-23 weeks gestational age where the conceptus may show signs of life (gasping, heart rate or limb movement) following delivery. For this reason, it is recommended that feticide is performed before termination of pregnancy at these gestations.[11, 12]

One crucial point is that there is no universally accepted definition of "extreme prematurity." While premature births are widely accepted as being those of less than 37 weeks gestation,[13], different studies have variously defined the population of "extremely premature" as being below 26, 27, 28 or 32 weeks gestation.[14] Throughout this thesis, therefore, gestational age will be explicitly stated where appropriate.

Babies requiring postnatal support may also be classified according to their birth weight – using the first measurement taken within an hour of birth. This may be based on absolute measurement or weight relative to that expected for gestational age. Absolute measurement has three classifications: low birth weight (LBW: < 2500g), very low

birth weight (VLBW: < 1500g), or extremely low birth weight (ELBW: < 1000g).[15] Relative weights fall into three groups: small for gestational age (SGA: $< 10^{th}$ centile), appropriate for gestational age (AGA: 10^{th} – 90^{th} centiles), and large for gestational age (LGA: $> 90^{th}$ centile).[16] These descriptions should be differentiated from *intrauterine growth restriction* (or *retardation* – IUGR), which represents a slower than otherwise expected *in utero* growth velocity, usually seen on serial ultrasound assessment. Specifically, it should be noted that IUGR may affect infants who are large for gestational age in early pregnancy but appropriately sized at birth.

Clarification is also required for the terminology used to describe the organisation of neonatal care. Particularly, it is important to distinguish between the classification of neonatal units and the categorisation of care that babies receive – determining the cot status, types of baby cared for in each, and the relevant staffing levels – as they are each frequently grouped into three, similarly-named levels, and details have changed over time. The "Report of working group of the British Association of Perinatal Medicine and Neonatal Nurses Association (NNA) on Categories of babies requiring neonatal care" was published in 1992,[17] and defined four levels of care for babies, including the types of babies who may meet criteria for the differing care levels, and the relevant resources (including equipment and staff) required for each baby being cared for.[17] The levels defined were: maximal intensive care, known as level one intensive care; high dependency, known as level two intensive care; special care; and normal care.[17]

Classification of neonatal units, on the other hand, was first defined by the British Association of Perinatal Medicine (BAPM) in 1996.[18] This categorised units by whether or not they provided intensive (either level one or level two – intensive or high dependency) care on an ongoing or temporary basis only, with a recommendation that care be organised on a regional basis.[18] This document specified the overall staffing requirements for the unit, as well as suggesting that nurse-patient ratios should not be less than one-to-four (in special care) with a recommendation that "[nurses] should not have responsibility for more than two infants receiving neonatal intensive care"[18] and noting that there were occasions where one-to-one nursing was appropriate.[18]

These two documents were merged in 2001 changing some of the terminology, but resulting in clearer guidelines. Care was designated as intensive, high dependency or special care, with recommended nurse staffing ratios of 1:1, 1:2, and 1:4, respectively, alongside a fourth category of "normal care" that did not requiring any additional

neonatal staffing.[19] Problems that babies might experience necessitating these differing categories of care were described, as were equipment requirements for each cot space.[19] Consequently, areas within hospitals became known as neonatal intensive care units (NICUs), high dependency units (HDUs) and special care baby units (SCBUs). Hospital units were, by contrast, designated as Level 1 (units providing special care only), Level 2 (units providing high dependency and short term intensive care) or Level 3 (providing a complete range of medical intensive care).[19] Additional staffing requirements at a unit level were also noted – for example, the necessity to have appropriately staffed (and equipped) laboratory and supporting services (for example, in radiology, microbiology, pharmacology and nutrition).[19]

Recommendations were changed again in 2009 with publication of the NHS Toolkit for High-Quality Neonatal Services by the Department of Health. This recommended that units were categorised as special care units (SCUs), local neonatal units (LNUs) and network neonatal intensive care units (NICUs) instead of Levels 1, 2 and 3.[20] This change in name for the description of units was further ratified by the second edition of BAPMs's "Service Standards for Hospitals Providing Neonatal Care" in 2010.[21] NICUs were the largest units dealing with the sickest babies and described as having dedicated medical rotas at tiers 1 and 2 (resident on-site, equivalent to the old senior house officer and registrar grades) as well as tier 3 (consultants, available 24 hours a day); LNUs would provide care for babies from lower-risk pregnancies, and have a dedicated medical rota at tier 1 plus a minimum of 1 dedicated consultant (tier 3); and SCUs have no dedicated neonatal facilities but are equipped to deal with unexpected emergencies, usually existing alongside a general paediatric service.[20, 21]

The categorisation of individual care levels for babies was however deferred and instead published as a separate document in 2011.[22] This added a new level of care, transitional care (TC), where mothers remained resident alongside their babies – resulting in areas within individual hospitals called transitional care units (TCUs). Thus, the current recommendations are as shown in table 2.1.

2.2 Neonatal care in England in the Nineties

Provision of neonatal care in England during the 1990s was mostly ad-hoc and determined according to local needs and policies; [23] a majority of babies born at less

Table 2.1: Recommended categories of care for babies. Adapted from guidance produced by the British Association of Perinatal Medicine (BAPM).[19, 21, 22]

Cot level	Nurse:patient ratio	Extra information
Intensive care	1 : 1	Medical staff immediately available. Continuous monitoring. Likely to require invasive respiratory support (mechanical ventilation).
High dependency care	1 : 2	Babies requiring non-invasive respiratory support, parenteral nutrition, or those considered to be at high risk who require continuous monitoring.
Special care	1 : 4	Babies who are unable to be cared for (yet) at home, but who are otherwise not unwell (e.g. may require supplementary oxygen, tube feeding or intravenous fluids, or phototherapy).
Transitional care	[No recommendation made]	Babies who are primarily cared for by their mothers but also require some medical treatment that may be delivered on a postnatal ward, e.g. administration of intravenous antibiotics or phototherapy.
Normal care	—	Babies who do not require any care and need not be in hospital.

than 28 weeks did not survive.[24] However, this was a decade of enormous change, particularly for clinical management, and regional networks were already developing in some areas.

2.2.1 Organisation of neonatal care

In contrast to the practice of regional organisation advocated in the United States,[25] there was little co-ordination of neonatal care beyond a district level in England and most hospitals provided care based on local considerations – particularly with regards to the type of obstetric service provided locally and in relation to other facilities in the region.

A United Kingdom-wide survey carried out in 1994 identified 204 neonatal units in England, and 46 units in the remainder of the UK.[26] This documented that the number of intensive care cots (confusingly, called level 1 cots in the paper, as per the recommendations of the time[17]) had increased in number by approximately 25% since the late 1980s, and that nursing levels appeared to be appropriate for larger units, but

smaller units – which formed a majority – were probably "minimally staffed".[26]

However, despite the 1992 BAPM recommendations described earlier for the classification of care provided for babies by individual units,[17] there was no universal system in use, and different units were able to use either the BAPM recommendations, former guidelines from the British Paediatric Association (1984),[27], those from the Northern Neonatal Network (1993) or to create their own.[27] There were few consistencies between these sets of guidelines, other than agreement that a baby receiving supplementary oxygen or intravenous fluids was considered to be in special care, and babies requiring a higher degree of respiratory support received intensive care.[27] This lack of standardisation between units made it difficult to compare outcomes between different units or regions of the country.

2.2.2 UK Neonatal Staffing Study (UKNSS)

In 1997, a repeat survey was carried out in all neonatal units in the United Kingdom as part of the UK Neonatal Staffing Study (UKNSS) – funded by the NHS Executive, and supported by BAPM, the Neonatal Nurses Association (NNA) and the Scottish Neonatal Nursing Group (SNNG).[28, 29] This sought to identify neonatal intensive care units within the United Kingdom through a census of all hospitals, and to allow "stratification [of those neonatal units] by primary organisational characteristics."[29]

Of the 246 respondent hospitals, 186 (76%) stated that they provided intensive care while 60 (24%) provided stabilisation or special care only. [28] Further assessment was limited to units providing intensive care, and investigations were conducted into the level of activity and the levels of neonatal staffing recorded in each unit.[29]

Due to the lack of standardisation between hospitals in definitions and recording of data, the comparator for unit activity that was reported as being "the most complete, comparable, and robust proxy measure available" [28] was the number of VLBW babies admitted per year. Definitions based on ventilation – either type or number of babies or total days of respiratory support – varied between units and could not be applied on a broad scale.[28]

As per the 1994 survey, the UKNSS demonstrated that nursing numbers were insufficient: 147 (79%) of the 186 NICUs had *actual*:*estimated* staffing ratios of less than one. This was conservatively estimated, based loosely upon the 1992 BAPM / NNA

recommendations,[17] using a level of 5.5 whole-time equivalent (WTE) nurses per intensive care (level 1) cot but only 1 WTE nurse per high dependency or special care cot.[29]

With relation to medical staffing, the UKNSS demonstrated that 25% (46) of NICUs did *not* have a minimum of one dedicated neonatal consultant (defined as having > 50% clinical time available for neonatal care).[28] Further inconsistencies between hospitals included the number of routine, "business" wards rounds conducted per week – which ranged from 1-18 in those units that conducted routine ward rounds; in some units, individual consultants saw their own patients.[29]

Although the UKNSS sought to assess staffing levels and activity at neonatal units throughout the UK, it provided only a very superficial summary. There was no breakdown of the number of nurses employed at different grades or in different roles, other than an attempt to identify the proportion of nurses with a "nationally-recognised qualification in neonatal intensive care" (see questions 13 and 14, appendix C). Similarly, the questionnaire asked only about the number of dedicated neonatal consultants and did not enquire about medical staffing numbers at junior levels. Combined with the lack of standardised care definitions between units, this limited the questions the survey was able to answer. In contrast, the survey was extremely successful in that it achieved responses from all hospitals. This may have been the result of using limited questions that were relatively straight-forward to answer, although it is noteworthy that the more specific the question, the fewer responses were achieved.[28] However, this reporting bias was restricted to the questions on unit activity – predominantly, with respect to the numbers of patients who received ventilatory support, and for how long – and is unlikely to affect the overall conclusions. For these reasons, the results of the survey should be taken as an accurate portrayal of neonatal care in the United Kingdom at that time.

There was, subsequently, a second phase to the UKNSS that sought to provide more detail. A prospective investigation, based on the findings of the initial UKNSS staffing survey, was undertaken in a random sample of 54 neonatal units conducted over thirteen months between March 1st 1998 and April 2nd, 1999.[29] Data were collected at each unit on workload (twice-daily over a period of 13 months); neonatal outcomes; economic data and resource usage (e.g. pharmacy and laboratory costs); and the psychological well-being of staff. [29, 30] The study population focused on babies born at <31 weeks

gestation and/or <1500g and used a predictive measure of mortality that was derived from ten factors (gestation, birth weight, standard deviation score (SDS), sex, mode of delivery, diagnosis, maternal antenatal steroids, admissions temperature, worst P_aCO_2 (partial pressure of carbon dioxide in arterial blood), mean FiO_2 (fraction of inspired oxygen), worst base deficit (or, base excess)) available at 12 hours of age.[30] This was despite a simpler score having been developed (clinical risk index for babies II (CRIB II), based on just five factors that were available at one hour of age) that had been shown to be more accurate using data from the UKNSS. [31]

Characteristics of nurse staffing in the prospective investigation were described based on availability for a single shift at a time: "provision ratios" were constructed of the *actual* to *predicted* numbers of staff per shift for all nurses and for those who were *qualified in specialty (QIS)* – in possession of a specialist neonatal qualification such as the English Nursing Boards 402, 405 or SNNG's A19 certificate.[30] The expected number of nurses for all nursing staff (P_{all}) on a shift was obtained using:

$$P_{\text{all}} = 1 + \frac{N_{\text{IC}}}{2} + \frac{N_{\text{HD}}}{2} + \frac{N_{\text{SC}}}{4} \tag{2.1}$$

where N_{IC}, N_{HD}, and N_{SC} correspond to the number of babies admitted into intensive, high dependency and special care cots, respectively.[30] The value for the predicted number of specialist nurses (P_{qis}) was similarly obtained – although not counting the number of admissions into special care[30]:

$$P_{\text{qis}} = 1 + \frac{N_{\text{IC}}}{2} + \frac{N_{\text{HD}}}{2} \tag{2.2}$$

In order to facilitate comparison with mortality, the ratios for each shift to which a baby was exposed during the duration of the baby's admission (i.e. until death or discharge from the hospital of observation) were averaged to provide a single value for each baby. These values were then used alongside various other organisational characteristics by the authors as covariates in regression models exploring effects on the observed mortality.[30] It is not clear precisely what variables were included during the model building – although good guesses can be made by considering the questions asked during the first part of the UKNSS (see appendix C). Nor is it clear how exactly the model was constructed, although we are told "[b]irthweight, unit organisational characteristics (size, consultant availability, nursing establishment levels), number of

nurses per shift and nurse provision ratio per shift were excluded in the final risk-adjusted mortality model." [30] Specifically, it is doubtful whether the stated significance value of 0.05 was correctly applied: this cut-off appears to have been used to interpret the Wald test results for individual factors, rather than for maximum likelihood testing between different models.

Results are then reported for gestation, the mortality prediction score, and the QIS nurse staffing provision ratio. This showed a relationship in favour of survival adjusted for gestational age (odds ratio (OR) 0.745, 95% CI: 0.67 – 0.83, $p < 0.001$), an extremely strong favourable association with the predicted mortality score (which also accounted for gestational age) of 0.008 (95% CI: 0.003 – 0.019, $p < 0.001$) and another favourable association with the specialist nurse provision ratio of 0.63 (95% CI: 0.42 – 0.96, $p = 0.031$).[30] Given the uncertainty surrounding the methods and the fact that this was an exploratory analysis, it is questionable how important these results are.

Of note, this study also initially saw an association between unit volume and mortality. After adjustment using the risk scores, however, there were no differences between units with low, medium or high activity (defined by number of LBW admissions per year: 0-34, 35-57 and 58+).[30]

2.2.3 Transport services

The disparate nature of neonatal care during the 1990s had an impact on transport services. Data from the 1994 neonatal survey showed that while the number of postnatal transfers remained relatively constant in the early 1990s – approximately 1,650 per year (range 1,510 – 1,799) – the number of antenatal transfers increased year-on-year from 1,368 in 1989 to 1,558 in 1993.[26] In 1999, another survey was carried out over a 3-month period of the 37 tertiary perinatal centres in the UK. This sought to identify the reasons and appropriateness for transfer, determined by whether national criteria produced by the Clinical Standards Advisory Group (CSAG) were met or not. There were 264 *in utero* and 45 postnatal transfers, of which 245 and 43, respectively, were due to lack of an intensive care cot – although 9 of the latter were, in turn, due to a shortage of neonatal nursing staff.[32] While these data are restricted to the largest, tertiary level NICUs, and therefore may not be representative of lower level units, the figures highlight the difficulties in determining the *actual* reason for transfer, as opposed to the stated reason.

2.2.4 Medical advances

In contrast to the relatively static organisation of neonatal services during the 1990s, care patterns for individual patients underwent a dramatic change – especially for those born prematurely – following the introduction of antenatal steroids and surfactant (see below) for routine care. Combined with changes in management strategies afforded by the introduction of techniques such as high-frequency oscillatory ventilation (HFOV) and non-invasive continuous positive airway pressure (CPAP), mortality rates fell throughout the decade.[33]

Steroids

Antenatal steroids were coming into vogue in the 1990s. Although numerous papers were published covering a wide range of studies, the first systematic review [34] attempted to provide a comprehensive overview of all unpublished studies. Contact was made with 42,000 obstetricians and paediatricians across 18 countries, and additional data were received from previously published trials in order to facilitate comparisons. This review demonstrated a reduction in RDS for babies whose mothers received steroids at any time antenatally, with an OR of 0.46, 95% CI: 0.41 – 0.60, compared to babies whose mothers were not treated. No differences were found in effect by fetal sex (OR for males 0.43, 95% CI: 0.29 – 0.64; for females, OR 0.36, 95% CI: 0.23 – 0.57), and although most babies included in the studies were between 31 and 34 weeks, the effect was more marked below this age (OR 0.38, 95% CI: 0.24 – 0.60). The authors concluded, "we have not been able to identify any subgroup of babies for which it can be concluded that corticosteroid administration before delivery is *not* associated with a reduction in the risk of neonatal respiratory morbidity." [34]

The use of steroids antenatally was further encouraged by a consensus statement released by the National Institute of Health in the United States.[35] This reported that there was good evidence to support the use of antenatal steroids (either dexamethasone or betamethasone) in pregnancies at risk of delivering prior to 34 weeks gestational age, with no difference by "fetal race or gender, or by the availability of surfactant replacement therapy."[35] A first guideline by the Royal College of Obstetricians and Gynaecologists (RCOG) was issued in April 1996, with subsequently updated versions in 1999 and 2004,[36] all based on an updated version of Crowley's review that had

been incorporated into the Cochrane Library.[36, 37] This was superceded by a further review in 2006 that included data from 4,269 babies born to 3,885 women in 21 trials;[38] however, the number of babies born before 26 weeks was limited, with only 49 babies in one trial identified by the review.[38]

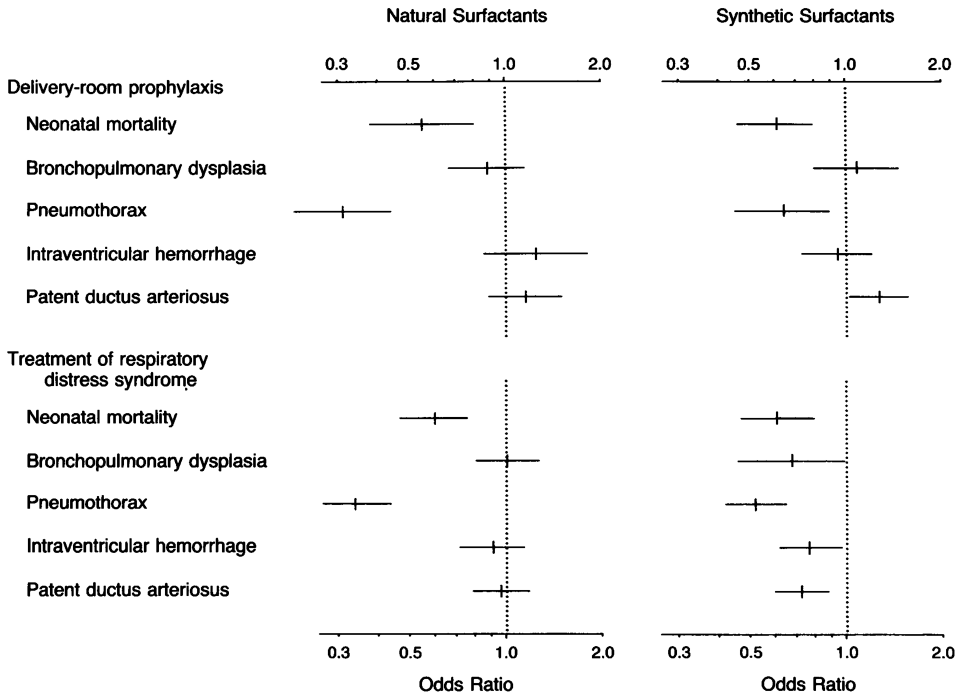
Surfactant

Surfactant deficiency had been recognised as important in the aetiology of hyaline membrane disease (the name given to the pathological diagnosis associated with the clinical picture of "respiratory distress syndrome") since the 1950s,[39, 40] but it was not until 1980 that the first report was published of surfactant therapy in humans.[41] This reported the use of an artificial surfactant in ten babies born between 28 and 33 weeks gestational age at a mean age of 12.3 (range 4 to 33) hours, eight of whom survived (one died of complications following an operation to correct oesophageal atresia with a tracheo-oesophageal fistula, and the other of sepsis caused by *Serratia spp.*).[41] In all babies, the P_aO_2 (partial pressure of oxygen in arterial blood), FiO_2 and $AaDO_2$ (alveolar-arterial difference in oxygenation) improved dramatically within 3 hours of endotracheal surfactant adminstration.[41]

There followed a period of intense study and debate around the best regimes for surfactant usage: between 1985 and 1992, for example, over 35 different randomised controlled trials took place.[40] Four meta-analyses in the Oxford Database of Perinatal trials (the predecessor to the Cochrane Reviews) subsequently produced by the Cochrane Collaboration[42] published in 1991 examined natural and artificial surfactants, and whether they should be used for prophylaxis or administered after several hours of age when the diagnosis of RDS manifested itself more severely.[40] These demonstrated a reduction in the odds of mortality to approximately 0.6 for both natural and synthetic surfactants administered either as prophylaxis or treatment for RDS (figure 2.1).[40] Additionally, the incidence of pneumothoraces was reduced by around a third with artificial surfactants, and by approximately 60% with natural surfactants, irrespective of the time of administration.[40]

There were four brands of surfactant in common usage in the UK in 1995: pumactant ("ALEC", synthetic surfactant), colfosceril palmitate ("Exosurf", synthetic surfactant), poractant alpha ("Curosurf", derived from pigs) and beractant ("Survanta" derived from cows), although the relative merits of each were still being ascertained,[39,

Figure 2.1: Collated results from four early meta-analyses of surfactant trials. Reproduced with permission from Jobe (1993) [40], Copyright Massachusetts Medical Society.



43, 44] with numerous randomised trials being conducted. These varied greatly in size and quality. For example, a small UK trial with just 75 subjects[45] compared poractant alpha with beractant, seeking to examine very acute effects; they didn't find any differences, although the group receiving poractant alpha appeared to have lower ventilatory and oxygen requirements after treatment.[45]

Other trials sought to answer similar questions to each other. In the United States, several large trials were published. The first of these, a study conducted within the Vermont-Oxford Neonatal Network (VON) enrolled 1,296 babies in a randomised controlled trial (RCT) from 4,895 admitted across 38 different centres[46]. This sought to compare colfosceril palmitate with beractant, but did not find any differences in death or chronic lung disease (CLD) between the two groups.[46] Another group compared beractant with calfactant ("infasurf", derived from cows).[47] They reported a double-blinded, prospective RCT in 13 institutions that enrolled babies of <2000g and <48 hours of age into either prophylaxis (if the mothers could be enrolled prior to delivery) or treatment arms and that aimed to compare short term effects. Despite the fact that

there were significantly more deaths in the calfactant-treated prophylaxis group (7 v. 2, p=0.01), the authors concluded that both beractant and calfactant were associated with improvements in respiratory status, but that these appeared to be greater and more sustained with calfactant.[47]

There was also a study comparing calfactant with colfosceril palmitate. This was carried out in ten centres and enrolled 894 of a potential 1,177 babies, although data were only available for analysis for 853 babies.[48] In this study, the authors also concluded that there were "several important advantages" associated with the use of calfactant.[48] In fact, their results had shown higher ORs for intraventricular haemorrhage (IVH) (OR 1.30, 95% CI: 1.08 – 1.57) and periventricular leukomalacia (PVL) (OR 2.03, 95% CI: 1.09 – 3.80).[48] Perhaps not surprisingly, this study and the one comparing calfactant with beractant were funded by the manufacturer of calfactant.[47, 48] They are not the only drug company to present misleading information, however; in 2012, the United States' Food and Drug Administration (FDA) sent letters to two companies (the manufacturers of calfactant and poractant alpha) warning that their advertising was misleading and overstated benefits of their drugs in relation to others.[49]

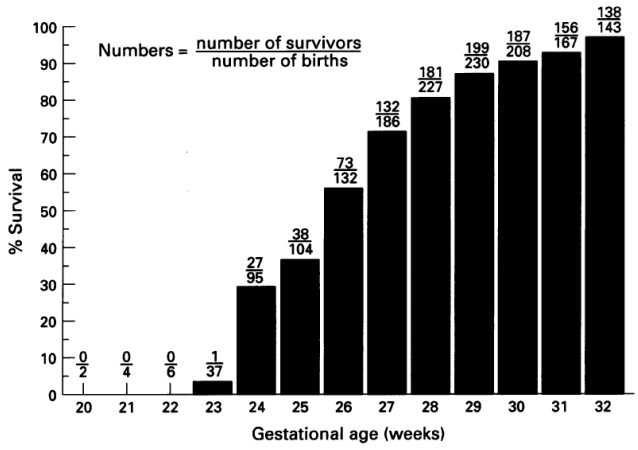
Further systematic reviews and meta-analyses around the turn of the century also attempted to provide definitive answers to some of the outstanding questions about surfactant usage. For example, a big question was whether to administer it as prophylaxis (in the delivery room) or to save surfactant for use as a rescue therapy (i.e. a few hours after birth, following formal diagnosis of RDS). This was answered by a systematic review in 1997 that identified six studies published between 1991 and 1995 that produced a combined OR of 0.55 (95% CI: 0.41 – 0.73) in favour of reduced neonatal mortality and an OR 0.59 (95% CI: 0.42 – 0.82) for reduced total mortality to discharge with prophylactic surfactant.[50] In the same review, the ORs for both pneumothorax and pulmonary interstitial emphysema were also reduced (ORs 0.60, 95% CI: 0.40 – 0.88 and 0.51, 95% CI: 0.32 – 0.79, respectively).[50]

From 1997, the Cochrane Library started publishing reviews on surfactant, including nine RCTs in their first review which focused on animal-derived, natural surfactants. This found a relative risk (RR) of death of 0.60 (95% CI: 0.47 – 0.77) as well as reductions in respiratory morbidity.[51] It was clear by now that, just like antenatal steroids, surfactant improved outcomes for extremely premature babies, and its use had become the standard for care.

2.3 Investigating extreme prematurity

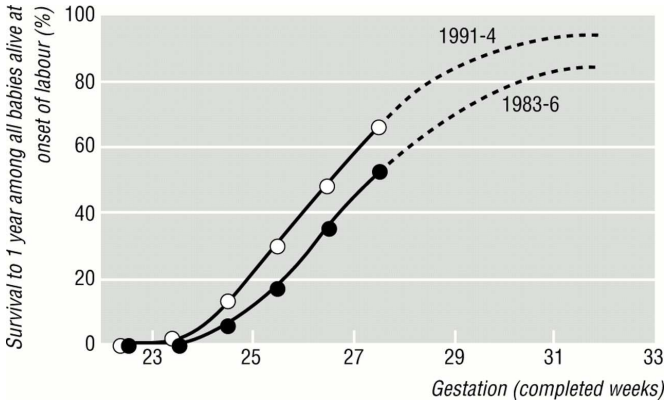
Several studies were carried out during this period that sought to determine the outcome for babies born extremely prematurely. The first of these was the Trent Neonatal Survey (TNS) of January 1991 to December 1993, which identified 1,684 live born babies under 32 weeks gestation.[24] Survival is shown in figure 2.2: fewer than half of those born between 22 and 26 weeks completed gestational age survived. The study was possible as the then Trent Regional Health Authority provided funding for external observers to visit each hospital and gather data on all babies born below 33 weeks gestational age or with a birth weight of less than 1500g. However, data on all births in this study were incomplete, as it did not cover stillbirths; information was instead obtained from the Office of Population Censuses and Surveys (OPCS). Due to organisational changes at the OPCS, data were unable to be supplied at that time for 1993,[24] which meant that estimated values had to be used. Furthermore, the legal definition of stillbirth changed in 1992 to additionally include births between 24 and 28 weeks gestation.[52] This may have influenced data recording. Both these factors therefore may have contributed to over-estimating survival.

Figure 2.2: Survival to discharge of live born babies between 20 and 32 weeks gestation in the Trent Region, England, 1991-1993. Reproduced from Archives of Disease in Childhood (Fetal and Neonatal Edition), Bohin et al, 1996,[24] with permission from BMJ Publishing Group Ltd.



While the TNS was able to provide useful information on resource utilisation within the Trent region, there were problems that limited its generalisability. Despite being a

Figure 2.3: Improvements in survival to 1 year of age between 1983-86 and 1991-94 in babies born at < 28 weeks gestational age who were alive at the onset of labour in the Northern Region, England. Reproduced from BMJ, Tin et al, 314:107-111, 1997,[53] with permission from BMJ Publishing Group Ltd.



well-defined geographical region that included a population of 4.6 million with around 63,000 births per annum, few babies were born at extremely low gestation each year.[24] This meant that several years of data needed to be combined for analysis, potentially causing temporal effects to be missed.

A similar study, from the Northern Neonatal Network, demonstrates the same problem.[53] This study collated information from across the area of the former Northern Regional Health Authority on every birth of 22 weeks gestational age and above, together with detailed follow-up at 2 years of age for those born at less than 28 weeks gestation in 1983, 1987 and 1991.[53] The graph in figure 2.3 shows the survival estimates for babies born less than 28 weeks gestation in 1983-86 compared with 1991-94; a clear improvement is seen between the two epochs studied.

It should be noted, however, that the survival estimates from the Trent [24] and Northern [53] studies are *not* comparable. This is because of the baseline populations used: in the former, survival is based upon those babies who are born alive,[24] whereas in the latter, the denominator is considered to be those pregnancies where there was evidence the fetus was alive at the beginning of labour care and monitoring.[53]

2.3.1 The EPICure study

To overcome the issues caused by the small number of babies born within each geographical region at extremely low gestations when assessing outcomes, the EPICure

study was set up.[54] This prospectively collected cohort included all births from 20 to 25 completed weeks of gestation between 1st March and 31st December 1995 in the whole of the United Kingdom and Republic of Ireland, and identified a total of 4,004 births. As in both the TNS [1] and the Northern Neonatal Network study,[53] babies born with congenitally lethal abnormalities were included.[54]

Of the 843 babies subsequently admitted to neonatal intensive care, 32 were excluded as they were determined to be $> 25^{+6}$ weeks gestation.[54] Gestational age was assessed by the study investigators according to the estimated date of delivery calculated from the mother's last menstrual period or, if there was a discrepancy of > 2 weeks, an obstetric ultra-sound scan performed before 20 weeks of pregnancy. For 29 babies, gestational age was based only on the paediatrician's best estimate; three of these were thought to be $> 25^{+6}$ weeks and were thus excluded. [54]

There was one major flaw in the study design that became apparent: collection of data for fetuses that were *not* born alive was extremely limited, with only basic data collected (date of birth, sex, estimated gestational age and whether there were any signs of life present). Furthermore, the lack of consistency between perinatal estimates of gestational age identified in those babies admitted into NICUs was a problem that had not been anticipated. This meant it was only possible to define accurately the survival and long-term morbidity of babies in relation to the populations of those babies born alive or admitted to neonatal intensive care, rather than for all deliveries as had originally been planned – as the limited data that had been collected on all births were insufficient to confirm gestational age. The lack of data also meant it was not possible to examine the effects of obstetric interventions or of antenatal decision-making.

The study investigators were, however, extremely rigorous with the data that were collected, with an investigator visiting study sites that returned incomplete data. All data were single-entered into a computer database, with a 10% sample being subject to double entry to ensure accuracy. This noted 17 mistakes amongst 15,280 items – an error rate of $< 0.01\%$.[54] Additionally, seemingly implausible data were manually checked with the reporting hospital.[54] It is therefore unlikely that there was serious misclassification of a non-differential nature and, while there may have been some differential misclassification (related to different interpretations between the reporting

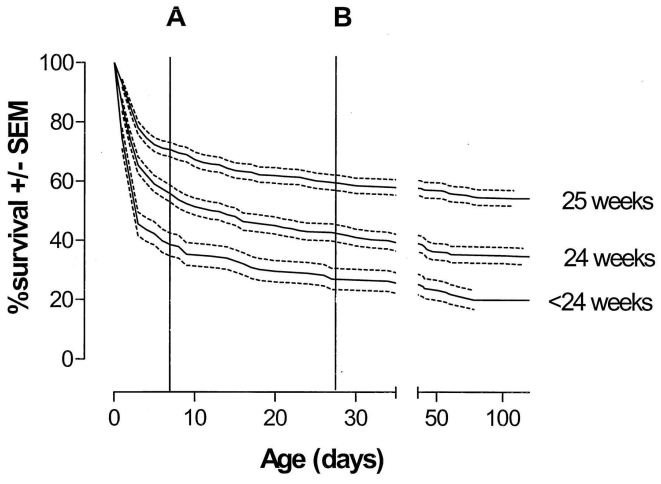
[1] Personal communication with Professor Elizabeth Draper, University of Leicester

hospitals), this too is unlikely to have had a large effect due to precise definitions provided on the data collection forms. A further strength was the well-defined geographical area that the EPICure study included: this was restricted by both physical as well as political boundaries, meaning that women were unable to seek care outside the study catchment area.

It is of particular interest to note the distribution of hospitals in which these babies were born and looked after. Of the 276 hospitals with maternity facilities in the UK and Republic of Ireland, 266 recorded births at less than 26 weeks gestation. Babies were admitted to the neonatal unit in 71% (190) of hospitals but, from 53 of these, babies were transferred elsewhere within the first 24 hours of life for ongoing care. Consequently, only 137 (51%) neonatal intensive care units were involved in looking after babies longer-term.[54] However, a further criticism is that geographical differences do not seem to have been considered in the analyses conducted at that time.

Of the 811 babies included in the study, 574 (70.7%) were the result of singleton pregnancies, and 14 (1.7%) were diagnosed with potentially lethal congenital abnormalities, of whom three survived (one each with tracheo-esophageal fistula, imperforate anus and meconium ileus);[54] 666 babies between 22 and 25 completed weeks of gestation were born in England.[55] Survival curves are shown in figure 2.4.

Figure 2.4: Survival curves for babies admitted to neonatal intensive care in the EPICure study (1995; n=811; 156 <24 weeks, 298 at 24 weeks and 357 at 25 weeks gestation).[A: 7 days; B: 28 days]. Reproduced with permission from Pediatrics, volume 106, pages 659-671, copyright ©2000 by the American Academy of Pediatrics [54]



Since the original study, the EPICure cohort survivors have been followed up at 2.5,[56] 6,[57] 11 [58] and 16 years of age, with a further 19 year follow-up currently in progress. Full survival statistics are shown in table 2.2; this highlights the discrepancies between the gestational age in use at birth and that calculated following admission to neonatal intensive care when it could be determined with more certainty. Two children who were severely disabled died between the 6 and 11 year follow-ups, and one further child was identified who had been missed from the initial cohort, hence no birth data were available for this subject.[58]

Table 2.2: Survival in the EPICure cohort up to 6 years of age. Data are collated from published reports by Costeloe, [54] Wood,[56] Marlow [57] and Johnson[58]. N/A: no published data available; ToP: terminations of pregnancy.

	Gestational age (weeks)					**Total**
	21	22	23	24	25	
All births (including ToP)	N/A	N/A	N/A	N/A	N/A	N/A
All births (excluding ToP) [a]	N/A	622	N/A	N/A	N/A	4004
Live births [a]	104	138	241	382	424	1289
Admissions [a]	3	17	121	313	389	843
Admissions (recalculated) [b]	3	22	131	298	357	811
28-day survival	0	3	38	129	209	379
Survival at EDD	0	2	28	107	192	329
Survival at discharge	0	2	26	100	186	314
1 year survival	0	2	25	98	183	308
6 year survival	0	2	25	98	183	308
11 year survival	0	N/A	N/A	N/A	N/A	307

[a] Prior to admission, gestational age was assigned as the best working obstetric estimate.

[b] Detailed data were available for babies admitted into intensive care hence gestational ages were able to be checked and corrected where necessary.

2.3.2 The European context

EPIPAGE – "étude épidémiologique sur les petits âge gestationnels" (*an epidemiological study of very preterm infants*) – was a study very similar to EPICure that was carried out in nine regions of France, approximating to one third of the French population. The study sought to examine the short and medium term effects of prematurity on mortality and survival in France, and the impact of obstetric and neonatal practices on those babies born extremely prematurely, as well as describing the longer term related outcomes.[59] Data were gathered on three groups of babies. The primary

group included all babies born at < 33 weeks gestational age during 1997 – except for two babies of 32 weeks born in a private establishment who were excluded from the results [14] – with continued collection over two years for births at 22-26 weeks completed gestational age. A second group of babies included those born at 33 or 34 weeks gestation during either April or October, while the final group was collected over a period of a week and comprised live born term babies of 39 or 40 weeks gestation.[59] Follow-up was planned to five years of age (to be completed by 2003), although this was mainly by postal questionnaire to the parents, with only the 5 year follow-up to include clinical examination by an independent observer.[59]

In total, there were 4,395 births between 22 and 32 weeks gestational age, of which 722 were late terminations; notably, termination of pregnancy was permitted at any gestational age in France at this time.[14] It is not possible to compare this with EPICure (which did not collect any information on terminations of pregnancy). However, it *is* possible to compare other data for gestational ages below 26 weeks. The proportions of all births that were *not* terminations of pregnancy that were still born were similar – 67.8% in EPICure (2,715 of 4,004 births) compared with 62.9% (351 of 558 births at 25 weeks gestation or less) in EPIPAGE – "étude épidémiologique sur les petits âge gestationnels" (*an epidemiological study of very preterm infants*) – – as was the proportion of babies who were admitted into intensive care (22.8% in EPIPAGE and either 20.3% or 21.6%, depending upon whether the recalculated gestational ages are used or not, in EPICure). When looking at the proportion of live births that were admitted into intensive care, the figures were also similar: 61.4% (127 of 207 babies) in EPIPAGE compared with either 65.4% (using original estimates) or 62.9% (using the recalculated gestational ages for admitted babies) of live births in EPICure.

Overall survival to hospital discharge in the EPIPAGE study was 85% of live births, but this hides a marked variation from 31% at 24 weeks gestation – with no survival below that – to 89% and above from 29 weeks onwards.[14] Indeed, the published survival rates to hospital discharge for this study were very similar to those in the EPICure cohort at 24 and 25 weeks gestation: 33.6% and 52.1% for 24 and 25 weeks respectively for those born in Great Britain and Ireland in 1995; [54] 31% (13/41 live births) and 50% (59/119) respectively for the EPIPAGE study, using data from 1997 only.[14] Available survival data from the 1997 cohort are shown in table 2.3.

Table 2.3: Survival until discharge in the EPIPAGE cohort of babies born in 1997. Data are collated from published reports by Larroque [14] and Bodeau-Livinec.[60] ToP: terminations of pregnancy

	Gestational age (weeks)					Total
	22	23	24	25	26	
All births (including ToP)	204	284	211	285	312	1296
All births (excluding ToP)	102	137	115	204	239	797
Live births	16	30	42	119	158	365
Admissions	0	6	27	94	141	268
Survival at discharge	0	0	13	59	89	161

2.4 Reorganisation into networks

By the beginning of the new century, evidence was accruing that coordination of neonatal care within regions had an impact on outcomes – particularly mortality, the easiest outcome to measure accurately. Regionalisation had been a long standing goal in the United States [25], but the economic and social contexts surrounding healthcare were markedly different from Europe and made extrapolation of the findings difficult.[61] Nevertheless, some European countries had started to implement regionalisation of care by the late 1990s.[61] This was particularly evident in France and Poland, countries where there were nationally-mandated guidelines relating to the care of women with high-risk pregnancies.[61]

2.4.1 Evidence for change

Evidence for improved neonatal outcomes for babies born in level 3 neonatal units compared with those born in units of lower levels was summarised in a meta-analysis in 2010.[62] This included articles reporting randomised controlled trials, cohort studies (prospective and retrospective) and case-control studies, which compared the level of unit of birth to neonatal mortality and survival to hospital discharge, and that were published between 1975 and 2010. Studies were divided between those examining the effect on babies with very low birth weight ($\leq$ 1500g) and those looking at babies born very prematurely ($\leq$ 32 weeks gestation).[62] The authors identified 41 articles, 37 of which examined regionalisation in relation to LBW babies and four in relation to prematurity. Two studies were conducted in England - one from Leeds published in

1986 and the other from London published in 1983. Both examined differences in pre-discharge mortality between babies who were transferred either before or after delivery rather than between different levels of hospital.[62]

Outcome for low birth weight

The total population reported in all studies – the majority coming from the United States (22 of 37 papers) – of VLBW babies was 104,944, and provided a combined odds ratio of 1.62 (95% CI: 1.44 – 1.83) for the occurrence of death prior to discharge in units of lower care levels compared to tertiary neonatal units.[62] There was no important impact on the effect when studies were limited to only those considered of "high" or "adequate" quality. When data were restricted to those available for ELBW babies only, the effect was even more pronounced, with an OR of 1.80 (95% CI: 1.31 – 2.46).[62] There was marked statistical heterogeneity for both the complete ($Q=153.14$, $p<0.001$) and restricted analyses ($Q=28.40$, $p<0.001$), indicating that there were important differences in how the combined studies were carried out.[62]

Outcome for extreme prematurity

Four studies were found that investigated the impact on populations of premature babies (of less than or equal to 32 weeks) of the unit designation for the units where babies were born. An odds ratio of 1.42 (95% CI: 1.06 – 1.88) was found for the three adequate or high-quality studies included for an increased risk of death in lower volume hospitals. Addition of data from the fourth study caused an increase in the OR to 1.55 (95% CI: 1.21 – 1.98) with no evidence of study heterogeneity ($p=0.28$).[62]

It is worth noting that the increased effect seen in the ELBW babies in this study may have been a reflection of the fact that birth weight only partially reflects maturity. Thus, there would have been effect dilution by the greater numbers of gestationally older babies included in the wider range of weights. Similarly, there are marked differences in maturity and, consequently, care that is required for babies born between 24 and 32 weeks gestation. The effect seen for premature babies may therefore also have suffered from effect dilution if gestational age was categorised inappropriately.

2.4.2 Contemporary evidence

The data reported in the meta-analysis, however, are predominantly based on recent studies, with only three of nine of the included high or adequate quality studies relating to very low birth weight babies – and none of the four studies relating to very preterm babies – having been published prior to 2000.[62] Instead, early evidence of the effects of neonatal health care organisation on preterm babies existed in the form of comparative studies between populations. An example of this was a comparison between Australia and the UK – specifically, Scotland, although data on England and Wales were also presented.[63]

The paper presented data on two studies from 1993-94, at which time Australian neonatal services were concentrated around 23 level III units with an average of more than 10,000 births per centre. By contrast, in Scotland, there were 17 level II or III units with only approximately 4000 births per unit.[63]

Recommendations for training requirements and staffing provision differed between the two countries, with the Australian system being more stringent than the UK system.[63] The first study investigated differences in pre-discharge hospital mortality from an invited selection of hospitals in both countries. The second used available population data to compare the neonatal mortality rate for babies weighing from 500-1499g at birth.[63]. Despite the non-random nature of the first study – meaning, as the authors highlighted, that "selection bias could not be excluded" – the results of both studies were consistent, showing improved survival in the Australian neonatal units compared to those in Scotland.[63] The authors argued that this was attributable to the differences "in the organisation of neonatal services, such as greater specialisation of medical and nursing staff, or higher nurse/patient ratios, or with differences in speed of implementation of effective treatment" rather than "genetic, social or environmental factors."[63] However, the only variables related to organisation of care that were included in the models were the unit designation, number of intensive care cots and the country of study,[63] hence it could be argued this is a tenuous conclusion.

Another study examined the differences in perinatal outcomes in Denmark and the UK region of Trent in 1994-95.[64] This looked specifically at babies of less than 28 weeks or 1000 grams, using data collected in Denmark for a study by the Danish Paediatric Society and in the UK from the TNS, supplemented by the Confidential Enquiry

into Stillbirths and Deaths in Infancy (CESDI).[64] Again, differences were highlighted between the organisation of health services, with supposedly a more hierarchical provision in Trent than in Denmark; however, as table 2.4 shows, the situation was more complicated than this.

Table 2.4: Hospital numbers in Denmark and Trent, by annual number of births. In eastern Denmark, all women with threatened premature labour are transferred to a single university hospital; in the western part of the country, all care is provided at the mother's local perinatal unit. Adapted from Field et al.[64]

Number of deliveries per year	**Denmark**		**Trent, UK**
	West	**East**	
< 1000	15	4	3
1,000-3,000	12	12	13 [a]
> 3000	3	2	
Total number of hospitals	30	18	16
Population	5.2 million		4.6 million

[a] Thirteen hospitals in Trent had between 1,000 and 6,000 births per annum.

The results showed that Trent had a smaller population but a similar total fertility rate to Denmark (1.80 v 1.76 respectively), and an overall higher rate of live births meeting the study criteria (birth rate of babies < 28 weeks or < 1000g per 1000 live births in Trent: 4.88, 95% CI: 4.50 – 5.29; Denmark: 3.31, 95% CI: 3.02 – 3.63) as well as a higher admission rate to intensive care (Trent: 4.17 per 1000 live births, 95% CI: 3.82 – 4.56; Denmark: 3.01, 95% CI: 2.74 – 3.31).[64] Postnatal events – use of surfactant, ventilatory support and oxygen dependency at 28 days of life – all showed extremely important ($p < 0.0001$) differences with much higher rates in Trent. Additionally, the mortality rate was higher in Trent compared to Denmark (42.3% v. 35.0%, $p=0.0238$).[64]

One of the potential influences considered by the authors for these differences was the centralisation of the health services into hierarchical networks. However, this was not thought to be an important influence – particularly as the worse outcomes seen in the Trent region were associated with the (supposedly) more centralised service. In fact, given the organisational differences between the eastern and western parts of Denmark, the influence of health care organisation cannot readily be assessed from the

data presented, and the different regions of Denmark might themselves have been a better focus for studying this question.

2.4.3 Implementation of managed care networks

In 2003, the Department of Health published a plan to introduce managed clinical networks (MCNs) for neonatal intensive care [65] based upon recommendations from BAPM.[19] As well as the accumulating evidence related to neonatal outcomes, workforce changes with the introduction of the European Working Time Directive (EWTD), new contracts for consultants and changes in the training of junior doctors [3] presaged the need for change: it would no longer be feasible to staff a multitude of smaller neonatal units.

The implementation of these clinical networks took place between 2004 and 2007,[20] although not all regions managed to achieve the same level of service provision. This was, in part, due to staffing shortages and problems with financial management (£70 million had been allocated by the government, but only 34% of this was actually spent[66]). Moreover, a lack of standardisation in data collection which prevented meaningful comparisons,[20] thus creating difficulty in knowing the extent of change that had occurred. The result of this was that, from a starting number of 24 networks, by 2007 only 23 existed due to the merger of the South Central and Thames Valley networks.[67]

2.5 EPICure 2

It was during this period of change that a second EPICure study was set up, although with some important differences. First, EPICure 2 only covered those babies born in England to mothers who were usually resident in England,[55] as this was where the majority of births had occurred during the first study and there had been no important differences between the regions.[55] Secondly, the gestational age range was expanded to include those of less than 27 weeks – that is, babies born at up to 26 weeks and six days gestational age were eligible for inclusion.[68] This was in order to provide a comparison group of babies who normally received full intensive care and also because of a paucity of data collection to date for this gestational age group.[55, 68]

The EPICure 2 study ran for the complete calender year of 2006, and study objectives (quoted verbatim from the study protocol[68]) were:

1. To quantify survival to discharge from hospital and major complications of extreme prematurity (Retinopathy of Prematurity, intracerebral haemorrhage, chronic lung disease of prematurity, laparotomy for NEC and/or perforation) in babies born before 27 completed weeks, by week of gestational age.
2. To compare outcomes for babies born $<$ 26 weeks gestational age in 2006 with those born in 1995.
3. To determine how antenatal complications and management in labour influence condition at birth in those born before 27 completed weeks gestational age.
4. To establish whether predictive values for death and adverse neonatal outcomes in the 1995 cohort still hold and to identify other factors that determine these outcomes.
5. To quantify changes since 1995 in the proportion of total births that are born alive and the proportion of livebirths admitted for intensive care at these gestational ages and to record significant changes in neonatal interventions.

Unlike the previous study, gestational age in the EPICure 2 was determined according to the earliest available antenatal ultrasound scan or, if no scans were performed, according to maternal LMP. Only in the absence of a certain LMP was clinical estimation used (78 cases of a total 3,133 confirmed to be between 22^{+0} and 26^{+6} weeks gestation). This was in accordance with national guidance of the time.[69] Two thousand, three hundred and twenty-six fetuses were alive at the onset of labour of whom 2,034 were born alive.[55] Three data collection forms were utilised to record details of the births; all three are available on the EPICure website[1]:

labour ward log sheet: a record of all births occurring to women of greater than 20 weeks gestation on obstetric or gynaecological wards;

[1] `http://www.epicure.ac.uk/publications/study-documents-for-professionals/`. Accessed 4th August, 2014.

perinatal notification form (PN:E2): a replacement for the standard CEMACH form (used to report perinatal deaths) with questions encompassing the standardised data set plus additional information related to obstetric management and complications, used for each individual birth whether live or stillborn;

case record form (CRF): a detailed data collection form for all babies admitted into neonatal intensive care units.

The PN:E2 form particularly was an example where the study investigators had learned from their experience with the first study: by gathering data on *all* births, it became possible to better estimate the impact of obstetric interventions on those babies who were born between 22 and 26 completed weeks of gestation.

2.5.1 Comparisons with the first EPICure study

A first concern for the investigators, though, was to describe the 2006 cohort and to examine any differences from the 1995 cohort that were found. As in the first study, there were important increases in survival to 28 days and to discharge with each increasing week of gestation ($p<0.001$ for both).[55] Among those admitted into intensive care, 83% were exposed to antenatal steroids, 14% had an admission temperature below 35°C, and 18% were transferred within 24 hours of birth; all of these measures showed very important improvements with increasing gestational age ($p<0.001$).[55] Moreover, 99% of babies received surfactant (with no differences evident by gestational age). These figures compared with figures from the entire 1995 cohort of admitted babies of 64.9% having been exposed to steroids antenatally, 40.4% having an admission temperature below 35°, 14.1% having been transferred within 24 hours of birth, and 84.1% of babies having received surfactant at any point.[54]

The suggestion from these two overlapping but different cohorts that care had improved between 1995 and 2006 (i.e. using the complete data sets from the EPICure and EPICure 2 cohorts) was able to be examined in a planned comparison by including only those babies born at less than 26 weeks in England between March and December in either epoch. This analysis identified 666 admissions in 1995 compared with 959 in 2006 – an increase of 44%, for which there were was no obvious explanation to the investigators and no further data with which to investigate.[55] Overall survival to discharge increased between the two epochs from 40% to 53% (95% CI: 48% – 58%), with

important increases at 24 (by 12%, 95% CI: 4% – 20%) and 25 (16%, 95% CI: 9% – 23%) weeks, but without reaching statistical significance at 23 weeks (9.5%, 95% CI: 0.1% – 19%).[55] Comparative data for babies born at 22 weeks in England in 1994 are not provided, but as only 3 of 478 births at 22 weeks in 2006 survived to discharge, it is unlikely that there were sufficient numbers to be statistically significant.[55]

2.5.2 Unit level effects

Additionally, a survey of neonatal units, heavily based on the original survey performed as part of the UK Neonatal Staffing Study,[28] was carried out by the EPICure investigators. This was called the Unit Profile Study (UPS) and, as well as staffing and activity data similar to that collected previously, additional information was gathered on obstetric and neonatal unit policies relating to the management of extremely preterm labour (the questionnaire is shown in appendix D).[68]

The combination of these two data sources – individual level information about births, and unit level data about perinatal centres – permitted the investigation of outcomes in relation to the place of birth as well as other organisational factors such as antenatal or postnatal transfer, or volume of activity. For this, the eligible population was restricted to those pregnancies with a live fetus at the time of the delivery admission to hospital. Unit levels were self-designated, breaking down into 52 level 1, 84 level 2 and 46 level 3 units, with 244, 829 and 1387 (440 after antenatal transfer) births occurring in each level, respectively.[70] As with the initial EPICure 2 paper by Costeloe et al,[55] differences in survival were most evident during the earliest points of perinatal care, with improved survival during labour, in the delivery room and during the first seven days of life in level 3 compared with level 2 units.[70] After adjustment for gestational age and birth weight, these findings persisted for early neonatal – but not antenatal – survival, leading to an improved overall odds ratio of 0.73, 95% CI: 0.59 – 0.90. There were, however, no differences in morbidity-free survival between units in the place of birth analysis.[70]

The analysis was also extended to look at differences by volume of work, defined according to the number of admissions per year and the number of consultants with more than half of their clinical time dedicated to neonatology.[70] This demonstrated a lower risk of mortality (adjusted OR 0.68, 95% CI: 0.52 – 0.89) for babies born in "high activity" hospitals – with $\geq$ 2000 days of respiratory support provided per year

and $\geq$ 4 neonatal consultants – compared with medium activity hospitals (500-1999 days respiratory support per year and one consultant).[70]

2.6 "Modern" neonatology

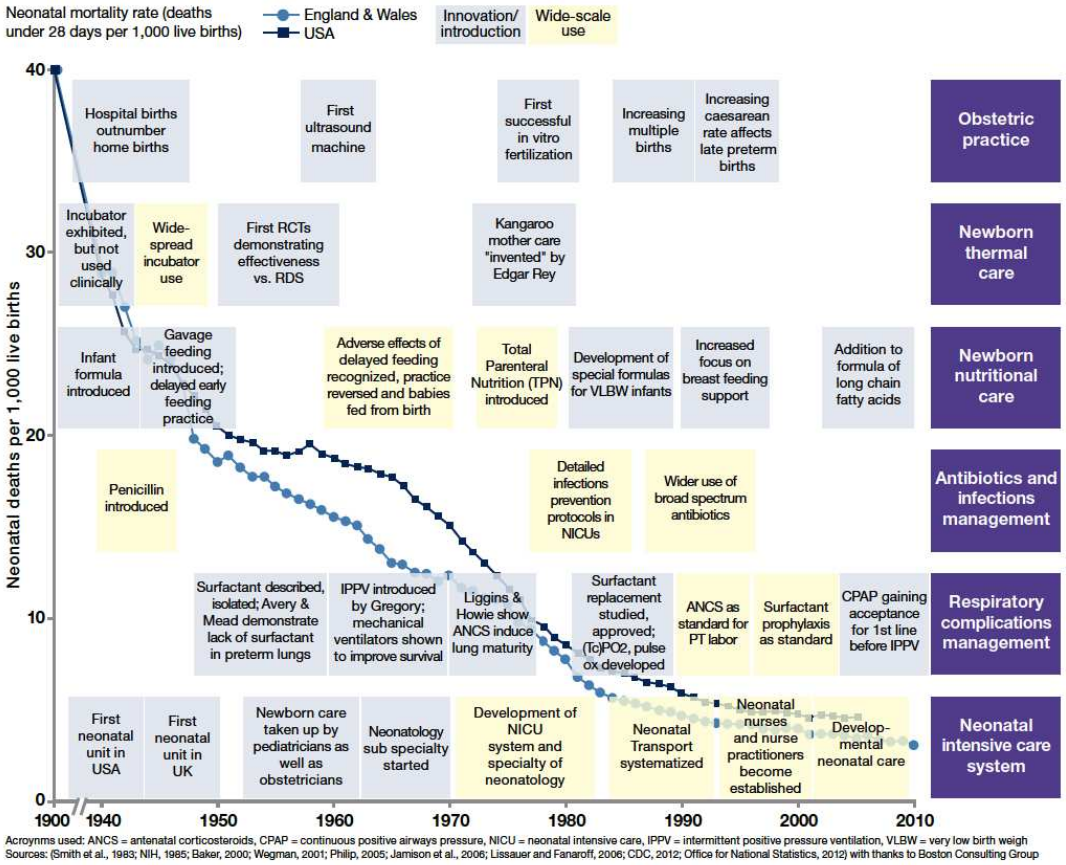
A graphic published by the World Health Organization (WHO) in *Born Too Soon: The Global Action Report on Preterm Birth*[7] illustrated that, by the mid-2000s, most of the large gains in neonatal mortality that were possible had already been made (see figure 2.5). Instead, attention was focusing on fine-tuning outcomes for specific groups of affected babies. An example of this was the introduction of therapeutic hypothermia as a standard of care to be provided by network NICUs for term babies who suffered perinatal asphyxial insults.[71] This followed reporting of the TOBY trial, which showed no difference in the rates of death or severe disability, but important differences in the rates of mild disability.[72]

The other large group of babies on whom attention was focused were those who were born extremely prematurely. This was largely due to the persisting morbidity profiles identified by longer term follow-up of the original EPICure cohort[73–77] as well as an acknowledgement that these babies had the longest durations of stay[55] and the largest financial impact on society.[78–80] Consequently, as well as continuing research into the effects of managed clinical networks in England and a number of other countries conducting large cohort studies of extremely premature births, a body of work was also being constructed related to perinatal ethical decision-making. All this was accompanied by expansions in computing that meant large data bases could be established and maintained across broad geographical areas – the era of "big data" was arriving.

2.6.1 Updated evidence for regionalisation

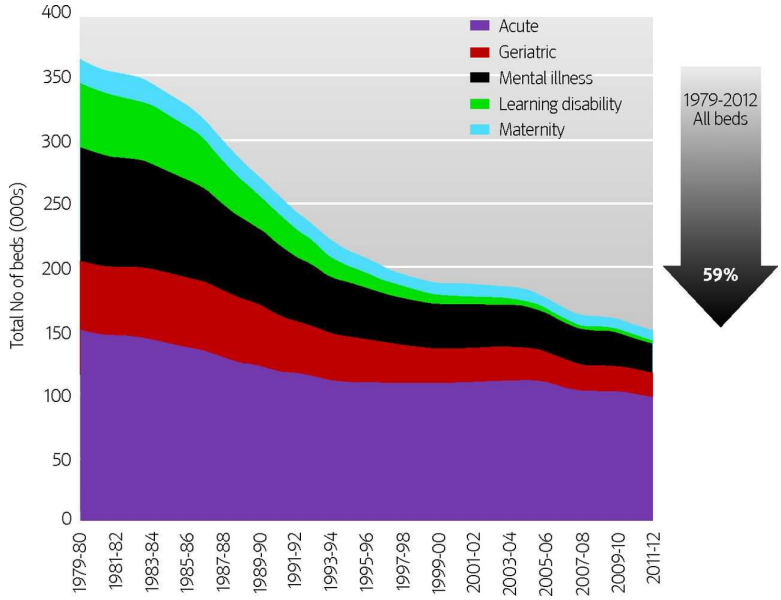
The introduction of managed clinical networks in England in 2003 coincided with an upturn in National Health Service (NHS) funding, with an increase from 7.0% GDP in 2000 to 9.4% in 2011.[81] More importantly, there was a general increase in activity: Accident and Emergency (A&E, or "casualty") department attendances increased by 50% between 2002/03 and 2010/11 [81] and births increased from 596,122 in 2002 to 708,711 in 2008.[82] In contrast, however, the number of beds available has fallen

Figure 2.5: Neonatal mortality compared with a timeline of innovations in neonatology in the developed world. Reprinted from “Born Too Soon: The Global Action Report on Preterm Birth”, the World Health Organization, page 87, 2012.[7]



consistently since the late 1970s.[83] This is particularly apparent for acute medical beds but also affects maternity care, as shown in figure 2.6 (adapted from Appleby).[83]

Figure 2.6: Changes in the number of NHS inpatient beds between 1979 and 2012 in England. Reproduced from BMJ, Appleby, 346:f1563, 2013,[83] with permission from BMJ Publishing Group Ltd.



With respect to neonatology, the implementation of MCNs suffered from a major flaw: they were *neonatal* networks that were established, rather than complete perinatal networks that also encompassed obstetric care.[3, 84] Furthermore, there was a lack of information with which to monitor the changes – although network implementation *did* act as a driver for new data collection systems.[3, 85]

Consequently, despite a number of studies that within the English context have examined outcomes in relation to hospital volume following the introduction of MCNs,[70, 86] only one study has examined a "before and after" effect.[85] This study compared outcomes for 27 or 28 weeks completed gestational age babies live born between 1st September 1998 and 31st August 2000 in England, Wales and Northern Ireland with those live born English babies who were admitted into neonatal intensive care in the calendar years of 2009 and 2010. Babies born in the latter period were more likely to be multiples ($p=0.03$) and to be of higher birth weight ($p<0.001$).[85] They were also more likely to be born in a hospital that provided $\geq$ 2000 days of intensive care per

year (18% and 49% for the two study periods, respectively, OR 4.30 with 95% CI: 3.83 – 4.82, p<0.001).[85] Interestingly, when clustering by neonatal network was accounted for in the analysis, the combined odds ratio remained similar at 3.25 (95% CI: 2.02 – 5.03) but there was significant heterogeneity across networks (p<0.001); these results were not published.[85]

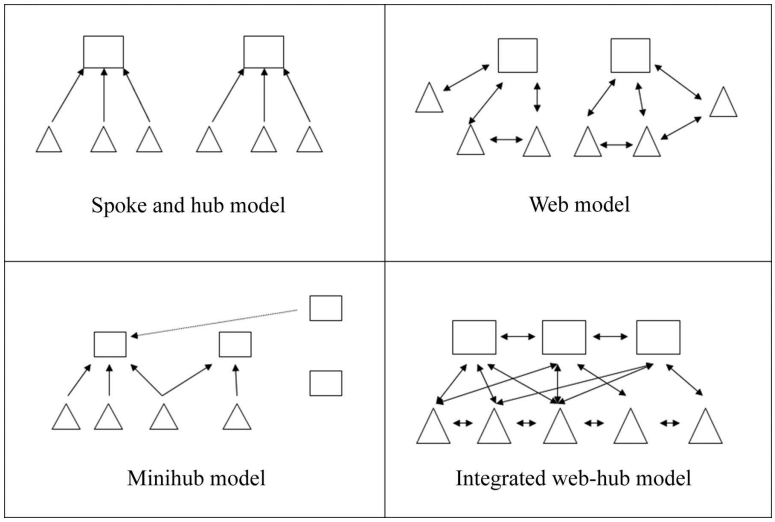
The same study went onto examine survival to 28 days of age. This showed 2,859 survivors of 3,248 (88%) live births in the first study period, and 2,677 of 2,919 (reported as 94% – 60 subjects with missing data were excluded) babies admitted into a neonatal unit, giving a p-value of less than 0.001.[85] These comparisons are invalid, however: although the data sets were both restricted to English-born babies, the use of different baseline populations (live births compared to neonatal admissions) precludes meaningful interpretation as many babies do not survive beyond the delivery room, thus introducing a major selection bias into the comparison.

Indeed, this study features all three of the “deficiencies in the literature” identified by Lorch and colleagues in a systematic review about “the Regionalization of Pediatric Health Care”.[87] As well as having a focus restricted to postnatal transfer status and survival, the before and after investigation of 27-28 week babies only accounts for differences between hospitals in terms of the volume of intensive care days provided and does not provide any adjustment for differences between hospitals in disease severity.[85, 87] Even more importantly, it focuses on 27-28 week gestation babies, with no discrimination between the gestational age group categories (due to a lack of data for the first study period) – when, in reality, it was only babies of 27 weeks or younger that were recommended to be cared for in level 3 units.[65]

The “Regionalization” review is, in itself, a notable paper albeit clearly influenced by the privatised health care setting in which the authors work.[87] First, it provides a conceptual framework that may be used to understand different forms of networked health care – varying from a “hub and spoke” model through to a “web” model, encompassing both “minihub” and “integrated web-hub” models, as shown in figure 2.7 [87] – as well as both the factors that may influence network characteristics and also the possible mechanisms through which networked care may improve patient outcomes.[87]

Secondly, the review presents a number of papers relating to both the provision of neonatal and paediatric intensive care, identifying deficiencies but also proposing recommendations of how these may be overcome in future work.[87] One suggested

Figure 2.7: Network models: possible configurations of neonatal units within a managed clinical network. Boxes represent regional (equivalent to level 3/network neonatal intensive care units in England) centres, triangles represent district general hospitals (level 1 or 2 equivalents), and arrows indicate the direction of patient flows between hospitals. Adapted from Lorch et al's "Regionalization of Pediatric Health Care", reproduced with permission from Pediatrics, 126:1182-1190, 2010.[87]



approach is to use a quasi-experimental design known as an "instrument variable (IV) technique".[87] This is a statistical method that may be used with observational data to effect a pseudo-randomisation based upon an "instrument" that is closely related to the exposure but unrelated to the outcome.[86, 88] Indeed, the better known *intention to treat* analysis is a modified form of the IV technique, using the *intention* as the instrument.[88]

The IV technique is employed in a study that examined unit designation (tertiary or non-tertiary unit) and volume (top quarter of units ranked according to annual number of days care provided to premature babies) against both mortality and a range of morbidities: ROP, necrotising enterocolitis (NEC), CLD.[86] This found an important decrease in mortality for babies born in a high-volume hospital (OR 0.70, 95% CI: 0.53 – 0.92 for all babies $\leq$ 32 weeks completed gestational age), consistent with the results of the "standard" regression analysis that demonstrated an OR of 0.73 (95% CI: 0.56 – 0.95).[86] The major problem again with this study is that the baseline population consisted only of babies born and then admitted into intensive care at each of the hospitals, rather than also including those who died.

2.6.2 Contemporary studies

Further information about the provision of neonatal care in other countries comes from a detailed report prepared in 2007 for the English National Neonatal Audit Programme (NNAP) subtitled “data for international comparisons”.[89] Notably, the report does not cover England, but does include provision by the NHS in Wales, Scotland and Northern Ireland, as well as Sweden, Canada, the United States and Australia. Comparative details of the health care systems in each country are provided as a baseline, with detailed chapters about the neonatal service provision in each country. It was noted that the US did *not* have the best outcomes despite having more resources than Canada or Australia – indeed, Victoria and New South Wales in Australia, and British Columbia in Canada, were highlighted alongside California as having particularly advanced networks.[89] All three of the UK countries were noted to be lacking nursing staff, although this was also a common theme in other countries.[89]

The report made a good attempt at assimilating multiples sources of information across each of five domains: national (or regional) neonatal statistics, organisation and commissioning of services, transport, financial costs and implications, and best practices. Yet, despite the fact that the information searches included the use of medical databases as well as searching grey literature such as government reports or national guidelines and contacting local experts, it is lacking in data that are directly comparable between nations.[89] What it does do well, however, is to describe the structures of different neonatal services in a qualitative fashion, providing a foundation on which to base understanding of other work. For example, we learn that Sweden had 45 neonatal units in 1993, and that three of them closed in 1997 due to decreased financial spending on health (GDP expenditure on health falling from 8.6% to 7.2% in the five-year period to 1995).[89] It is a geographically spread country – the fourth largest land mass in Europe – with a concentration of inhabitants in the southern regions; there is no dedicated neonatal transport service but outcomes were the best of all the countries considered in the report.[89] However, we do not actually find out the size of the country in terms of the number of inhabitants, nor what the annual birth rate is.

This context is useful when it comes to interpreting EXPRESS – the Extremely Preterm Infants in Sweden Study.[90] EXPRESS was conducted between 1st April 2004 and 31st March 2007, a three year period during which there were 305,518 births

in Sweden, of which 1,011 were preterm babies below 27 weeks gestation born to 887 different mothers; we also learn the population is approximately 9 million people.[90] Approximately a third of the EXPRESS babies were still born, and 70% were born in a tertiary level unit. Four hundred and ninety-seven survived to one year of age, although there is a slight discrepancy between papers regarding precisely when some of the deaths occurred: one reports that 58 babies died in the delivery room,[90] whereas another reports 69 deaths prior to admission[91] – presumably 11 babies did not die while being transported from labour ward to the intensive care unit.

EXPRESS is of interest not only because it provides data from a cohort collected contemporaneously to the EPICure 2 cohort, but because the outcomes are so different within a health care system that has many parallels to the English system. Survival, particularly, was much better: EXPRESS demonstrated survival rates at 1 year of 10%, 53%, 67%, 82% and 85% of all live born babies at 22, 23, 24, 25 and 26 weeks completed gestational age, respectively.[90] By comparison, EPICure 2 calculated survival at discharge and found rates of 2%, 19%, 40%, 66% and 77% for the same gestational age categories.[55] Follow-up of the EXPRESS children at 2.5 years of age, however, showed poorer rates of morbidity-free survival than the three year follow-up performed by EPICure 2, with 0% (33% in EPICure 2 – albeit, only one of three babies) at 22 weeks, 30% (53% in EPICure 2) at 23 weeks, 34% (65%) at 24 weeks , 44% (72%) at 25 and 49% (79%) at 26 weeks.[55, 92] Also in contrast to EPICure 2,[70] birth in a Swedish hospital with a level 3 unit was not associated with improved survival after multi-factor adjustment.[90]

Another Scandinavian cohort, from Norway, published their outcomes to discharge[93] and to five years[94, 95] at a similar time, albeit the cohort was set up over a two year period encompassing 1999 and 2000. They reported survival rates using a denominator of all births that were roughly equivalent to survival rates seen in EPICure 2 with a denominator of live births only;[55, 93] 153 of the 174 (88%) deaths prior to neonatal admission were stillbirths.[94]

2.6.3 Ethical guidelines

By definition, results of a five-year follow-up study cannot be published until at least six or seven years after the original intervention, in order to allow time for data collection and analysis, and many studies require longer than that. These necessary delays in

publication, in tandem with the variation in outcomes at the borders of viability, from 22 to 26 completed weeks of gestational age, mean that ethical decision making about the many choices faced by a mother and her offspring is extremely challenging.

A number of studies have sought to explore this further. A comparison across 13 European countries was carried out in 2009 seeking to evaluate national policies related to preterm birth.[96] Data from a variety of different sources were collected, including scientific publications, government and national authority websites, and media reports; however, little was obtained of sufficient quality to permit comparison. The authors concluded with a number of recommendations, the first of which was that premature birth should be recognised as an important health concern.[96] More specific guidance does exist for at least a majority of European countries when it comes to specific management, however: a survey of national guidelines for resuscitation at the borders of viability obtained responses from 19 of 28 countries.[97] This documented a variety of approaches: responses in relation to the resuscitation of 23 or 24 week gestation babies ranged from no resuscitation to full resuscitation, whereas all the responding countries offered resuscitation to babies of 25 weeks or greater (albeit three countries on an individualised basis).[97]

Attitudes among different groups of staff have also been shown to vary. In the winter of 2000, a study carried out in the two major neonatal units in Nottingham achieved a 49% response rate from 142 questionnaires sent out to midwives, obstetricians, neonatal nurses and neonatologists.[98] There were important differences between obstetric and neonatal staff in terms of who would use external fetal monitoring and who would administer antenatal steroids at different gestational ages (the neonatologists were more interventionalist), but not in terms of who would attempt resuscitation of a baby at birth.[98] All groups of staff were more pessimistic about survival rates amongst those admitted into intensive care without significant morbidity than estimates available in the published literature.[99]

In England, a collaborative attempt was made to provide a framework for confronting such ethical dilemmas. The Nuffield Council on Bioethics produced a report in 2006 entitled “Critical care decisions in fetal and neonatal medicine: ethical issues” for which members of the working party were drawn from medical and academic disciplines encompassing ethics and philosophy, law, neonatology, obstetrics and gynaecology, social anthropology, and public policy as well as the Head of External Relations

at the baby charity, Bliss, the Disability Rights Commissioner and a lawyer.[11] The report drew heavily on contemporary English evidence – particularly results from the EPICure study[54] – as well as providing a detailed background on legal and regulatory considerations.[11] Throughout, case examples are used to ensure information is presented contextually; discussion about circumstances that arise in consequence to difficult decisions is also included in the discussion.[11]

A key point highlighted by the Nuffield report is the potential conflict between "sanctity of life" – the viewpoint where any form of life is better than none – and "quality of life" – where some forms of life may not be worth living.[11] This was highlighted early on in the document, with an emboldened conclusion that "the Working Party adopted the concept of 'intolerability"' and continued "although a presumption in favour of life is rightly at the root of all medical care, it cannot be absolute in situations where there are clear indications that the life to be experienced will be an intolerable burden on the child".[11] The Working Party further came to the conclusion that "the moment of birth... [is] the significant moral and legal point of transition for judgements about preserving life." They further provided proposed guidelines for when to offer intensive care for different categories between 22 and 24 weeks gestational age.[11] Possibly the only fault with the report was its length – which at over 250 pages presents a daunting obstacle when the time available to make a decision may be short.

2.6.4 Big data

While the Nuffield report provides excellent guidance, it doesn't help to disentangle the confounding effects of perinatal decision-making in observational studies on cohorts born at the limits of viability. Understanding decision-making processes instead requires improved knowledge of the decisions that are currently made when such ethically challenging circumstances arise. This will help to assess the impact of guidelines and provide an evidence base for future investigations of disease processes, treatment options and decision-making.[11]

"Big data" offers a potential solution here. In the modern era, the term has come to mean extremely large data sets along with the implication that those data can be combined in meaningful ways.[100, 101] This suggests existing data sets could be standardised and linked or merged to provide additional information. Additionally, storage capacity for all types of data increases year on year[102] along with accompanying data

collection. Yet, as the Snowden revelations highlighted, there is a balance between using information for the public good and intruding on individuals' privacy.[103] These concerns are particularly important when applied to health care, due to conflicting desires to preserve information for accurate linkage while maintaining anonymity of the research subjects.[104]

2.7 Neonatal health care data in England

England is fortunate in comparison to countries without a nationalised health care system as the NHS collects large amounts of data about the services it provides and the patients it serves. In understanding care and decision-making at the borders of viability, there are several data bases in existence that could be helpful, thus meaning the requirement for collection of additional information may be limited. There is also the possibility that additional information can be combined with data sets such as EPICure – for example, to further investigate differences that were seen between the two study cohorts.

There are three sources of birth data in England that are collected routinely by central government, stemming from the desire to satisfy two objectives. The first source of data is the Birth Register, collected by the Office for National Statistics (ONS) and based on registrations of births by parents.[6] The second source is the system historically used by midwives to notify health visitors of a new birth;[6] this became known as NHS numbers for babies (NN4B) when it was formalised in 2002 as a registration system for allocating NHS numbers.[105] These systems are designed to facilitate national registration of births and deaths which, together with data about migration, permit calculation of population estimates for inter-census years.[106] The NN4B "message" is now sent to the NHS Health and Social Care Information Centre (HSCIC),[107] – known prior to April 2013 as the NHS Information Centre [108] – which is the governmental authority with responsibility for these data. The NHS HSCIC are also the custodians of the third data source, Hospital Episode Statistics (HES), a data set that records and warehouses data from hospitals. The purpose of these data is primarily administrative but also includes assessment of care delivery and public health planning.[107, 109, 110] A diagrammatic overview of the information flows following a birth in 2005 was constructed by Hilder et al and is shown in figure 2.8.[111]

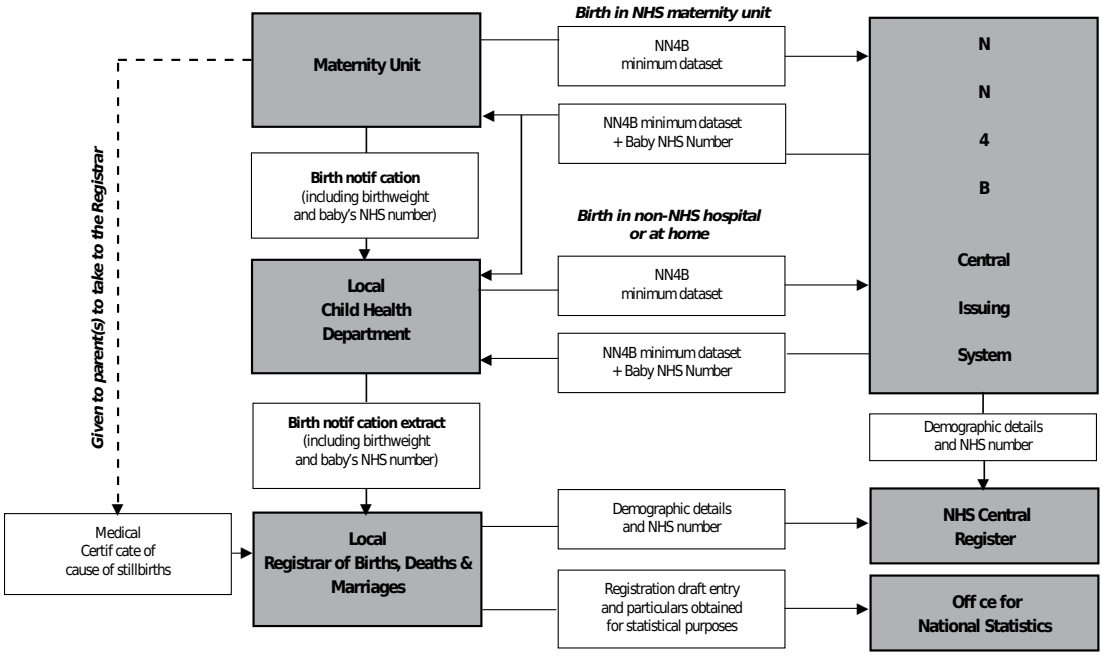


Figure 2.8: Information flows following a birth in 2005: NHS numbers for babies acts as a central issuing system for National Health Service (NHS) numbers which are notified to the local registrar who, in turn, notifies the Office for National Statistics and the NHS central register. Reproduced with permission, Crown Copyright ©2007, from Hilder et al,[111] Health Services Quarterly.

Additional sources of information also exist. For babies who are admitted into neonatal intensive care, there are the National Neonatal Audit Programme (NNAP), run by the Royal College of Paediatrics and Child Health (RCPCH),[112] and the National Neonatal Research Database (NNRD), run by the Neonatal Data Analysis Unit (NDAU), based at Imperial College, London.[113] Both systems use data entered into BadgerNet, a proprietary software available to NHS organisations.[114] Current costs are approximately £25 per admission with an initial set-up fee of around £15,000 plus interfacing costs.[115] There are also the confidential enquiries. These date back to 1952 when the first Confidential Enquiry into Maternal Deaths (CEMD) was launched; forty years later, CESDI was set up.[116] These were both combined into CEMACH in 2003,[116] before this role was taken over in 2012 by Mothers and Babies: Reducing Risk through Audits and Confidential Enquiries across the UK (MBRRACE-UK), coordinated by the National Perinatal Epidemiology Unit (NPEU).[117]

2.7.1 Birth and death registration data

The Births and Deaths Registration Act 1953 made it compulsory for all births to be registered within 42 days by the parents, and for deaths to be registered within five days.[118] In 1986, registrations of neonatal deaths were changed to include spaces for both fetal and neonatal causes of death, and in 1992 the gestational age limit for registrations of stillbirth was lowered from 28 to 24 weeks.[52, 119] This resulted in a change in the perinatal mortality rate from 7.1 to 8.3 per 1000 total births as the definition excluded births below registration cut-off.[52] However, information collected by the registrations of births and deaths is limited: birth weight was first notified by health departments in 1975 [111] but only recorded comprehensively from 1983,[120] gestational age is not recorded at either birth or death, nor is ethnic origin – although mothers' country of birth *is* recorded, and social status may be derived from the parental occupations that are recorded.[120, 121] Only in 2002 was NHS number first recorded consistently alongside the birth registry data.[105, 111]

2.7.2 Hospital Episode Statistics

Hospital Episode Statistics, in contrast to the registration data, are collected directly from secondary care providers in England and Wales in relation to individual patient

'episodes' – so called because they represent the time under the care of a single consultant, thus potentially meaning a patient may experience several episodes during a single hospital admission.[122] For instance, an elderly patient admitted through the emergency department (episode 1) following a fall at home; she or he is then admitted under orthopaedic surgery to repair a fractured hip (episode 2), before requiring transfer to general medical care for treatment of a chest infection acquired during recovery (episode 3); finally, our hypothetical patient is transferred to a rehabilitation ward (episode 4), prior to her or his final discharge home.

HES provide a wide variety of functions, from basic administration and local planning through to population surveillance and development of policy.[122] The HES Data Dictionary for inpatients describes the available variables.[123] Within the maternity data set, each record contains six "baby tails" with detail about each live or still born fetus delivered from that pregnancy.[124] However, there are major concerns expressed regarding the clinical accuracy of collected data, and clinical engagement is consistently identified as an ongoing issue.[110, 125] Particular problems have been highlighted with coding of ethnicity,[126] although this was shown to improve from just under 25% of records containing missing data to fewer than 9% of responses coded as either "not stated" or "not known" between 2004 and 2010.[126] It should be noted that birth records within HES were excluded from this analysis – although they had the highest level of missing data at 12.9%.[107] Marked differences in completion rate were also noted by reporting trust.[107] A similar finding was reported by a study examining the quality of delivery information within the maternity HES data using records from the 2009-10 financial year.[127] They identified 11 out of 136 trusts that had "divergent coding practices", as well as stating that seven trusts had not reported information for any of the variables they were examining in the baby tails.[127]

2.7.3 Other data sources

Other data sources related to neonatal health available for potential linkage come from BadgerNet [114] or the Confidential Enquiries.[116] The National Neonatal Research Database is the primary data set available from the first source and contains data on every baby admitted to one of the contributing neonatal units.[113] In the 2010 report – the most recent year for which country-wide data are reported – data from 155 contributing neonatal units were included;[128] however, 173 units reportedly contributed

data to the NNRD overall during that time period.[86] There are also a number of units who do not (yet?) use the BadgerNet system, hence do not have their data included in the NNRD.[86]

The lack of clarity about the precise population coverage of the NNRD is compounded by a consistent use of babies discharged from neonatal care within a defined period,[128, 129] rather than describing results in terms of those who are admitted or even born in a given time period. These issues lead to concerns that there may be a selection bias inherent within the National Neonatal Research Database based upon both geographical and temporal criteria.

On the other hand, the NNRD does provide a comprehensive database of all babies admitted into intensive care in the participating units, and coverage is almost complete within England.[86] This compares well with data collection from the Vermont-Oxford Neonatal Network (VON), a network of centres around the world that was set up in 1989 and allows for benchmarking for individual hospitals (as long as they are members), as well as enabling quality improvement programmes and research to be organised across multiple sites.[130]

The other source of information is from Confidential Enquiry data. While there are few maternal deaths each year, collecting data on all still births and deaths below a year of age generates substantial amounts of data such that the confidential enquiries usually focus on a specific issue each year.[116] This has been continued by the MBRRACE-UK programme. Indeed, EPICure 2 was run in conjunction with CEMACH, the precursor to MBRRACE-UK, in order to prevent duplication of effort in data collection.[55]

2.7.4 Combining data sources

The potential with these different sources for data linkage has been investigated in recent years. Hospital Episode Statistics were first linked with ONS data in a feasibility study published in 2002.[124] This study found that, in a sample consisting of data for February 1997, only 62% of the registry data were matched by HES – an indication that the registry data were most likely complete, and that there were substantial amounts missing from HES.[124] Substantial variation was identified in matching by region, although the data appeared to be representative, meaning there were no important differences identified between the matched and unmatched records. There was generally

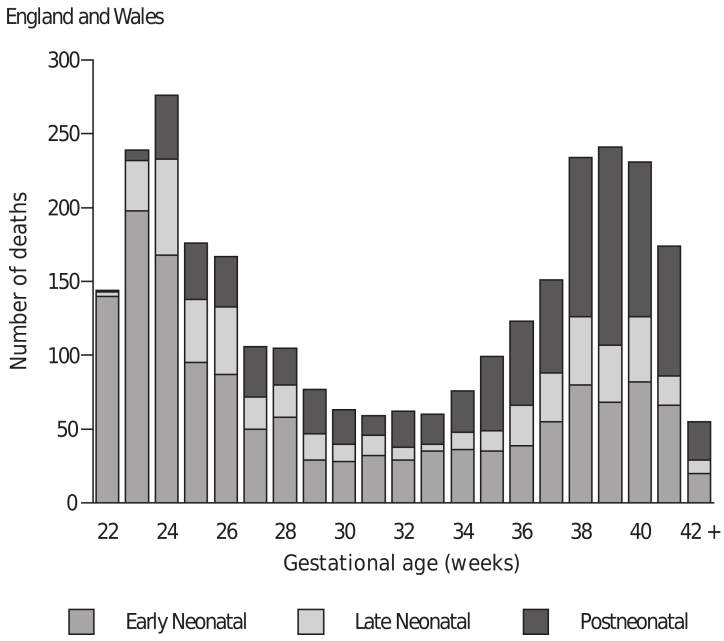
a high level of agreement (above 90%) between the data sources for birth weight, although this was lower (above 80%) for records with a birth weight below 1500g.[124]

Next was a pilot study linking NN4B with the registry data. [111] This examined data from the first three months of 2005, seeking to link NN4B and ONS registry data. Of the two data sources, 99.1% and 99.0% were matched, respectively.[111] Data were concordant for all identical variables in 83.4% of the linked cases, with failure to match post code the most common error in 12.5% of cases.[111] There were also important differences in birth weight which were less apparent by 500g category than when direct comparisons were made. Overall, the discordance rate was reported as 1.5%, and 1.4% when using categories of birth weight.[111] However, the most important differences were found for gestational age – only comparable between these data sources for the still birth data. This showed that in over 80% of cases there was complete agreement, but in 13.5% there were differences in reported gestational age of one week and, in 4.7%, a difference of two or more weeks.[111]

The methods from this pilot study were then used to perform matching of NN4B and registry data from 2005 in order to investigate mortality by gestational age.[105] This identified plausibility concerns with the data, particularly for births registered as being of below 22 weeks gestation: 100 births reportedly had a birth weight over 1000g, hence were excluded from the main analyses.[105] Overall, there was an infant mortality rate (IMR) of 5.0 per 1000 live births. However, there was a marked difference by gestational age in *when* the deaths occurred: extremely preterm babies were more likely to die in the early neonatal period whereas term babies were much more likely to die post-neonatally (figure 2.9).[105]

The same data and methods were also used to investigate the effect of socioeconomic status on infant mortality rate.[131] This was measured using the Carstairs deprivation index for all records, and using the National Statistics (NS) socioeconomic classification (NS-SEC) scheme in sensitivity analyses for the approximately 10% of records with relevant occupational data recorded; birth weight and gestational age were both considered as potential effect modifiers.[131] The authors found there was a generally good correlation between the two socioeconomic indices they had available, and that IMRs did not vary by socioeconomic status for LBW babies but did vary markedly for babies of a normal weight (> 2500g). Conversely, IMR varied by socioeconomic status for preterm babies – although not so much for babies of 37-42 weeks gestation.[131]

Figure 2.9: Infant mortality by gestational age in 2005 using linked data from NHS numbers for babies and Office for National Statistics birth and death registry data. Reproduced with permission, Crown Copyright ©2007, from Moser et al [105], Health Services Quarterly. Early neonatal: < 7 days; late neonatal: 7-28 days; post-neonatal: 29 days – 1 year.



Ethnicity, sex and young maternal age were also independently found to be important predictors of outcome.[131]

Two subsequent studies bear the most relevance to answering questions about increased admissions between the EPICure studies. These examined the results of performing linkage between HES and ONS data for 2005-06 [132] and 2007,[133] with the former reporting data for the same year in which the EPICure 2 cohort was recruited.[55, 132] For both studies, linkage was performed by a commercial third party who provided linkage algorithms that were then modified in conjunction with the study investigators.[132] Despite this, duplicate matches were present in the returned data: for 2006 maternity records (consisting of maternal details with baby tails), of 601,623 records linked between HES and registry/NN4B data, only 584,795 were unique; by comparison, for HES baby records, 674,534 linked records were supplied which was restricted to 538,325 after exclusion of duplicate mothers or babies; the maternal data were used in all analyses.[132]

There was a high proportion of missing information within the data: 25.8% of the 566,313 singleton maternity HES records were missing date of birth, 25.2% birth weight, and 45.1% gestational age.[132] The combined registry/NN4B data were much better in this regard, missing fewer than one percent for each of these variables – and only 189 of the NN4B (and no registry) records were missing date of birth.[132] Where information was present in both HES and the combined data, discordance rates were low, with lack of agreement in 7% of records relating to post code, and just 0.2% disagreement for date of birth.[132] There was also good concordance for birth status, with 114 of 2,514 registry still births recorded as live births in HES, and 748 of 563,799 live births recorded as still births in HES; the major issue was that birth status was not recorded in 27.5% of maternity HES records.[132] For sex, the discrepancy was also small, but 586 males and 2,978 females were recorded incorrectly in HES if the registry data are accepted as the gold standard.[132]

The greatest issues came with the data on birth weight and gestational age, principally due to the data missing from HES,[132] although concordance rates for the latter were higher than in the pilot study linking NN4B with birth registry data.[111] Unfortunately, tabulated data from 2006 are no longer available following a reorganisation of the ONS website.[134] Consequently, we only know that 90% of matched data were concordant for gestational age during that time period; in 2005, the discrepancies were greatest at the extremes – above 42 and below 30 weeks gestational age.[132]

2.8 Data linkage

A key consideration when interpreting studies using linked data is the actual process used for the data linkage, as errors or uncertainties may exert undesirable effects on the results.[104, 135] Broadly, there are two methods that may be used: deterministic, where there is a unique identifier (this may be created using multiple variables), or probabilistic, using a mathematical algorithm to assign a score (or "weight") that can then be used to select matches based on a threshold.[135, 136] Alternatively, some combination of the two methods may be used. Furthermore, data linkage may be applied to either one or more than one data set. In the case of a single source, linkage is used for "deduplication" – in order to identify duplicate entries – else two data sources

are merged for use as a single data set; if there are more than two data sources, each additional source is merged into the master data set sequentially.[135, 137]

2.8.1 Deterministic linkage

Deterministic linkage requires the two data sets being merged (or the single data set being deduplicated) to have a unique identifier that matches between the two sources. Using this variable (or group of variables), matching rows are aligned. Matches may occur on a one-to-one, many-to-one (or one-to-many) or many-to-many basis; only the one-to-one matches are considered as "true matches": the rest require further investigation somehow. This may ultimately be by manual review – but this can be time-consuming and hence is usually only performed when data for review have been restricted to a manageable size. In order to achieve this, linkage is often performed iteratively, with relaxed or altered matching patterns for each step of the process.

Examples of a deterministic linkage strategy performed in an iterative, deterministic manner include all of the studies mentioned in section 2.7.4.[105, 111, 124, 131–133] For instance, the pilot study of linkage between NN4B and registry data records an algorithm containing 15 steps. The first of these matched data from 153,572 subjects (of 155,126 NN4B and 155,034 registry data) based on NHS number. There were then 11 stages where separate attempts were made to match live and still birth data, before a further two rounds of trying to match any case, regardless of birth status. Finally, 154,885 records were linked, 15 following manual review. Unmatched subjects from the registry data were 136 live births and 13 still births; from NN4B, the corresponding numbers were 191 and 50.[111]

Similarly, the studies in which HES and NN4B/registry data were linked for 2005-06 [132] and 2007 [133] used an algorithm with nine steps for the maternity data, achieving a match rate over 90%. By contrast, the algorithm for the baby data used five steps but matched fewer than 85% of the data, almost all on the first step which used "exact sex + partial date of birth + exact NHS number".[132]

2.8.2 Probabilistic linkage

Probabilistic linkage has developed since the Second World War in tandem with the evolution of modern computing.[136, 138] For each subject, information is compared variable by variable between the two data sets; for each variable, a weight is calculated

based upon whether or not the information matches. "Matches" do not necessarily have to be identical: for example, phonetic algorithms exist to enable matching of words (particularly, names or addresses) that may commonly be spelled differently (e.g. the inclusion or omission of a space from the standard post code as used by the ONS for geographical locations in the UK[139]).[140]

Fellegi and Sunter were the first to describe a methodology for record linkage in 1969,[141] using "advanced mathematics".[136] The basic principles are discussed in greater detail in chapter 5, which describes methods used for the linkage work carried out as part of this project. Essentially, probabilities are pre-specified for the chances of achieving a true matched (denoted 'M') or unmatched ('U') pair between corresponding variables from the two data sources available for comparison. These are then used to calculate a "weight" ('w') for each paired variable, as shown in equation 5.3 (chapter 5). For each subject (row), weights from each of the variables used in the matching are added together to provide an overall weight. Each subject is then ranked according to this total. Thresholds are then chosen to distinguish linked pairs, possible links and not-linked pairs (often denoted L, P, NL respectively).[137, 138]

The major consideration with this "classical" approach to record linkage is that it requires some kind of prior estimates of the M and U probabilities. This may be obtained by basing the linkage on knowledge gained from a previous study, or by basing estimates on some other prior knowledge: for example, estimates of likely error rates in the data sets.[137, 138]

An alternative is to calculate weights using an algorithm. Two approaches are described that may be of use: one was developed to aid linkage of data for an Italian cancer registry [142] and is known as the Contiero (or, "*EpiLink*") algorithm.[137, 142, 143] This calculates weights according to the estimated error rates and the frequency of responses for each of the variables used in the data linkage (see section 5.3.3 and equations 5.6 and 5.7 for further detail).[143] The other method utilises the computationally-intense Estimation-Maximisation algorithm,[143, 144] which is itself based upon regression techniques, to obtain estimates of the weights.[143]

2.8.3 Linkage issues

Important considerations in linkage analysis relate to the error rates. There are two main ways that errors [104, 136] may occur:

missed matches, also known as a false non-match or a "synonym error", where two records that truly match are not linked by the analysis;

false matches, also known as a mismatch or a "homonym error", where two records that do not truly match are linked by the analysis.

Missed matches will decrease statistical power as fewer true matches will be identified, [104, 136] hence also leading to underestimates of prevalence.[104] Conversely, false matches will identify more matches than there truly are, thus overestimating prevalence.[104, 136] Further problems may arise from duplicate records being matched (i.e. where a record from one data set is matched to multiple records within the other data, also known as "confusions").[104]

In order to minimise these issues, it is not uncommon to expend the greatest amount of time in linkage analysis on data preparation and standardisation.[136, 138] Duplicate records should be removed, variables need to be formatted identically in each data source, including any transformations (e.g. using a phonetic algorithm [140]) that will be used.[136, 138] Error may be assessed using common epidemiological measures of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), (discussed in more detail in chapter 5 and shown in equations 5.8 – 5.11).[136, 138]

To date, error reporting for linkage analyses has been poor. A well-conducted systematic review investigating error measures in probabilistic record linkage studies across the international literature found studies in English, French, Portuguese and Spanish. However, only 33 articles met inclusion criteria and just six of these reported information for sensitivity, specificity and PPV (the authors do not mention whether they looked for reporting of NPV).[145] When reported, sensitivity ranged from 86% – 99.2%, specificity was $\geq 99\%$ and PPV was between 68% and 99%; the suggestion was that linkage improved with increased numbers of variables used for matching, rather than with sizes of databases.[145]

Another study attempted to collate evidence relating to how differences at individual or group levels may affect validity of matching and cause discrepancies in linkage rates for different sub-populations, thus affecting outcomes and interpretation.[135] Again, 33 articles were identified by a search strategy covering three databases, but only including articles published between 1991 and 2007. Age, gender, ethnicity, site of

origin of the data, socioeconomic status and health status were considered as possible factors that may lead to differential linkage and consequent bias.[135] No differences in linkage rates were found by age in five of 18 relevant studies, with variable, non-consistent differences shown in the others. For gender, five of ten studies showed no difference, three demonstrated poorer linkage in men than women and two the opposite. Differing rates of linkage by minority ethnicity were found in 7 of 14 relevant studies, although five studies found no difference in linkage rates.[135]

The most important effects were seen with respect to the geographical location where the data were collected. Thirteen of the fourteen studies showed important differences by region – examples included greater numbers of unlinked records for city-dwellers, variable linkage by hospital type, or altered consent rates for participation by geographical area.[135] There was also marked variation by socioeconomic status, with nine of eleven studies showing poorer rates of linkage with lower socioeconomic status. It was postulated that this may be related to lower rates of consent and compliance with follow-up and/or data entry, meaning there were more missing data items for these participants.[135] Interestingly, there were few consistent differences found according to health status in eighteen studies but singleton/multiple status was found to be important in one study that also involved multiple admissions, and five studies found lower linkage rates for premature or LBW babies.[135]

To aid both readers in interpretation and authors for reporting of data linkage studies, guidelines have since been developed.[146] These cover four "domains": data sources, variables chosen for linkage and the data preparation, record linkage, and ethics and security issues; fourteen different points are included between them.[146] No study has yet attempted to assess the impact of these guidelines.

2.9 Statistical and epidemiological considerations

Assessment of potential error is an important consideration in all epidemiological studies, not only those involving record linkage. Incorrect conclusions may arise as a consequence of errors in measurement, divisible into two types, random or systematic error, or by misinterpreting an effect if, for example, potential confounding variables have not been taken into consideration.[147]

The consideration of alternative possibilities for results obtained from epidemiological studies has lead to the development of guidelines for reporting and evaluation of different types of trial. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines exist for cohort, case-control and cross-sectional studies.[148] For randomised and other clinical trials, the Consolidated Standards of Reporting Trials (CONSORT) guidelines [149] were created, and PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses – also exist.[150]

2.9.1 Confounding

Confounding is different from measurement error because the impact on results is not caused by differential reporting among subjects, but by a factor not accounted for thus far in the study design or analysis. Instead, the relationship between an exposure and an outcome is *confounded* when some (or all) of the effect is influenced by another variable. Such a *confounding* variable is one that is associated with both the exposure of interest and the outcome under consideration, but does not fall on the causal pathway between the two.[151] Consequently, it is not possible to identify confounders solely through statistical testing: given that there is an association between a third variable and the exposure and outcome in a study, some *a priori* knowledge is required to determine the direction of association. Identifying confounding factors becomes much easier when causal diagrams – also known as directed acyclic graphs (DAGs) – are used.[151, 152]

A special form of confounding is known as *interaction*, or *effect modification*.[148] This occurs when different effects are seen for an association when there is stratification by a third variable. For example, the study examining the effect of socioeconomic status on infant mortality using linked using registry data and NN4B found a strong association between deprivation and an increased neonatal mortality rate (NMR).[131] This was confounded by low maternal age, sex and ethnicity of the baby, and registration status. Combined, these factors reduced the size of the effect between deprivation and mortality from an OR of 2.0 (95% CI: 1.79 – 2.24) to 1.57 (95% CI: 1.41 – 1.75) for the comparison between the least and most well off.[131] Low birth weight, however, acted as an effect modifier, leading to no effect (OR 1.14, 95% CI: 0.97 – 1.34) in those born small and an important effect in those of a normal size ($\geq 2,500$g, OR 1.35, 95% CI: 1.09 – 1.67, $p < 0.05$).[131]

2.9.2 Chance

Random error is more commonly referred to as *chance*, and is accounted for by using appropriate statistical techniques to measure the effect and then providing an indication of the level of support (or, *evidence*) that lies behind it.[153, 154] This may be achieved using p-values, confidence intervals or – more commonly – a combination of the two.[153, 155]

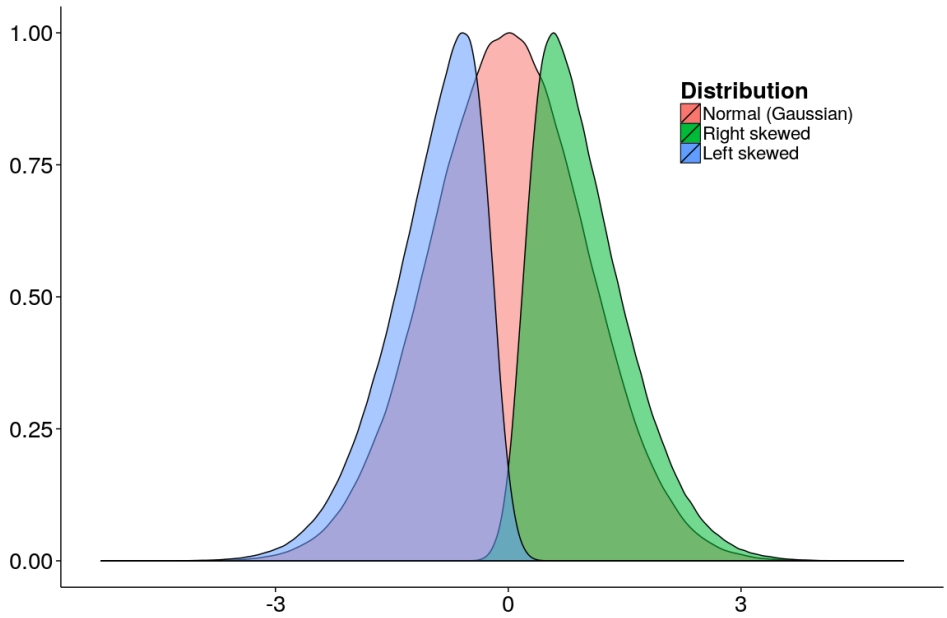
P-values provide a single point-estimate of the probability of accepting (or rejecting) the null hypothesis[153, 156, 157] – but only that: their interpretation needs to consider numerous other factors, including what the null hypothesis was, what the alternative hypotheses may be, what the sample size was (and, consequently, the statistical power there might have been[158]).[153–156, 159] Perhaps most importantly, the results need to be contextualised within the clinical (or research) scenario, as highly statistically significant results may be obtained that are of no clinical relevance or, conversely, an important effect may actually be being missed by the study.[153–156]

A further issue arises when multiple tests are carried out: by chance, some results will be statistically "significant" – regardless of the cut-off chosen. An example is provided by a study using functional magnetic resonance imaging (MRI) to assess the cortical responses of a salmon during questioning about the emotional states of humans represented in different photographs. An area of brain activity was identified that was 81mm^3 in size with a "cluster-significance of $p = 0.001$" – but functional MRI relies upon multiple statistical tests being carried out concurrently, and the description of the salmon tested stated it was "approximately 18 inches long, weighed 3.8 lbs, and was not alive at the time of scanning."[160]. The recommendation of that study was that functional MRI needed to "utilize multiple comparisons correction",[160] but the utility of tools such as Bonferroni's correction to adjust for multiple testing in epidemiological studies is debatable,[159, 161, 162] with many preferring the added information provided by confidence intervals.[154, 155, 158]

Confidence intervals provide a range of estimates around the single obtained point estimate upon which to base certainty.[153, 155, 157] Specifically, a confidence interval provides the range of effect for which a percentage certainty (frequently set at 95% but may be some other suitable threshold, for example 90% or 99%) is given of the true effect being within that range. Hence, for example, if the study were to be repeated

100 times, each time selecting a sample from the same study population, 95 of the studies would contain the true effect for that population within their 95% confidence intervals.[155]

Figure 2.10: The normal (Gaussian) distribution, with super-imposed right-skewed and left-skewed density plots.

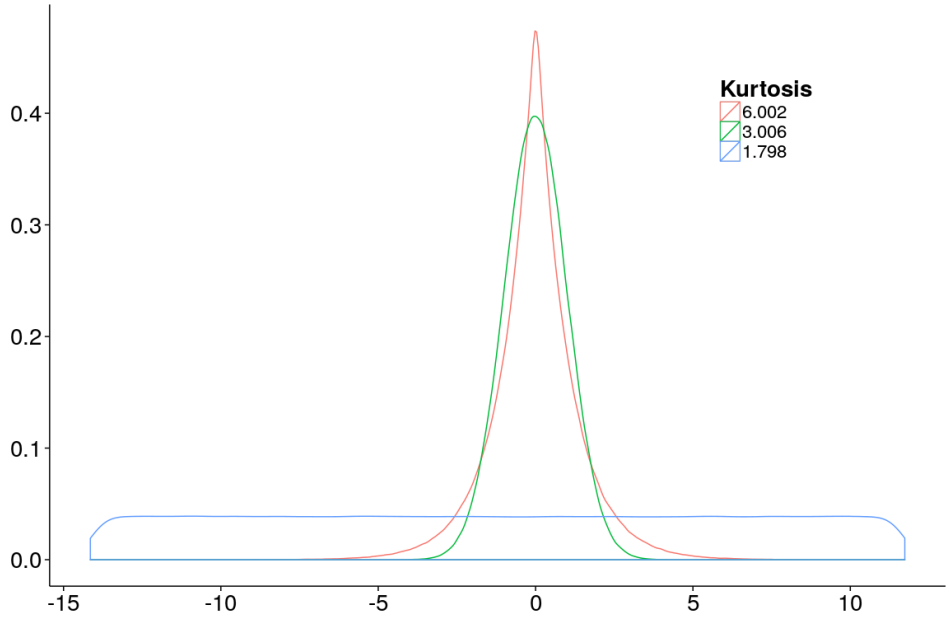


When choosing appropriate statistical methods to use for investigations, the underlying distribution of the data should also be considered, as this affects the choice of test to use. Distributions may be described in terms of their *skewness* and *kurtosis*. The former refers to how the data are balanced around their central tendencies (i.e. mean, median and mode) whereas the latter describes the relative "pointedness" or "flatness" of the distribution. The normal distribution has a skewness of zero and a kurtosis of three (although this is frequently "corrected" for by statistical software).[157] Examples are shown in figures 2.10 and 2.11.

2.9.3 Bias

Measurement errors that are not random are either due to problems with subject selection or with the information available from those subjects, hence may be divided into *selection bias* and *information bias*.[148] These affect all types of studies – including

Figure 2.11: Three distributions with varying kurtosis. The normal distribution is shown in green.



those self-reported as "randomised"[163] – and, while many quality assessment tools have been developed for appraisal of evidence,[164] historically few are validated.[163] A study that investigated agreement among six different grading systems for clinical evidence found extremely poor agreement between them all.[165] More recently, the success of tools like the Grading of Recommendations Assessment, Development and Evaluation (GRADE) can be assessed through popularity indices such as use by international organisations or numbers of citations.[166]

Selection bias

Selection bias may be a feature of different study types.[148] Prospectively collected studies may be biased such that they enrol subjects differently according to their risk profile; for example, those of lower socioeconomic status may be less likely to participate, but more likely to suffer the disease process. Retrospectively collected data may be biased if the sample subjects do not accurately reflect the baseline population; for example, assessing disease incidence using a local population but selecting patients from the local hospital that also acts as a tertiary referral centre for the wider district.

Information bias

Information bias may take many forms, varying upon the study design, and relates to systematic error occurring within the information flow from the time an event occurred through to the end of the statistical analysis. For example, in retrospective studies, information bias can arise if a subject fails to remember events that have happened previously; prospective studies aim to eliminate such *recall bias* by collecting data contemporaneously. Alternatively, *investigator bias* may be introduced by interviewers or at the analysis stage, hence the use of blinding,[165] and recommendations that analysis strategies should be specified in advance.[148, 153] Information bias may even be introduced after studies have been completed, with *publication bias* being a well-recognised problem. Fortunately, a systematic review that identified 26 studies investigating publication bias did not itself identify any evidence of publication bias in the results![167]

2.10 Workload assessment

Assessment of work may focus on numerous different factors, dependent upon the level it is being examined at. Broadly, in the context of clinical work load, these levels may represent the individual providing care, the unit or hospital, or the region; [168] however, little health services research occurs at the level of the individual health care provider.[169–171] One reason for this paucity of studies may be that health services studies are commonly conducted to assess fees for service, but workload may also impact on health outcomes.[168, 171]

Within neonatal medicine in England, workload assessment has been focused at unit or regional levels, primarily investigating the effects of nurse staffing.[28–30, 172, 173] One recent study examined nursing workload in order to estimate time required per (cot) category of care.[172] Nursing teams in three neonatal units were observed over a number of clinical shifts covering both night and day; observations covered all work related to babies admitted in the units.[172] Results were presented for "nurse minutes per baby hour", according to clinical categories of care provided; these were defined using two different scales that were additionally adjusted during the analysis phase.[172] Time spent caring for babies was related to category of care, although this

was only statistically significant ($p < 0.01$) when babies receiving CPAP were reclassified as receiving level two (high dependency) care.[172]

This study raises more questions than it answers, and highlights another issue: that concepts involved in workload assessment overlap multiple disciplines – for example engineering, biomedicine or psychology – thus leading to confusion with terminology or to ideas being poorly described. In the study just referred to, the numerator is really defined as *nurse time spent at a task* and the denominator is *category of care for the baby.*[172] This was different from the UKNSS, where staffing is defined as *number of nurses per unit* – and unit size was stratified according to patient volume, and numbers of both consultants and nurses.[28, 29]

It is therefore imperative to ask what the components are of the work that we seek to assess. The answer to this needs to be interpreted in the context of data that are realistic to obtain – that is, data that are possible to collect or, better, data that are already available. In the case of the latter, questions of importance become: which factors should – or could – be used as denominators, and which as numerators, and what are the available interpretations for each of these options?

2.10.1 Staffing

"Staffing" is merely a numerical concept: the number of staff. Assessment of staffing levels, however, is most readily comprehended with respect to a structural unit. This may be a hospital bed, ward or even a region containing multiple hospitals; alternatively, as in the study described earlier,[172] numbers of staff may be compared with numbers of patients.

The few studies on neonatal nurse staffing have mostly focused on nurse-to-patient ratios, defined using BAPM (or other) categories of care.[19] However, a systematic review and meta-analysis of nurse-to-patient ratios in neonatal medicine found only six studies, one of which had published results in two journals thus resulting in seven articles.[174] Associations with lower nurse-patient ratios were found with both increased (three studies) and decreased (one study) mortality; there was significant heterogeneity amongst all trials, and it was not possible to combine any of the results.[174] A similar – although slightly broader – systematic review defined the target population as children below the age of eighteen, and included other measures of nurse-staffing besides nurse-patient ratios such as "nurse hours per patient day", skill mix, education and

experience levels. Although it was extremely comprehensive in its methods, the authors found only eight studies.[175] Details of all excluded studies – along with justifications – and those included were provided. Only two studies reported on mortality: the UKNSS and a study conducted within a general paediatric population; these produced conflicting results which the authors suggested may be attributable to the different patient populations.[175] One other study used an exclusively neonatal population, and one study was conducted in multiple settings, including one neonatal unit.[175] Another slightly older paper examined neonatal staffing specifically in England in a discursive, non-systematic review.[176] This highlighted changes in the workforce, resulting in part from implementation of the European Working Time Directive and changes to nurse education and postgraduate medical training, and leading to increasingly widespread employment of ANNPs.[176]

To understand these studies more readily, it is helpful to have a framework in which to place them. One such mode for conceptualisation involves the elements *structure*, *process* and *outcome* and was used to develop a framework specifically examining nurse staffing.[177] The authors used a "realist logic" strategy employing "realist evaluation, realist review and logic modelling" to develop their model. The strategy initially focused on two pieces of work by a "highly-influential author", extracting potential causal pathways between nurse staffing and both patient and nurse outcomes.[177] Following this, the authors looked at a single patient outcome, *failure to rescue*, that had been covered in a previous systematic review. Using studies within the review as sources, the final stage was to extract possible mechanisms of action for the effects of nurse staffing on potential outcomes and to place them into a logical model.[177] This model is shown in diagram 2.12.[177]

Figure 2.12: Conceptual framework proposed by Subirana et al[177] for the effects of nurse-staffing on patient and nurse outcomes. **NOTE:** *Permission was **not** granted by the publisher to include this figure.*

Perhaps not surprisingly, the "structures" identified in this conceptual framework are similar to those identified in investigative work[175] and, indeed, other frameworks that have been proposed suggest similar factors.[171] For instance, "a conceptual model of physician work intensity" has been proposed that identified factors divided into three groups: patient, provider and practise.[171] This relates back to the three "structural

units" or "levels" described at the beginnings of sections 2.10 and 2.10.1, and suggests that it might be best to organise thinking along two dimensions: structure, process, outcome versus recipient (whether patient or bed), provider (medical, nursing and/or other staff), organisation (for example, unit, hospital or regional health authority).

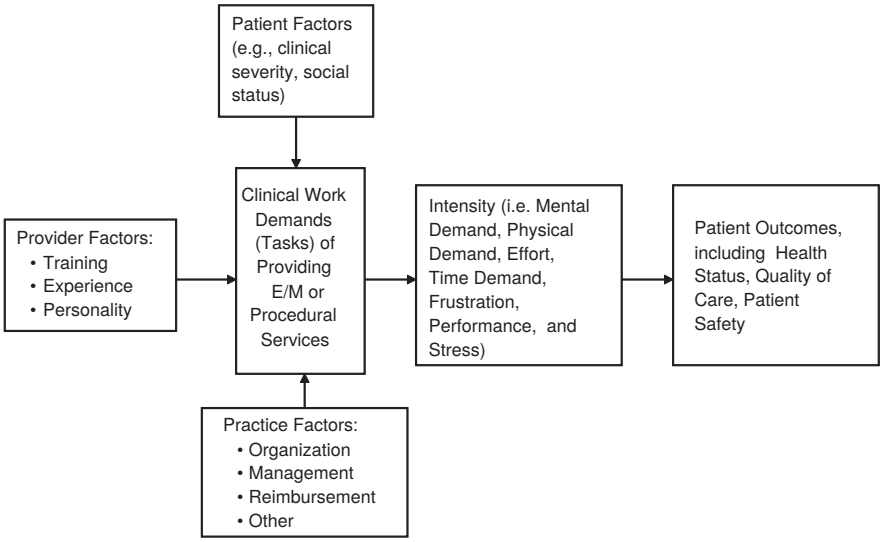


Figure 2.13: Conceptual framework proposed by Horner et al[171] for the effects of doctors' work intensity on patient outcomes. Reproduced with permission from Medical Care 2012, 50:654-661.

2.10.2 Activity

An important component of research into workload relates to the measurement of activity. As suggested by the frameworks described above, this can be measured at the individual or at grouped levels, and may refer to recipients or providers of health care. What is less clear is whether "activity" is a structure or an outcome – indeed, it can also be asked, is "staffing" a structure or an outcome?

In reality, "workload" – either in terms of assessment or with respect to workforce planning or (in the scenario described by this thesis) the effect on health outcomes – may be described as a "wicked" [178] problem: one that cannot be solved, at least not in the "tame" way that scientific problems are usually expected to be solved.[178] This *wicked*:*tame* dichotomisation was first presented in the context of social policy and planning.[178] Ten core characteristics of wicked problems were detailed by the

authors; the equivalents for problems that were tame were not listed, although examples were provided in the text. Some of these provide a great deal of clarity about wicked problems: they "have no stopping rule", meaning there is no single, ideal solution; similarly there are no "true-or-false" solutions, only "good-or-bad" ones; and, even, a "definitive formulation of a wicked problem" is not possible.[178]

What *is* possible instead is to solve "tame" problems within the "wicked" environment, although by definition changing one part of the system will have knock-on consequences elsewhere.[178] *Activity* with relation to work may be considered as the number of objects passing through a system: for example, babies treated in a cot or a neonatal unit. This corresponds to the queuing theory proposed by Little and shown in chapter 11 in equation 11.5.[179] Alternatively, without a time-component, *activity* may be considered in relation to the amount of work that is carried out; this may be called *intensity*, and aspects were discussed in section 2.10.1 in relation to staffing.

2.11 Summary

This chapter provides the background for the remainder of the thesis. It commenced with neonatology: first detailing important concepts and then describing the organisation of neonatal care in England in the 1990s in section 2.2. This included discussion about the UK Neonatal Staffing Study as well as transport services and the medical advances of the time – antenatal steroids and neonatal surfactant. The 1990s were also the era that large population studies into extreme prematurity really commenced. Among the first was the EPICure study – from whence some of the questions to be tackled by this thesis first originated – which is described in section 2.3 along with an example of the European context.

Around the turn of the century, evidence was increasing for the centralisation of neonatal care into larger hospitals. This resulted in the implementation of managed clinical networks which mostly occurred between 2004 and 2007. In the middle of this period, during 2006, the EPICure 2 study was carried out. Section 2.5 describes this, including a comparison with the first EPICure study and a description of the UPS that was carried out concurrently.

In section 2.6, the chapter starts to shift focus. Emerging evidence regarding the implementation of MCNs is presented, as well as discussion about the European land-

scape. The improvements in care seen at the end of the previous century had led several countries to implement large population cohorts, and these in turn raised issues around potential ethical pitfalls at the borders of viability; these are all discussed. However, ethical problem-solving requires detailed knowledge and this, in turn, requires "big" data. Consequently, section 2.7 is devoted to the sources of health care data that are available in England and that may be of use for neonatal research: registry data (including NHS numbers for babies), Hospital Episode Statistics, the National Neonatal Research Database, and other data sources.

The utility of multiple data sources may be increased by combining them hence, in section 2.8, I focus on the methods that are available to achieve this. Both deterministic and probabilistic linkage are discussed, as are related issues like errors that arise during record linkage. However, such considerations are not solely limited to record linkage: there are many epidemiological and statistical considerations to be aware of. In particular, errors may arise as a consequence of confounding, chance, or bias, discussions about each of which are also presented. The following – and penultimate – section of the chapter then focuses on a topic which tends to span disciplines: workload assessment. This is divided for discussion into *staffing* and *activity*, although it is also discussed that workload is a "wicked" problem, and therefore an ultimate answer will remain elusive.

Chapter 3

Aims

This thesis, which arises from the larger EPICure research studies into extremely preterm births, seeks to explain the demographic, clinical and organisational factors that resulted in a large increase in the number of extremely premature babies (from 22 to 25 completed weeks of gestation) admitted to neonatal intensive care between 1995 and 2006 in England [55] and, where possible, aims to extend that knowledge to the present time. Three specific aims relating to the overall project will be explored:

1. To investigate whether the apparent 44% increase in admissions to the neonatal unit of babies born between 22 and 25^{+6} weeks gestation in England between 1995 and 2006 is true and, if so, the reasons behind the increase.

2. To evaluate the impact of obstetric interventions on perinatal outcomes using data from the EPICure 2 cohort.

3. To explore the organisational changes that have occurred in neonatal care in England since the 1990s.

Further detail on each of these is provided below.

3.1 Validation of the EPICure data sets using routinely collected Hospital Episode Statistics.

In order to determine if the apparent 44% increase in admissions of extremely premature babies to neonatal intensive care is true, this investigation will validate findings

from both EPICure studies using data from Hospital Episode Statistics (HES). All live births identified in the EPICure cohorts will be compared with those recorded in HES. Consideration will be given to exploring patterns of variation due to changes in population make-up that may have impacted the admissions rate – for example, through differing ethnic or socio-economic profiles.

3.2 Identification of obstetric antecedents of extreme prematurity.

The effect of obstetric interventions – specifically, the use of antenatal steroids and tocolysis, and mode of delivery – on outcomes at birth for babies born at less than 27 weeks gestational age will be investigated using data available from the EPICure 2 cohort. Condition at birth and death during labour or in the delivery room will be considered as the primary outcomes.

3.3 Organisational changes in neonatal care in England

Changes in the organisation of neonatal care will be assessed using data from the UK Neonatal Staffing Study (UKNSS, conducted in 1997) and Unit Profile Study (UPS, 2006). Additionally, given that the network model of care was still at an early phase of implementation at the time of the UPS 2006, a new Unit Profile Study (compatible with the previous surveys) will be distributed to neonatal units in England. This contemporary survey will facilitate a crude estimate of the current effects of organisational practices on neonatal care, as well as permitting the exploration of differences in unit profiles between 1996-7, 2005-6 and 2010-11. A particular focus of this study will be on nurse staffing and unit activity. Consideration will be given to changes in geographical organisation over time (i.e. with the introduction of managed clinical networks), combining area based information with unit level data using regression techniques to obtain mutually adjusted estimates of effect. A further goal of the 2011 survey will be to provide baseline staffing data by network level for use by the British Association of Perinatal Medicine and others.

Chapter 4

Methods overview: the available data

Seven data sets play a key role in this thesis: the first two are from birth cohorts of extremely premature babies in England; these are each paired with an administrative data set that contains supplementary information; the three remaining data sets are related to the organisation of neonatal care in England. This chapter provides detail on all of these data sets, including their provenance, key features and governance issues, as well as introducing methodological concepts common to all the work in this thesis.

I commence with describing the data available from the two EPICure cohort studies, carried out in 1995 and 2006. These data are used alongside Hospital Episode Statistics (HES, introduced later) in the Admissions Validation Study (AVS), a linkage exercise performed to validate data collection of the respective sources; additionally, EPICure 2 data are used to investigate obstetric antecedents of prematurity.

Alongside the second EPICure cohort, the same research team conducted the Unit Profile Study (UPS), detailing functional and structural aspects of neonatal health care organisation at that time. The Unit Profile Study (UPS) of 2006 was itself based upon the UK Neonatal Staffing Study (UKNSS) that took place in 1997; data from both studies are also used in this thesis. Additionally, an updated version of the Unit Profile Study was carried out in 2011 as part of this thesis. These studies are presented in this order. Together, data from the three are used to investigate changes in staffing and activity that may have occurred over time.

Next, I discuss the HES data in more detail. Acquiring these data necessitated

obtaining specific permissions as they included personal identifiable data (PID). This is described in the following section, along with ethical, data and statistical considerations relevant to the rest of the work in the thesis. The chapter ends with a summary of the information that has been presented.

4.1 The first EPICure study – 1995

The EPICure studies were two prospectively collected observational studies of extremely premature births. The first was carried out between March 1st and December 31st 1995, inclusive, with data collected on all births that were believed *at the time of birth* to have occurred between 20 and 25 completed weeks of gestation. There were no exclusion criteria, and the 276 hospitals where data were collected represented 100% of the services providing maternity care across the whole geographic area of Great Britain and Ireland. Follow-up of the cohort has occurred at 2.5, [56] 6,[57] 11 [58] and 16 years of age [180], with a further follow-up at 19 years currently under way.[181]

Data collection for the cohort was carried out in two phases. So as to avoid duplication of effort, data relating to all births under 26 weeks gestational age were collected in conjunction with the the Confidential Enquiry into Stillbirths and Deaths in Infancy (CESDI) into log books containing only nine variables (identification number, hospital of birth, date of birth, gestation, birth weight, gender, number of fetuses, signs of life at delivery, and whether admitted to into a neonatal intensive care unit) located on the labour wards. For babies that were admitted onto a neonatal intensive care unit, a comprehensive "case record form" was then commenced. This comprised a 24 page form with questions relating to:

- entry criteria;
- maternal data;
- demographic data;
- status at discharge or death (divided into pre- or post-estimated date of delivery (EDD));
- drugs at death or discharge;

- nutrition and growth;
- ROP;
- head ultrasound scan data;
- other problems; and,
- contact information for follow-up.

A complete list of data items included is listed by Costeloe et al in a supplementary appendix to their paper,[54] and is replicated in appendix A.

All data from the study were collected by health care staff involved in the provision of care, with the labour ward log books being returned to the main research office in London on a monthly basis.[1] Where gaps or data ambiguities were identified, one of the principal investigators visited local hospitals to review individual sets of case notes; only if the case review forms could not be completed following this were the data considered "missing". Data were then single-entered into a computerised database by two study nurses, with double-entry being used for a random 10% sample in order to verify accuracy. This was reported as finding an error rate of 0.1% – 17 errors among a total of 15,280 data points entered.[54]

Available data

Of the 4,004 fetuses delivered during the study period, 843 live born babies were admitted into a neonatal intensive care unit. Gestational age for these babies was then reassessed by the principle investigators using the criteria shown in table 4.1. Eight hundred and eleven babies were confirmed as having a gestational age that met the inclusion criteria, with subsequent attrition as shown previously in table 2.2.[54]

For this thesis, data were provided on the 668 babies who were born to an English mother in an English hospital.[2] This data set had been used extensively prior to this point, and contained over 600 original and derived variables as well as information relating to subsequent follow-up of survivors at 2.5, 6 and 11 years of age. Cleaned variables were noted at a meeting between myself and Ms. Enid Hennessy in April

[1] Personal communication with Professors Kate Costeloe and Neil Marlow.

[2] Data were provided to me by Ms. Enid Hennessy, trial statistician, based at the Wolfson Institute of Preventive Medicine, Queen Mary, University of London.

Table 4.1: Hierarchy of criteria used to determine gestational age in the EPICure studies. Adapted from Costeloe et al (2000 and 2012) [54, 55]. LMP: last menstrual period; USS: ultrasound scan (performed at < 20 weeks in 1995, or the earliest scan performed at any time in 2006).; EDD: estimated date of delivery.

Situation in which EDD to be determined	EPICure (1995)	EPICure 2 (2006)
If LMP and USS both present but differ by > 14 days	USS	USS
If LMP and USS both present	LMP	USS
If LMP only present	LMP	LMP
If USS only present	USS	USS
If neither USS or LMP present	clinical judgement	clinical judgement

2011, and a copy of the original case record form annotated with variable names was provided. Data were formatted as a Stata [182] "`.dta`" file.

4.2 EPICure 2 – 2006

EPICure 2 was conducted during the entire calendar year of 2006, and was limited to births that occurred between 22^{+0} and 26^{+6} weeks gestation to mothers who were usually resident in England and delivered at one of the 182 maternity hospitals there, with local midwifery units individually contacted early the following year to check for any missed deliveries.[55, 68] As for the first EPICure study, data collection was performed by clinical staff on the labour wards in conjunction with the confidential enquiry investigating perinatal mortality – renamed in 2003 to the CEMACH[116] – with site visits once per month from the principle investigators to ensure data completeness.[55] Each fetus delivered required a new notification pack to be opened; these contained a "perinatal notification" form ("PN:E2" form), and bar coded sticky labels to aid unique identification. The birth was recorded in a labour ward log book using one of the sticky labels, specimens were collected from the placenta or umbilical cord where possible, and the labour ward staff completed as much of the detail in the PN:E2 form as possible. For babies admitted to neonatal intensive care units, the form remained with the baby throughout the duration of the hospital stay for completion by subsequent care givers, else for still births or babies that had died in the delivery room the relevant parts of the form were completed and returned to the study office at the Homerton Hospital in London.[55, 68, 183] Consequently, many more detailed data were available for those

babies who were *not* admitted into intensive care, in contrast to the situation with the first study in 1995. All data collection forms and other documentation are available on the study website at `http://www.epicure.ac.uk/publications/study-documents-for-professionals/`.[1]

Once data had been collected, they were double-entered onto a computerised database designed specifically for the study.[55] The individual data items collected were similar to those from 1995: they were again listed in a supplementary appendix by the study authors and are repeated in full in appendix B in this thesis.[55]

Available data

Four thousand, one hundred and five rows of pseudonymised data were initially provided, with a subsequent 4,144 rows of identifiable information (date and time fetus delivered, maternal and baby NHS numbers, maternal date of birth and home postcode) provided for the matching studies carried out in this thesis. Both data sets contained a unique EPICure ID number that could be used for linkage if required. In the identifiable data set, 4 rows were determined to be duplicates and hence removed.

In both data sets, the 3,133 rows corresponding to the births confirmed by the principle investigators as occurring between 22^{+0} and 26^{+6} weeks gestational age were marked. Criteria used to determine gestational age are shown in table 4.1.[55] The main working data set comprised more than a thousand original and derived variables; these were provided in Stata [182] "`.dta`" format and discussed at the same meeting as for the first EPICure data set, with Ms. Enid Hennessy in April 2011; annotated collection forms were again provided.

4.3 Unit Profile Study – 2006

The EPICure Research Group additionally were responsible for running the first Unit Profile Study, conducted in 2006 (UPS 2006). This was defined as a survey of hospitals providing neonatal care within a maternity setting in England, hence excluding neonatal intensive care units situated within children's hospitals such as Great Ormond Street Hospital (in London) or Alder Hey Hospital (Liverpool) as well as stand-alone, midwifery-led birthing centres. The main EPICure contact at each of 182 units that

[1] Last accessed: 7th September, 2014.

were identified as meeting these criteria was mailed by post with a copy of the 4-page form and asked to complete the survey in conjunction with both nursing and medical colleagues.[68]

The forms were originally meant to be mailed out in January, 2006, with complete returns envisioned approximately 6 weeks later, by the middle of February.[68] However, this was subsequently revised to the end of July (see appendix D, which shows a draft version of the form) and, ultimately, the mail-out did not occur until the autumn.[70]

The survey contained four sections, relating to "activity", "cot capacity and staffing", "maternity unit policies", and "paediatric policies" (see appendix D). The first two sections gathered data similar to that collected by the UK Neonatal Staffing Study (UKNSS, described below), while the remaining sections carried an additional set of questions aiming to distinguish exactly what protocols were in place in each unit for dealing with babies born at the edges of viability (appendix D). Data from these latter sections have *not* been used in the work presented for this thesis.

Activity questions in the Unit Profile Study related to the previous complete year, 2005, or respondents were asked to specify an alternative time period to which the data related. Units were asked about the total number of deliveries in the hospital, total number of admissions and number of admissions born < 1500g, and the total number of both babies provided with and days provision of ventilatory support or CPAP.

In the section on cot capacity and staffing, there were questions relating to the number of cots, how the unit categorised those cots (the BAPM system had changed in 2001, but not all units were using the updated system), unit designation (i.e. level of care provided), whether cardiac or surgical services were provided on-site, number of consultants and number of business ward rounds, and the whole-time equivalents (WTE) of nursing staff employed, including the numbers who were *qualified in specialty* (QIS).

Data were provided to me by Professor Elizabeth Draper of the EPICure Research Group as a Microsoft Excel file containing four sheets. These were saved separately as comma-separated value ('csv') files, for subsequent use in statistical analyses. Column headings were self-explanatory with respect to the form used for data collection; data were available from all 182 (100%) of the hospitals surveyed.

Additionally, members of the EPICure Steering Group, independent of the principle investigators, assigned activity levels to the responding units based on both the number

of days of respiratory support provided (< 500, 500-1999, ≥ 2000 days) and number of consultants with more than 50% of their time dedicated to neonatology (four or more, one to four, and no consultants). This created three unique categories: high (≥ 2000 days of respiratory support *and* ≥ 4 consultants), medium (500-1999 days of respiratory support and one or more consultants), and low activity (< 500 days and no dedicated consultant cover for neonatology).[70]

4.4 UK Neonatal Staffing Study – 1997

The UK Neonatal Staffing Study took place prior to the Unit Profile Study, in 1997. As part of a wider investigation into the effects of neonatal staffing on outcomes, a census of all hospitals providing maternity care in the United Kingdom (UK) was conducted – a total of 250 at that time. The study was run by the UKNSS Collaborative Group, led by Dr. Janet Tucker, and was posted to each unit, along with two posted reminders and one telephone call. Hospitals were asked to respond using their most recently available annual data; 98% of hospitals responded for the previous year, with the remaining 2% using data from 1995.[29]

Overall, a 100% return rate was achieved from the existing 246 units in the UK: three hospitals that were contacted no longer provided any neonatal services, and two units had merged; 186 (76%) provided neonatal intensive care and the remainder special care with provision for stabilisation of a sick baby only.[28]

The survey itself was a simple, single-sided questionnaire (shown in appendix C) that collected information similar to the UPS nine years later. After clarifying the year for which data were presented, the census asked about the total number of deliveries in the hospital, and – for the neonatal service – the total numbers of admissions, admissions < 1500g, babies ventilated or given CPAP, days of ventilation or CPAP provided, cots, consultants and nurses (in WTE). Hospitals were also asked if they provided neonatal surgical or cardiac services, and if there were any major changes planned for the coming year.

The data for units in English hospitals were shared with the EPICure Research Group for analysis in combination with the Unit Profile Study as a csv file. Permission for this was confirmed in a letter from Dr. Tucker to Professor Costeloe on the 6th

February, 2008[1], which commented,

> "[These data sets present] .. a unique and important opportunity to study changes in the pattern of neonatal care and of the availability of trained staff in English hospitals over the past decade." [2]

The planned analysis did not happen immediately; when the data were finally checked by me in September 2012, it became apparent that there were only observations for those units classified as providing intensive care, rather than for all maternity units in England. Unfortunately, in the interim period, the original data had been destroyed and therefore were no longer available.[3] Consequently, analyses using the UKNSS data set in this thesis could only be performed using the restricted data set of 146 units in England that provided neonatal intensive care in 1997.

Additionally, there was an oddity with the data from the UKNSS. This was the inclusion of a variable – and data – for high dependency (HDU) cots: there was no record of these being collected in the original data collection form (appendix C), nor were they reported on by the UKNSS.[28]

4.5 Unit Profile Study 2011

In 2011, the Unit Profile Study was repeated (UPS 2011). The survey was again carried out by the EPICure Research Group, this time collaboratively within the Neonatal Economic, Staffing and Clinical Outcomes Project (NESCOP) with the following participating organisations: the Neonatal Data Analysis Unit (NDAU), based at Imperial College, London, Bliss (the charity "for babies born too soon, too small, too sick"), the British Association of Perinatal Medicine (BAPM) and the National Neonatal Audit Programme (NNAP) of the Royal College of Paediatrics and Child Health (RCPCH).

4.5.1 Preparing the questionnaire

I prepared the survey form, drafting questions and explanations based on the two previous surveys (UKNSS and UPS 2006) and comments from an initial meeting between

[1] Personal communication with Professor K.L. Costeloe, shared with myself.

[2] Letter from Dr. Tucker to Professor Costeloe, provided to me by Professor Costeloe.

[3] Personal communication between Professor Draper and Dr Tucker, copied to myself and Professor Neil Marlow, in October 2012.

collaborators in August 2011[1]. Additionally, I included three new questions relating to neonatal staffing coverage of junior doctors:

- Provide the total number of the following in whole time equivalents (WTE) *actually* working or on annual leave on the medical rotas in your unit on 22nd November 2011:
  1. Tier 1 (ST1-3/SHO).
  2. Tier 2 (ST4-8/SpR).
  3. ANNPs (in a medical role).
- Provide the total number of the following in whole time equivalents (WTE) who *should be* working on the medical rotas in your unit on 22nd November 2011:
  1. Tier 1 (ST1-3/SHO).
  2. Tier 2 (ST4-8/SpR).
  3. ANNPs (in a medical role).
- On 22nd November 2011, did you have dedicated medical rotas at:
  1. Tier 1 (ST1-3/SHO level)?
  2. Tier 2 (ST4-8/Registrar/middle-grade level)?

Each collaborating party was given the opportunity to provide feedback, with two draft versions being circulated in October 2011.[2] Further trials of the survey were performed by sending the questionnaire to several external reviewers who were asked to complete it.[3] An additional question (number 16, relating to the absolute numbers of nursing staff working on the 22nd November 2011) was requested by collaborators interested in the economic analyses, and by NNAP.

[1]Present at the meeting were: Neil Marlow (Professor of Neonatology, IfWH, UCL), Elizabeth Draper (Professor of perinatal epidemiology, University of Leicester), Andy Cole (Director, Bliss), Neena Modi (Professor of Neonatal Medicine, Imperial College London); I was not there, but received a comprehensive summary afterwards from Professor Draper.

[2]Draft questionnaires were sent to the following representatives of each organisation: Stavros Petrou (Professor of Health Economics, Warwick University), Andy Cole (Director, Bliss), Neena Modi (NDAU), Elizabeth Draper (BAPM/EPICure), Neil Marlow (EPICure).

[3]These were: Dr. Bryan Gill (BAPM), Dr. Michael Watkinson (NNA), Dr. Maggie Redshaw (NPEU), Kim Davis (NNAP/RCPCH) in an email circulated on the 30th September, 2011.

The questions finally used were again very similar to both the previous UPS in 2006 and to the UKNSS: this time, the survey was divided into three sections, relating to *activity*, *staffing* and *infrastructure* (or facilities). The four pages of the questionnaire were formatted throughout to provide obvious, precise questions in a large font followed by more detailed, smaller-sized explanations of what data were being requested. Boxes were provided in which to enter responses, along with extra space for additional, free-text notes; each page was numbered and contained a separate box with the return address in it.

Two time periods were covered by the survey. First, data relating to the calendar year of 2010, or the most recent complete year for which data were available, were collected. Second, a census of *current* staffing was made; these data related to "the present time", defined as Tuesday, 22nd November, 2011.

Within the 'activity' section were questions common to the two previous surveys. These related to the total number of women delivering in the hospital as well as numbers of admissions, babies provided with respiratory support and total days of respiratory support provided. The 'staffing' section contained similar questions to the UPS 2006 and UKNSS about consultant medical staffing and nurses, as well as the additional questions noted above about junior medical staff and the absolute numbers of nursing staff on a single day. The final section, about 'facilities', sought information about unit categorisation (including BAPM method used), number of cots, provision of surgical and/or cardiac survice, as well as two new questions relating to the number of "rooming-in" rooms available and the lowest gestational age at which babies were routinely admitted into the unit (as opposed to transferred *ex utero* to another unit). The final version of the paper questionnaire is reproduced in appendix E.

Additionally, a customised web-based survey was designed to match the paper version. This was created using a software called "Opinio" [184] that was provided by University College London for research purposes, and required not only inputting the questions and specifying formats for the responses, but also adjusting the cascading style sheets used to display the survey. The software then created a unique survey URL for each individual invited to participate; this allowed easy tracking of who had responded (or not) and whether they'd actually finished the online survey, or merely looked at it but not pressed the "submit" button. It also facilitated using another

email address for the 'sender', meaning responses came direct to me if there were any issues.[184]

4.5.2 Hospitals surveyed

The list of hospital units providing neonatal care was obtained by combining the 2006 Unit Profile Study list with a publicly available list published on the Neonatal Data Analysis Unit (NDAU), based at Imperial College, London website, supplemented with information obtained from NNAP at the RCPCH. As previously, hospitals for inclusion were defined by the presence of both maternity and neonatal services on the same site (i.e. the provision of a perinatal service). One hundred and seventy hospitals were identified for inclusion.

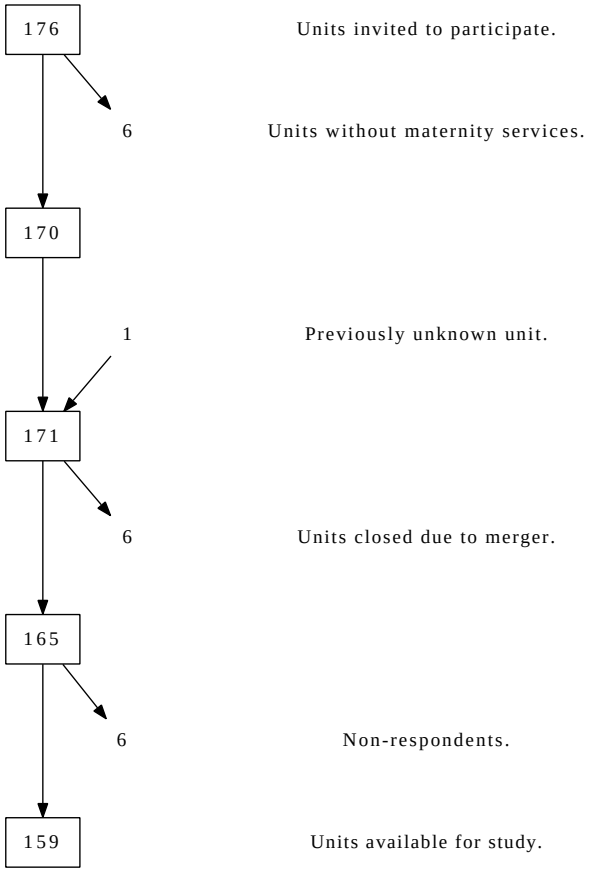
Copies of the UPS 2011 were mailed to a named contact at each hospital during the first week of November 2011, and the email invitation to complete the online questionnaire was mailed on November 15th; both of these letters were written by me but signed by Professors Marlow and Draper on behalf of the collaborating partners. Respondents were asked to return questionnaires either via the mail or using the custom web page as soon as possible after the 15th November.

Five reminder emails were sent to non-respondents over the following two months with links to the online survey. Following this, non-respondents were contacted individually by telephone and email over a period of six months; the survey was declared closed on the 30th June, 2012. I was assisted in contacting hospitals that hadn't responded by Dr. Laura McCormack, EPICure Research Manager.

Responses were received from 74 respondents via the web interface (32 completed entries, and 42 partially completed) and 134 respondents who completed the posted questionnaire. Six hospitals did not respond at all, despite multiple attempts to contact them; all of these were SCUs, hence would not affect analyses of provision of intensive care over time. The total number of units responding by the end of June 2012 (excluding 35 duplicate returns, three of which were duplicate paper returns from different individuals within the same hospital, the rest duplicated between the online and paper versions) was 159. This included one hospital that had not been previously identified as providing a perinatal service, Pinderfields Hospital – albeit, it replaced the service previously operating at Pontefract Hospital, and both hospitals were part

of the same NHS Trust. Twelve other units that were contacted had also merged. A flow chart of the numbers is shown in figure 4.1.

Figure 4.1: Flow chart of respondents to the Unit Profile Study (UPS) 2011.



4.5.3 Data management

Postal survey returns were entered by myself and Dr. McCormack into two separate bespoke MySQL [185] databases which I created. These data sets were then merged, with any discrepancies between them clarified by examining the original response recorded on the paper questionnaires. The online survey data were downloaded and converted into a format compatible with the processed postal data. The postal and online data sets were then appended to each other and duplicates compared, retaining the entry

with the most complete data.

4.5.4 Ensuring data accuracy

Following initial data entry and merging, assessment was made of the accuracy of data returns. This was performed using R.[186] Variables were prioritised for data checking according to their utility and perceived data integrity. Priority was given to ensuring that those variables common to the previous two staffing surveys were accurately recorded. Specifically, some units appeared to have been confused by questions 14 to 16: questions 14 and 15 related to the "total number of budgeted whole-time equivalent nurses who provide hands-on clinical care" – also known as the *establishment* figure – while question 16 asked about the (actual) numbers of staff who were working or meant to be working on the 22nd November, 2011. Confusion arose either when questions 14 and 15 were interpreted as requiring the WTE number of staff present and working on a single day, or when question 16 was completed with the number of WTEs rather than the absolute number of staff.

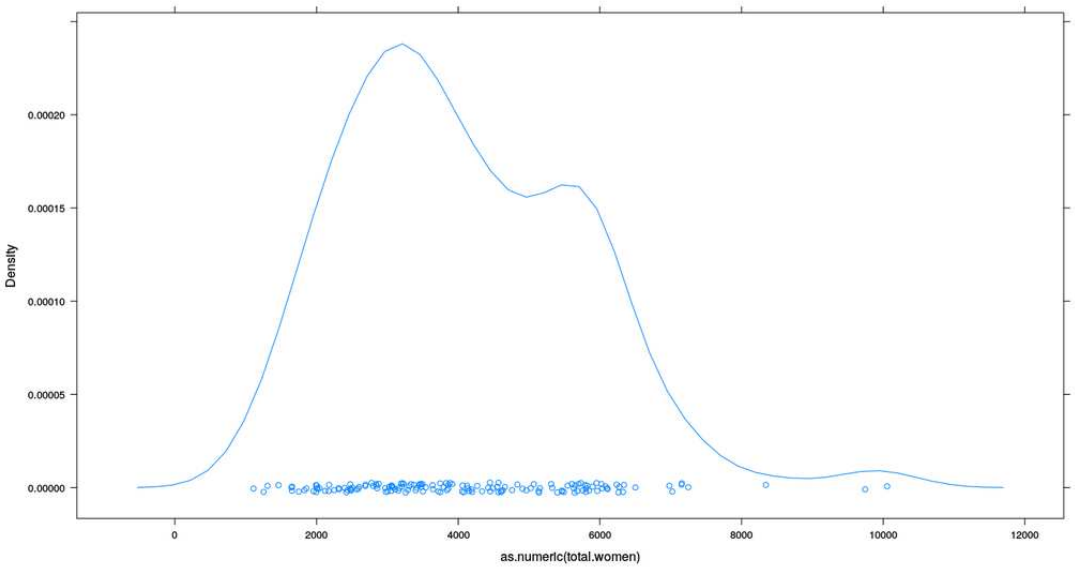
A further source of confusion perceived by the Neonatal Economic, Staffing and Clinical Outcomes Project (NESCOP) collaborating group was around how the numbers of cots were reported by different units: were units reporting the total number of cots in the building, those for which they were staffed as an establishment, or only those which were actually open on the 22nd November 2011?[1] The desire was to collect the latter of these – in order to provide an estimate of the total capacity of English neonatal services at that point in time; however, I suggested it would be still consistent with the previous surveys, as the exact detail hadn't been specified in either of the previous two surveys.

Numeric variables were plotted as kernel density plots (for example, as shown in figures 4.2a and 4.2b, for the total number of women delivering in each hospital and the total number of babies admitted into neonatal intensive care during 2010 (or the most recent complete year for which data were available), respectively). This permitted visualisation of outliers, allowing reasonable limits to be selected for further investigation and identification of units – for example, by selecting hospitals that reported greater than 7000 deliveries or 900 neonatal admissions per year. Scatter plots were used to

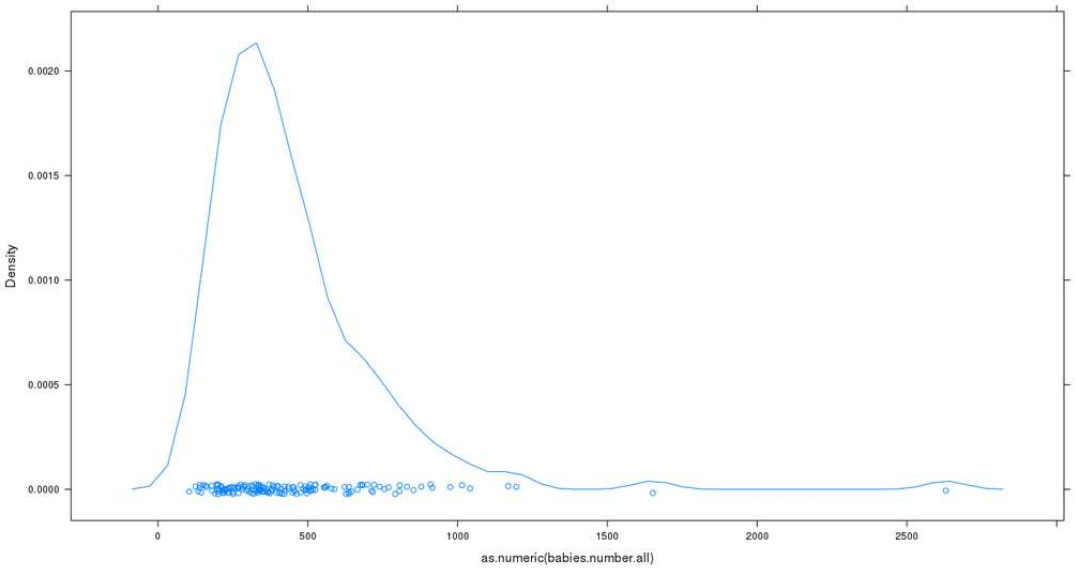
[1]NESCOP meeting, 4th March, 2013; notes available at `https://nescop.org.uk/trac/wiki/Meetings/20130304` (password-protected).

Figure 4.2: Example density plots, used for cleaning and data consistency checks of data for 2010 (or the most recent 12-month period for which complete data were available), returned by English maternity hospitals in the Unit Profile Study 2011.

(a) Total number of women who delivered ("total.women).



(b) Total number of babies ("babies.number.all") admitted.



check data consistency between variables within individual unit returns; an example with labelled outliers is shown in figure 4.3.

Categorical data were tabulated according to self-reported unit designation (i.e. level) to identify whether any data were systematically missing or appeared erroneous (table 4.2). Histograms were used to visualise spread of data among categories.

Following unit identification, apparently erroneous or missing data were checked by comparing them with the original data returns then, where necessary, contacting the original survey respondents again. This was conducted during a three month period over the summer of 2012, commencing two weeks after the survey closed at the end of June. In total, data were updated for 48 units, with a further eleven proving impossible to check and hence were dropped from analyses by converting the responses to missing data.

Figure 4.3: Example scatter-plot of the total number of women delivering ("total.women") compared with the total number of babies ("babies.number.all") admitted to neonatal units during 2010 (or the most recent 12-month period for which complete data were available) used in data cleaning and consistency checks for the Unit Profile Study 2011. LGI: Leeds General Infirmary; RDE: Royal Devon & Exeter Hospital.

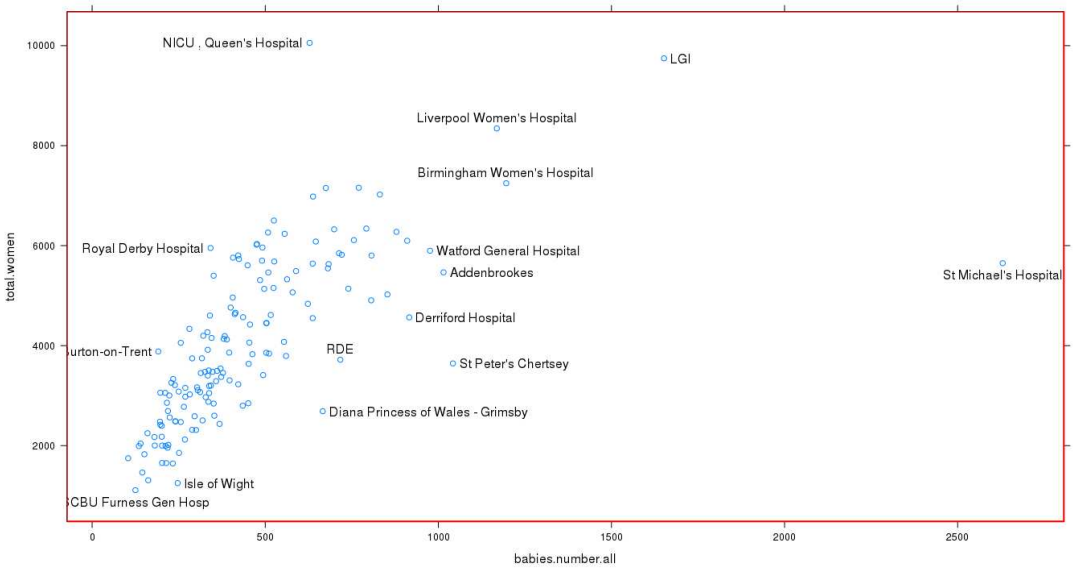


Table 4.2: Descriptive table of categorical variables from the Unit Profile Study (UPS) 2011, including some common to those available from the UK Neonatal Staffing Study and UPS 2006.

Variable	N (%)	Special Care Unit		Local Neonatal Unit		Network NICU		χ^2 P-value
Levels		n	%	n	%	n	%	
2010 data supplied	159 (100)							0.9226
No		10	13.3	6	15	7	15.9	
Yes		65	86.7	34	85	37	84.1	
Tier 1 (SHO): presence of a 24-hr on-call rota	156 (98.1)							< 0.001
No		24	32	21	52.5	0	0	
Yes		51	68	17	42.5	43	97.7	
Missing		0	0	2	5	1	2.3	
Tier 2 (SpR): presence of a 24-hr on-call rota	155 (97.5)							< 0.001
No		51	68	24	60	4	9.1	
Yes		24	32	13	32.5	39	88.6	
Missing		0	0	3	7.5	1	2.3	
Year of BAPM classification	153 (96.2)							0.5489
1992		3	4	1	2.5	2	4.5	
2001		71	94.7	36	90	40	90.9	
Missing		1	1.3	3	7.5	2	4.5	
Cardiac service	158 (99.4)							0.0129
No		74	98.7	39	97.5	39	88.6	
Yes		1	1.3	0	0	5	11.4	
Missing		0	0	1	2.5	0	0	
Surgical service	158 (99.4)							< 0.001
No		74	98.7	39	97.5	25	56.8	
Yes		1	1.3	0	0	19	43.2	
Missing		0	0	1	2.5	0	0	

Continued on next page...

Table 4.2: (Continued.) Descriptive table of categorical variables from the UPS 2011, including some common to those available from the UKNSS and UPS 2006.

Variable	N (%)	Special Care Unit		Local Neonatal Unit		Network NICU		χ^2 P-value
Levels		n	%	n	%	n	%	
Number of consultants on-call	158 (99.4)							0.0236
0-4		0	0	0	0	3	6.8	
5-7		40	53.3	30	75	23	52.3	
8-10		30	40	7	17.5	16	36.4	
11+		5	6.7	2	5	2	4.5	
Missing		0	0	1	2.5	0	0	
Number of consultants with >50% neonatal PAs	155 (97.5)							< 0.001
0		26	34.7	23	57.5	0	0	
1-4		37	49.3	10	25	4	9.1	
5-7		10	13.3	1	2.5	26	59.1	
8-10		0	0	4	10	12	27.3	
11+		0	0	0	0	2	4.5	
Missing		2	2.7	2	5	0	0	
Number of fixed consultant rounds (mean avge/day)	158 (99.4)							< 0.001
< 1/day		26	34.7	17	42.5	0	0	
1+/day		45	60	21	52.5	22	50	
2+/day		4	5.3	1	2.5	22	50	
Missing		0	0	1	2.5	0	0	
Lowest GA routinely cared for	159 (100)							< 0.001
< 24 weeks		5	6.7	0	0	43	97.7	
24-26 weeks		25	33.3	0	0	1	2.3	
27-28 weeks		43	57.3	4	10	0	0	
29-31 weeks		2	2.7	17	42.5	0	0	
32+ weeks		0	0	19	47.5	0	0	

Of note, the three questions that I had included in the survey relating to junior medical staffing were not well completed. This appeared to be for several reasons. First, there was confusion about whether "staff on the medical rota who provide hands on clinical care" included doctors only dedicated to neonates or also those who provided cross-cover with general paediatrics. Secondly, there was no provision for respondents who had different numbers of staff available at different times of day: for instance, some hospitals have increased numbers of doctors during the afternoon and evening periods, when there may be increased general paediatric attendances in accident and emergency (A&E) departments. Thirdly, it was unclear to respondents how they should account for "extra" doctors who may provide on-call coverage – for example, junior doctors working in community paediatrics sometimes work on-call in paediatrics at the local hospital during out of hours periods (evenings/nights and weekends).

4.6 Hospital Episode Statistics

Hospital Episode Statistics (HES) are an administrative data set relating to clinical "episodes" that take place either in an NHS hospital (including private patients treated by NHS consultants), or in an independent establishment treating NHS patients, that have been routinely-collected since 1989.[110] Each "episode" represents the care provided by one consultant.[122, 132]

For this thesis, two years worth of data were obtained in order to perform a linkage exercise (detailed in chapters 5 and 6) with the EPICure and EPICure 2 cohorts. Data governance and confidentiality measures are described below in section 4.8. The objective of the linkage exercise was to obtain supplemental socio-economic and ethnicity data for the EPICure cohorts to facilitate further investigation of the increase in admissions, and to mutually validate administrative data collection (i.e. the HES data sets) against the data collection performed for the EPICure cohort studies.

4.6.1 Data permissions

Responsibility for the HES data warehouse resides with the NHS Health and Social Care Information Centre (HSCIC); however, permission to use the data is – since April 2014 – granted by the Confidentiality Advisory Group (CAG) of the Health Research

Authority; [187] prior to this, the responsible body was the National Information Governance Board (NIGB). I applied to National Information Governance Board (NIGB) in August 2011 using the "fast track" process (subsequently known as "proportional review"), which allowed for permission to be granted by a subpanel of the Board, rather than having to wait until the next full meeting for discussion. Application in this way was possible as we met one of the pre-specified eligibility criteria:

> "Time limited access to undertake record linkage/validation and to pseudonymise the data" [188]

Despite this, the application still took 6 months, with final approval received on March 27th, 2012. Following this, access to the data was negotiated with the NHS HSCIC; additionally, they agreed to undertake an initial round of data linkage. This used a deterministic methodology and is discussed in more detail in Part II.

4.6.2 Initial management of the HES data

In September, 2012, the HES data were made available for download via the NHS HSCIC's "Data Depot". The download contained four files:

- `COHORT_OLD_NEW.txt`
- `COHORT_UNMATCHED.txt`
- `output_06.txt`
- `output_95.txt`

The first two contained details on the members of the EPICure 2 cohort that had been able to be matched (or not) with the 2006 HES data. The other two files contained the data sets for the respective years.

The files were renamed "`*.csv`" and the contents imported into R, [186] using appropriate filters. In the 1995 data set, 575,509 rows of information were provided and, in 2006, 631,499 rows: each row corresponded to a single birth. One hundred and twenty three variables were provided for each year, although a large proportion of these were empty (89 variables in 1995 and 58 in 2006); a further proportion contained information that was either non-discriminatory, e.g. the data entered for "admission checkflag" was

the same in every row, or non-informative – for example, "admission method" details how the patient was admitted to hospital; a majority of the data provided for this project were merely coded "Other: babies born in health care provider." [123]

Additionally, some of the provided data were either administrative or derived variables: there was a variable representing the year of collection (in financial, rather than calendar, years), several different system identifiers (study id, encrypted HES id, "epikey"), age at admission and length of stay (both derived). Potentially useful variables included information uniquely about the mother (date of birth, age, ethnic category, sex) or the baby (date of birth, birth weight, delivery method, delivery place, gestational age, live birth (birth status), birth order, resuscitation method), geographic information (post code, Health Authority of residence, lower and middle super output areas), socioeconomic information (using domains from the Indices of Multiple Deprivation (IMD)), as well as information relating to the admission (dates, level of care provided, destination following admission). The complete set of variables provided are detailed in appendix F.

4.7 Statistical analyses

For all investigations, statistical analyses were performed using R,[186]. Simple descriptive analyses were performed first. For the Admissions Validation Study, this was used to direct the linkage analyses. In the remainder of the investigations, the descriptive analyses were used to inform choice of potential confounding variables [151] prior to using multiple regression techniques. Specific methods are described in chapters 5 (data linkage), 8 (obstetric antecedents of extreme prematurity), and 11 (analysis of trends in organisational factors of neonatal units over time).

Following usage, all Hospital Episode Statistics data that did not match known EPICure subjects were securely deleted using the Gutmann technique; [189] remaining data were archived for future use by the EPICure Study Group. Information from the three staffing studies were also retained in storage for future use. Aggregated unit and geographical level datasets are available for collaboration on request.

4.8 Governance

4.8.1 Ethical considerations

Overall approval for the EPICure 2 study was received from City and East London Research Ethics Committee (REC) in August 2005 (reference: *05/Q0605/107*). Additional approval was obtained from the Patient Information Advisory Group (PIAG) (Reference: "*PIAG 3-07(f)/2005: EPICure 2: Population based outcomes for births before 27 weeks gestational age in England in 2006.*") in order to forgo the requirement to obtain consent prior to data collection in 2006. Ethical approval for the first EPICure study is no longer recorded (for instance, as was customary at the time, it is not mentioned in the original paper by Costeloe et al,[54]) but was granted at the time[1]. Both sets of approval from 2005 for EPICure 2 included permission to perform comparisons with data obtained in the first EPICure cohort. EPICure 2 was subsequently re-approved in 2008 by the Northern and Yorkshire REC for the continued follow-up of the cohort at $2\frac{1}{2}$ years of age (reference: "*08/H0903/51*"). This was due to City and East London REC having become defunct, meaning it was not possible to submit a substantial amendment to them at that time.

The use of HES data for the Admissions Validation Study (AVS) was therefore discussed with the Northern and Yorkshire REC, and a substantial amendment for this was submitted to them and subsequently approved in March 2012. Permission for the study was then granted by the Ethics and Confidentiality Committee of the National Information Governance Board for Health and Social Care. The NIGB were "an independent statutory body established to promote, improve and monitor information governance in health and adult social care,"[190] which had taken over responsibility for granting Section 251 (of the NHS Act 2006: permission to allow data collection in situations where it was either not practical or not possible to obtain patient consent[187]) from PIAG (reference: *ECC 1-02(FT3)/2012 EPICure 2 Outcomes for births before 27 weeks gestation in England*). Due to delays in obtaining the data and in performing the subsequent analysis, an extension was granted in March 2013 to allow for data linkage for the AVS to be concluded. The NIGB subsequently has been replaced by the Confidentiality Advisory Group (CAG) of the Health Research Authority (HRA).

[1]Personal communication with Professor Kate Costeloe and Professor Neil Marlow, August 2014.

The Unit Profile Study 2011 was also discussed with the Northern and Yorkshire REC. However, their criteria no longer required ethical approval to be obtained for surveys. This was in keeping with national guidance, which specified that service evaluation and audit did not require Ethical Review Board approval.[191] Permission to conduct the Unit Profile Study of 2006 had already been granted alongside the EPICure 2 cohort and, as indicated earlier, this was accompanied by a letter from Dr. Janet Tucker giving permission to use data from the UKNSS.

4.8.2 Data security

Data with personal identifiable information were required for the studies – particularly, for the AVS. However, where possible, pseudonymous datasets were used for statistical analysis. At all stages during investigations, all data – including pseudonymised datasets – were stored on physical media that was encrypted at the level of the block device (e.g. hard drive or cdrom), with concurrent encryption of virtual (SWAP) memory using `cryptsetup` [192] packaged for Debian GNU/Linux.[193] Identifiable datasets were only accessed from within the University College London network. Storage of the administrative databases with patient contact details was restricted to computers employing full disk encryption located in the EPICure study offices at University College London. For this, a systems-level security policy (appendix G) was produced; this was approved by the National Information Governance Board in December 2011.[1]

A Data Protection Act policy is in effect at University College London, and appropriate authorisation was sought (UCL Data Protection Act registration numbers: Admissions Validation Study: *Z6364106/2011/10/20*; Unit Profile Study: *Z6364106/2011/06/67*; both registered under *section 19, research: health research*). Additional registration was not required for the EPICure cohort data for these studies as they were pre-existing data sets already stored within UCL.

Any files that required external transfer during the course of the work for this thesis were without patient identifiers and protected using the same standards as described for university systems, in accordance with existing UCL policies.

[1] Personal communication between Adam Goodwin, ECC Security Review team, and Rick Borges, NIGB Deputy Operations Manager at the NIGB Office, cc'd to myself on December 19th, 2011.

4.8.3 Funding

Costs of the research were provided from the EPICure 2 grant, awarded by the Medical Research Council (MRC, file reference: G0401525). Dr. Andrei Morgan was employed in a combined clinical/academic post, joint funded by University College London Hospitals NHS Trust and University College London. Educational tuition fees were jointly covered by Dr. Morgan and Professor Neil Marlow.

4.9 Summary of data sets and common methods

This chapter presented the data sets to be used in the different analyses undertaken as part of this thesis. Data are available from the first EPICure study in 1995 for the 668 babies born in England at between 22^{+0} and 25^{+6} weeks gestation who were admitted into a neonatal intensive care unit. These are used in the Admissions Validation Study, along with the 575,509 rows of data available from the 1995 Hospital Episode Statistics data. The data representing the 3,133 confirmed births from the EPICure 2 are also used in the AVS, in conjunction with 631,499 rows of HES data from 2006. Additionally, the EPICure 2 data are used to conduct an investigation into specific obstetric antecedents of prematurity: antenatal steroids, tocolysis and Caesarean delivery.

Three other data sets were described in this chapter: the Unit Profile Survey of 2006 was a survey of 182 different perinatal centres, defined as those hospitals that provided both a maternity *and* neonatal service, in England. This survey had a 100% response rate. It was modelled on the UKNSS which surveyed all 246 perinatal centres in the UK; however, data were only received corresponding to the 146 maternity hospitals that provided neonatal intensive care in England. In order to permit contemporary estimates of the same measures, a further UPS was run in 2011; the response rate dropped slightly, with only 159 of 165 (96.4%) units responding – although the non-responders were all special care unit (formerly known as Level 1) and do not affect analyses comparing the provision of neonatal *intensive* care over time.

The chapter closed with some generic governance issues: a description of the ethical approvals required, particularly, to use the HES data for the AVS; discussion of the data protection measures that were required for data security; and an introduction to the statistical methodologies that will be used in the thesis.

Part II

Admissions Validation Study

Chapter 5

Data Linkage

This chapter presents the methods used in the Admissions Validation Study (AVS), designed to link data from the EPICure cohorts with Hospital Episode Statistics (HES) data. The chapter commences with an overview of the rationale for the study, followed by a detailed account of the methods used, summarised at the end. Results follow in chapter 6.

5.1 Study context

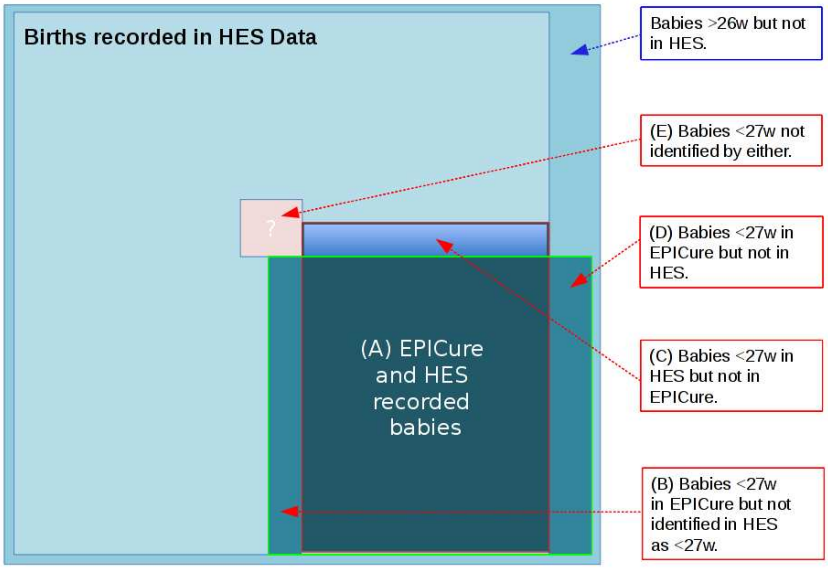
The prime objective of the AVS was to garner additional information with which to investigate the causes behind an observed increase of 44% in the admission rate of babies born between 22^{+0} and 25^{+6} weeks gestational age to neonatal intensive care units in England between 1995 and 2006.[55] This was unexpected, particularly as the birth rate increased by only 0.7% over the same time period, from 613,257 to 635,748 per year.[194] Key drivers for a change in the preterm birth – and, consequently, admissions – rate may include shifts in population make-up, specifically, as a consequence of migration; changes in attitude, both among health care personnel, and in wider society; and differences in data collection methodologies and completeness.

HES data were chosen for linkage with the EPICure cohorts because they contain supplementary information on ethnicity (of the mother) as well as socio-economic data (using the Indices of Multiple Deprivation [123]), they are available for time periods corresponding to when both EPICure cohorts were collected, from 1st March – 31st

December 1995 and 1st January – 31st December 2006, and nearly all births in England are registered within HES – very few births happen in non-NHS locations.[195]

Within the total number of births in each time period, the EPICure and HES data may overlap with each other in a number of ways. Figure 5.1 shows an example of this, using a large blue square to represent all babies born during the study period, which in this case is 2006. The overlapping, slightly smaller box represents all babies who were included in the HES data set: this is likely to be a subset of 'all babies' due to the incomplete coverage already described. The third box, outlined in green, represents those babies who are already included in the EPICure cohorts; and the final, red-edged box represents those babies in HES who meet the EPICure inclusion criteria.

Figure 5.1: Potential overlap among all births of EPICure 2 and Hospital Episode Statistics data collection in 2006. The largest blue square represents all babies born during the year. The overlapping, slightly smaller box represents all babies who were included in HES. The green-edged box represents those babies who were included in the EPICure 2 cohort; the red-edged box represents those babies in HES who meet the EPICure 2 inclusion criteria, i.e. have a gestational age of 26^{+6} weeks or less but not identified in either data set.



As demonstrated, this leaves five groups of patients who potentially meet the EPICure criteria:

A) Those matched in the EPICure dataset and in HES who already meet the inclusion criteria for EPICure.

B) Those identified by EPICure but with an incorrect gestational age recorded in HES

C) Those not previously known to EPICure but who are identified in HES as meeting the gestational age inclusion criteria.

D) Those in EPICure who have no data at all recorded in HES.

E) Those not identified in either HES or EPICure as having a gestational age of <27 weeks.

Of these, groups A and B (as gestational age and other criteria have already been checked for the EPICure cohorts whereas the HES data are unverified) only need supplementary data from HES. There will be no extra data available for group D. The most important group will be group C: these are subjects (extra to those in the EPICure cohorts) who are identified by HES as being of less than 27 completed weeks gestational age at birth. Group E will be subjects identified from the HES data – and missed from the EPICure data – who are potentially of the correct gestational age, but the gestational age identifying data in HES is either missing or thought to be wrong.

Thus, there is a possibility the EPICure datasets do not include 100% of participants (as believed) and that there may be other individuals who fit the criteria for inclusion in the cohorts. This is particularly true for children who were born in the first EPICure cohort, in 1995, as there were greater difficulties with subject ascertainment due to the wider geographical dispersion of the neonatal units. If this resulted in fewer than 100% of births being identified, it would have contributed to the observed rise in the admissions rate. The second objective of the AVS, therefore, was to investigate both whether EPICure data collection was complete *and* to assess whether the routinely-collected, administrative data, Hospital Episode Statistics (HES), could potentially be used as a reliable source for the routine surveillance of the outcomes of pregnancies that end at extremely premature gestations.

Two broad methods are available for matching data: deterministic, using known uniquely identifying variables (for example, NHS number), or probabilistic linkage, which depends upon statistical likelihood of two rows of data (one from each data set)

matching. Deterministic linkage requires an exact match for each variable included in the matching process, hence may be of limited use if there are any ambiguities or errors in the data. Conversely, probabilistic data linkage, also known by the term "fuzzy matching", is a process that allows for variation in the matching data, assigning a weight to each comparison pair and then selecting those pairs that rank highest as the most likely matches. Both "match" and "non-match" thresholds can then be applied, with manual inspection of the data if desired. This may be preferred because it allows for ambiguity in one or more of the matching variables, and is particularly useful for the present circumstances as the HES data – specifically, for 2006 – are known to contain inaccuracies. [132] A further benefit is that both the weights assigned to the matching variables and the predetermined thresholds for the identification of matching or non-matching pairs may be varied, allowing for great flexibility in identifying true matches. For this study, deterministic linkage between the HES and EPICure data sets was performed in advance by the NHS Health and Social Care Information Centre (HSCIC); hence, the major focus of the work described in this thesis is on the probabilistic matching of EPICure and HES data.

In summary, data linkage was performed between Hospital Episode Statistics and the corresponding EPICure datasets for the years 1995 and 2006. The primary objective of this exercise was to obtain supplementary data from the HES datasets with which to investigate an apparent increase of 44% in admissions to neonatal units among extremely preterm babies. The secondary objective was to assess the completeness of both EPICure and HES data collection. Probabilistic methods were chosen in preference to deterministic linkage because NHS number was not available in 1995 and because it was thought that there might be problems with deterministic linkage due to data inaccuracies.

5.2 Deterministic linkage

Deterministic linkage was performed by the NHS Health and Social Care Information Centre between HES data for 2006 and the EPICure 2 cohort in advance of data being made available to the EPICure group. Matching was not possible for 1995 as the algorithm used by the NHS HSCIC requires NHS number, which was only available for

EPICure survivors under current follow-up at 19 years of age,[1] and not for all births.

The full matching protocol[2] used by the HSCIC for linking HES with other data sets is as follows:

1. Exact sex, date of birth, postcode and NHS number.
2. NHS number, sex, and date of birth.
3. NHS number, sex, postcode and partial date of birth.
4. NHS number, sex, and partial date of birth.
5. NHS number, postcode.
6. Sex, date of birth & postcode (where NHS number does not contradict the match and DOB is not 1st January and the postcode is not on the ignore list).
7. Sex, date of birth & postcode (where NHS number does not contradict the match and DOB is not 1st January).
8. Exact NHS number only.

Data for the EPICure 2 cohort consisting of unique EPICure ID number, maternal postcode, NHS number and date of birth, child's NHS number and date of birth, birth weight, gestational age at delivery, sex, number of fetuses and birth order, and hospital of birth were provided to the NHS HSCIC via their secure data depot in August 2012. A total of 4,144 rows of data were transferred; 4 were identified as containing duplicate data by the HSCIC, and this was subsequently confirmed by Professor Draper, who was able to check the data against the original collection forms.[3]

Two files, one containing a list of the matched data and the other the unmatched EPICure 2 unique IDs, were returned from the HSCIC approximately 6 weeks later along with the complete HES birth data sets for 1995 and 2006.

[1] Personal communication: Dr. Laura McCormack, EPICure Study Manager, and Professor Kate Costeloe.

[2] Personal communication with Stephen Cowley on 4th September 2012.

[3] Personal communication with Professor E.S. Draper, 4th September, 2012.

5.3 Probabilistic linkage

Probabilistic linkage analyses of two datasets depend upon comparing each row from one data set with every row in the other. The total number of row pairs (one from each dataset) that are therefore available for comparison (denoted as 'N_{TRP}') is equal to the product of the number of rows in each dataset (denoted as D_1 and D_2 for the two datasets, respectively). This is shown in equation 5.1:

$$N_{\mathrm{TRP}} = D_1 D_2 \tag{5.1}$$

However, of the theoretical total number of row pairs, there is only an *actual* maximum number of matches ('N_{M}') of whichever is the lesser of D_1 or D_2 (or the total number of rows in one dataset, if they both contain the same number of rows). Therefore, the chance of any one row matching can be calculated. This is shown in equation 5.2:

$$\text{probability of single row matching} = \frac{N_{\mathrm{M}}}{D_1 D_2} = \frac{N_{\mathrm{M}}}{N_{\mathrm{TRP}}} \tag{5.2}$$

5.3.1 Data availability and choice of variables

Each analysis, for 1995 and 2006, was prepared separately, in two stages. First, the overall data sets were cleaned, with variables renamed and empty variables removed. The second stage then consisted of restricting the data to only the variables that would be used for matching, in order to reduce the size of the data sets and thereby limit the amount of computer resources (memory, processing power, hard disc space) that would be required.

1995

For the 1995 EPICure cohort, only data from the 668 babies who were admitted into neonatal intensive care were available (appendix A). This was because the complete data set – comprising the 4,004 births reported by the EPICure team in 2000 [54] – had been destroyed. Furthermore, although some of the original data collection forms were available, approximately 50% of the forms had been lost (and possibly destroyed)

when a study centre where one of the original principal investigators re-located.[1] The data from all available subjects were therefore used for matching, as HES includes data on live births occurring at any gestation, hence each of the babies admitted into neonatal intensive care in 1995 should have a corresponding entry.

Only 34 of the 122 supplied variables from HES contained data; a complete list of variables is provided in appendix F. Broadly, those available in 1995 were as follows: administrative data (e.g. HES identification information), admissions and discharge information, demographic information, geographic information (excluding any derived information), and birth information including birth weight, gestation, sex of baby, delivery place and method, and resuscitation method. Columns were provided for each of nine possible 'baby tails' that exist within the HES maternity data set; however, only the first of these contained data as information had been requested on a per-baby basis. In total, there were 575,509 rows of data; 127,958 of these were duplicates, but all corresponded to babies where there were insufficient data to adequately discriminate between multiple births (due to missing data in one or more of the matching variables); no data were deleted.

From the available data sources, the variables chosen during the second stage of pre-analysis processing for direct inclusion in the matching exercise were: baby's date of birth, sex, gestational age and weight at birth, birth order, total number of babies in the pregnancy, and the mother's number of previous pregnancies, as well as discharge date – although this was not expected to correlate highly between the data sources as HES relate to inpatient "episodes" rather than the complete length of hospital stay collected by the EPICure study. Maternal age at delivery was included in preference to maternal date of birth in order to minimise errors from data entry affecting results; date of death was also derived for HES, using "date of discharge" and "discharge method" (which indicated whether the patient died or was discharged home). "Ethnic category" was recoded in HES to match the EPICure categorisation and was included as a matching variable even though supplementary information on ethnicity was one of the desired results. Derived variables and ethnicity were included in the matching for 1995 (but not 2006) in order to increase the data available for subject discrimination as postcode information was unavailable.

[1] Personal communication with Dr. Alan Gibson, July 2012.

2006

The data sources available for linkage in 2006 were more complete than in 1995. The EPICure 2 data that had been provided to the NHS HSCIC for deterministic linkage were available; additionally, the remainder of the variables were available to provide supplementary information in order to differentiate potential subject matches between the data sets. The 4,144 rows of data the NHS HSCIC had been provided with for the deterministic match, after the terminations of pregnancy (which are not collected in HES birth data) were removed, were reduced to 3,376 rows. This was further restricted by removing the 626 still births that occurred before 24 weeks gestational age – as these deaths are not recorded in HES – meaning a total of 2,750 individual subject records were available for matching from the EPICure 2 data.

In the HES data, 65 of the provided 123 variables were populated (appendix F), although those actually of use for matching were restricted further as some variables provided information supplementary to that in the EPICure data set (for example, relating to socioeconomic status). The broad categories of data available were similar to those from 1995; in addition, geographic information was supplemented with derived information representing different domains of the Index of Multiple Deprivation. A total of 631,499 rows had been provided; 98 were found to be duplicates. Unlike the 1995 data, it was not thought that these data corresponded to different individuals as the majority of identified duplicates contained specific values. For example, birth weight was recorded in 51 of the duplicates, ranging in value from 200–6,200 grams and including values such as 338, 570, 652, 690, 720, 760, 880, 890, and 1,196g; additionally, there were very few duplicates identified. Consequently, these rows of data were removed and a total of 631,401 records were used in the data linkage.

Data standardisation

Using the available HES and EPICure data, two smaller data sets were prepared for each epoch to be examined (1995 and 2006) containing identical variables: each variable matched on data type (for example, `integer`, `factor` or `POSIXct` date format) as well as specific properties (such as the number of categories for factors). A full list of variables included in the matching exercises is shown in table 5.1. Additionally, a unique identifier

Table 5.1: Variables selected for matching the Hospital Episode Statistics and EPICure data sets in 1995 and 2006 that were common to both sources of data.

Variable name	1995	2006
Date of birth	✓	✓
Gestational age at birth	✓	✓
Sex	✓	✓
Discharge date	✓	×
Date of death	✓	×
Birth weight	✓	✓
Delivery method	✓	×
Number of babies	✓	✓
Birth order	✓	✓
Maternal date of birth	×	✓
Maternal age	✓	✓
Previous pregnancies number	✓	✓
Ethnic category	✓	×
Post code	×	✓

was created prior to reducing the data sets, to facilitate later re-combination with the original data.

Date variables (dates of birth for mother and baby, date of discharge and date of death) were were converted to `POSIXct` format [196] for all data sets. Post code was stripped to only contain alpha-numeric characters, i.e. all spaces and extra symbols such as "/" or "-" were removed. Sex was classified as "male", "female" or "indeterminate" in 1995, and as "male" or "female" in 2006 due to a change in the coding of HES data from 1996-97.[123] Gestational age at birth in weeks, birth order, number of babies, maternal age and number of previous pregnancies, and birth weight were provided as integers, the latter standardised across the data sets as grams (g). Delivery method and ethnic category were both categorised according to the EPICure criteria: delivery method was divided into "Caesarean in labour", "Caesarean not in labour" and "vaginal"; for ethnicity, categories were "black", "Indian/Pakistani/Bengali", "white", "other" and "unknown".

For both years, data were recoded if they seemed implausible. Dates of birth (of babies) recorded as occurring before January 1st or after December 31st were recoded as missing; similarly, age for women who were recorded as giving birth younger than 10 or older than 65 years old and gestational age for babies who were recorded as being

of less than 20 or greater than 45 weeks at birth were both recoded. Finally, for the 1995 matching, births recorded as occurring during January or February were removed from the HES data during the matching process.

5.3.2 Blocking variables

When data sets are large, the number of potential matches can be extremely high (see equation 5.1). Consequently, it can be beneficial to use "blocking" variables which are assumed to be correctly entered data but may not be highly discriminatory in order to reduce the number of potential matches and, therefore, the required computing power. An example would be to use "month of birth" in this way, thereby reducing the number of matches to approximately a twelfth. However, the counter balance to this is the loss of power if data are *not* accurately entered: true matches may be missed.

As HES data from different years are known to contain errors [132, 133], and because the probability of a single row matching was extremely low (0.1161% for 1995 and 0.4355% in 2006, equation 5.2), it was felt that the use of blocking variables would reduce the chances of true matches being identified too much, hence they were not used.

However, due to the projected large number of comparison pairs data sets (384,440,012 in 1995 and 1,736,352,750 in 2006), it was not possible to match the EPICure and entire HES data sets at the same time. This was due to limitations with R, which maps "objects" to memory and consequently restricts the size of the data that may be handled; however, a number of packages have been developed to help deal with this problem. [197]

Instead, for each epoch, the entire EPICure data were sequentially matched with a single day's worth of HES data at a time; e.g. in 2006, the EPICure data set was first matched with HES data for babies recorded as being born on the first of January, then with the babies recorded as being born on the second of January, then the third of January and so on until the end of the year. For 1995, the same process was followed, but matching started from March 1st; for both years, subjects without a date of birth recorded in HES were recoded as being born on "day 0" and then included in the linkage.

To achieve this work flow, a bespoke function was written making use of the "RecordLinkage" [143] and "ff" [198] packages from R. Furthermore, the required out-

put was specified in advance: either the list of calculated weights for the entire set of matches *or* the ID numbers for linked pairs above a predefined threshold were produced. This restricted the size of the data actually being handled by the computer's CPU at any one time to manageable proportions, while writing anonymised results to disc for future analysis. For each analysis, the function was therefore run twice: first, to obtain the range of weights and numbers of individual values with which to estimate the thresholds for possible links, and then a second time to obtain the corresponding IDs in each data set for weights above the set threshold.

5.3.3 Linkage criteria

Matching was performed for both study epochs in the same way. Each of the three matching algorithms available in the "RecordLinkage" package were used.[143] The most straight forward of these calculates weights (w) stochasticly, based on Fellegi and Sunter's work, whereby both the M probability (i.e. that both records of a pair are from the same subject) and U probabilities (where records in a pair belong to different subjects) are specified in advance. [143] The calculations are performed as follows:

$$w = \left\{ \begin{array}{ll} \log_2(M/U) & \text{if records are the same;} \\ \log_2(1-m)/(1-u) & \text{if records are different.} \end{array} \right\} \tag{5.3}$$

The values that are chosen for the M and U probabilities may have an important impact on the results of the data linkage. Consequently, they should be chosen carefully, ideally guided by prior knowledge about the data sets and chances of successfully matching subjects, but where this is not possible, best guesses may be used. Prior work by Dattani et al [132] provides some data on which the 2006 estimates of these values may be based. However, as not all of the variables to be used for matching had prior estimates, it was decided to perform one round of matching using best-guess "guestimated" values, and a second round of matching using the Dattani et al estimates. The guestimates were derived using the following rules:

***M*-probability** based on the estimated accuracy of record completion: for example, sex is likely to be recorded extremely accurately in both data sets, hence there will be a higher probability that an identified match is a true match, whereas there are more opportunities for errors to be introduced when recording gestational age and thus the likelihood of an identified match being a true match is lower.

***U*-probability** based on chance agreement: the likelihood that two subjects would match if the subjects were chosen randomly.

For the M-probabiliities, date of birth, mother's age at delivery, baby sex and number of babies were considered to have a high probability ($\geq 90\%$) of having been entered correctly; for other variables, the estimated probabilities varied as low as 20%. Best guess U-probabilities for date of birth and death were set at $1/365 = 0.00274$, and for discharge date, 1/500, as HES is likely to be discrepant from EPICure data in this respect; for birth order, number of babies and number of previous pregnancies at 90% as pregnancies of lower birth are more common, as are lower parity women; and sex at 0.49 so as to account for those of indeterminate sex. Gestational age at birth and maternal age were based on approximate number of categories with a slight adjustment for unequal distributions. Birth weight was assigned a U-probability of 1/1000, i.e. 0.001. The full set of values, along with corresponding weights, are shown in table 5.2.

In the comparison round of matching, using the Dattani estimates, data were available for date of birth, postcode, number of babies in the pregnancy, sex, birth weight, gestational age and ethnicity; of these, absolute numbers were provided for number of concordant and discordant pairs for number of births per pregnancy and sex, and percentages of concordant pairs for the remaining variables. It was therefore possible to calculate probabilities for these variables using equations 5.4 and 5.5 (C = concordance rate, D = discordance rate, and P_{nm} = percentage not missing):

$$M = CP_{\mathrm{nm}} \tag{5.4}$$

$$U = CDP_{\mathrm{nm}} \tag{5.5}$$

Where no prior information was available from the Dattani et al estimates for variables to be used in the matching, the guestimated values were used in supplement.

The second method of matching uses the algorithm designed by Contiero, on which the *EpiLink* software is based. [143] For this method, the overall weight (w_o) for each subject-pair can be calculated as:

$$w_o = \frac{\sum w_i s_i(x_i^1, x_i^2)}{\sum w_i} \tag{5.6}$$

Table 5.2: Probability estimates for linkage analyses between Hospital Episode Statistics (HES) and EPICure data based on "best guess" (guestimate) and prior knowledge (adapted from data linkage performed by Dattani et al between HES and NHS Numbers 4 Babies data sets).[132]

Matching variable	Baseline "guestimate"				Dattani et al [132] estimate			
	m	u	w_{m} [a]	w_{nm} [b]	m	u	w_{m} [a]	w_{nm} [b]
Date of birth	0.90	0.00274	5.794	-2.3	0.7405	0.0015	6.202	-1.347
GA at birth	0.80	0.02	3.689	-1.589	0.4941	0.0494	2.3028	-0.6308
Sex	0.999	0.49	0.7123	-6.2344	0.7208	0.0062	4.756	-1.270
Discharge date	0.20	0.002	4.6052	-0.2211	—	—	—	—
Date of death [a]	0.20	0.00274	4.2904	-0.2204	0.30	0.002	5.0106	-0.3547
Birth weight	0.60	0.001	6.3969	-0.9153	0.7405	0.0074	4.606	-1.342
Birth order	0.87	0.95	-0.08797	0.95551	0.8153	0.0033	5.510	-1.686
Delivery method [a]	0.80	0.80	0	0	0.67	0.1	1.902	-1.003
Ethnic category	0.20	0.10	0.6931	-0.1178	0.7308	0.095	2.040	-1.212
Mother's age at delivery	0.95	0.05	2.944	-2.944	—	—	—	—
Mother's date of birth	0.90	0.0001	9.105	-2.302	—	—	—	—
Postcode	0.90	0.001	6.802	-2.302	0.9291	0.065	2.660	-2.579
Number of previous pregnancies	0.60	0.90	-0.4055	1.3863	—	—	—	—
Number of babies	0.95	0.95	0	0	0.8153	0.0033	5.510	-1.686

[a] w_{m} = weight if pairs match.

[b] w_{nm} = weight if pairs do not match.

[c] Date of death and delivery method were both modified using an adjusted "best guestimate" for the second linkage analysis performed using estimates from Dattani et al.

where s_i is the value of the comparison between the ith records from each of the data sets x and y, and w_i is the weight attached to that particular (variable) comparison. In turn, this is calculated from the frequency of responses (f) and estimated error rates (e):

$$w_i = \frac{\log_2(1-e)}{f} \tag{5.7}$$

Default values for this comparison within the RecordLinkage package are set at 0.1 for the error rates and the inverse of the unique number of responses to each value for the frequency. As the data in the current matching exercise were analysed sequentially by day, both the error rates and frequencies were explicitly set according to the default

values for the overall data sets.

The final method of matching uses an automated method to assign weights based on maximum likelihood, and is known as the *Estimation-Maximisation (EM)* algorithm. [143] This did not require any parameters other than the names of the data sets to be passed to it.

5.3.4 Missing data and sensitivity analyses

Probabilistic linkage is inherently designed for data sets where there are concerns surrounding data completeness and/or quality. Subject variables with missing data are assigned a weight of zero, and corresponding weights for non-missing data may be adjusted in line with the levels of missingness identified. For this study, it was originally intended to perform sensitivity analyses using different values for the available parameters. Due to the presence of large quantities of missing data, I planned to vary the probability estimates used in the Fellegi and Sunter analyses, and to alter the estimated error rates used in the Contiero and EM analyses. However, due to time limitations, it was not possible to carry this out and the only direct comparison (i.e. using different parameters) was between the two versions of the stochastic (classical Fellegi and Sunter) linkage – using the guestimate and Dattani probabilities.

5.3.5 Thresholds

It was initially planned that thresholds would be estimated from graphs of the weight distributions. However, in several of the analyses, the resultant data set – containing just a single variable but with the number of rows equivalent to the N_{TRP} – was too large to fit into memory and so could not be processed. Instead, data were tabulated (using the `table.ff` command from "ff"[198]) to identify appropriate cut off points. For the estimation-maximisation analyses, this value was confined to being between 0 and 1; for the other analyses, there were no restrictions.

5.3.6 Clerical review

Following linkage, a master data set was created of "true" matches. This was done by appending the four data sets (one from each of the different analyses) of retained ID pairs from each time period together into a single data set. The combined data

sets were then reduced to only contain rows with unique ID pairs, and merged with the respective original EPICure and HES data sets for the epoch using the unique identification numbers that were generated when the data sets were originally reduced in size for data linkage.

After the creation of a combined data set for each year, key variables were chosen and rows corresponding to duplicate entries of a single EPICure ID were manually reviewed. This allowed each row within a set of potential matches to be marked as "true match", "not a match", or "undecided". In the case of a true match, that row and all other potential matches with either the EPICure or HES ID were removed from further consideration; for rows marked as "not a match", only that row was removed; and in the "undecided" case, the row remained in the data set for future consideration. The review process was repeated using both the EPICure and HES subjects as the base for the comparison, and with different supplementary data items, until it was no longer possible to discriminate any further true matches.

5.4 Error measures

Using the reviewed data set of true matches, it was possible to assess the accuracy of the matching processes. For each method of linkage (Fellegi and Sunter – using both the baseline guestimates and Dattani estimates; EpiLink (Contiero); EM), the true matches were merged with the saved unique pairs. This permitted the number of "true matches" to be identified. Knowing the number of true matches and the total number of linked pairs above a set cut weight, as well as the total number of row pairs and the maximum number of matches, it was possible to construct a table as shown in figure 5.2.

From these values, it was then possible to derive the following measures of error:

Sensitivity Proportion of true matches correctly identified

$$= \frac{N_{\mathrm{TL}}}{N_{\mathrm{M}}} \tag{5.8}$$

Specificity proportion of non-matches that are correctly identified as non-links:

$$= \frac{N_{\mathrm{NM}}}{N_{\mathrm{TNL}}} \tag{5.9}$$

Postive predictive value (PPV) Proportion of links which are true matches:

$$= \frac{N_{\mathrm{TL}}}{N_{\mathrm{L}}} \tag{5.10}$$

Negative predictive value (NPV) Proportion of non-linkes that are not matches:

$$= \frac{N_{\mathrm{TNL}}}{N_{\mathrm{NL}}} \tag{5.11}$$

These measures were calculated for each type of the matching algorithm used, to facilitate comparison between the techniques, and to guide future attempts at linkage using either of the data sets involved (particularly, the HES data).

5.5 Summary of linkage methods

This chapter described the methods used for the AVS, designed to compare the EPICure data sets with Hospital Episode Statistics. In turn, the linkage exercise was conducted in order to assess completeness – both of the EPICure and HES data sets – and to garner additional information with which to investigate an observed increase in admissions of extremely premature babies to neonatal intensive care between 1995 and 2006.

After discussion of the study context, in section 5.1, methods on data linkage were presented. This may be performed in two ways: deterministicly, merging the data sets according to a predetermined selection of variables, or probabilisticly. The former approach was employed by the NHS HSCIC and is briefly described; however, the primary focus of investigation was on probabilistic linkage.

In preparation, all data sets underwent two stages of cleaning: the first to standardise the variables between EPICure and HES, and the second to restrict the data to eligible subjects as well as the precise variables required for the matching. This was because the data were of sufficient size to be affected by available computer resources (particularly, memory and processing power). Linkage was performed using the classical Fellegi and Sunter algorithm with both "best guess" estimates as well as estimates based on findings from a similar matching exercise performed by Dattani and colleagues.[132] Further analyses were performed using the Contiero and EM algorithms.[143]

Full methods of the probabilistic linkage are discussed in great detail in sections 5.3.1 – 5.3.6, including choice of thresholds and clerical review. In section 5.4, there follows detail of the error measures that were used to assess the results.

Figure 5.2: Known and calculated values for matching algorithms, used in assessment of linkage error. Data linkage is performed by pairing data from two data sets, followed by manual verification of linked pairs to identify true matches. Values for cells were identified in the following manner:
(1) The total number of row pairs, maximum number of matches, total number of linked pairs and number of true matches within those linked pairs were identified.
(2) The numbers of false links, false non-links, total non-links and number of non-matches were then derived.
(3) Finally, the true number of non-matches among the non-linked pairs was calculated.

Identified by linkage algorithm		Identified by manual review: **Match**	**Non-match**	**Total**
	Link	**(1)** True link:[a] N_{TL}	**(2)** False link: $N_{\mathrm{FL}} = N_{\mathrm{L}} - N_{\mathrm{TL}}$	**(1)** Total linked:[b] N_{L}
	Non-link	**(2)** False non-link: $N_{\mathrm{FNL}} = N_{\mathrm{M}} - N_{\mathrm{TL}}$	**(3)** True non-link: $N_{\mathrm{TNL}} = N_{\mathrm{NM}} - N_{\mathrm{FL}}$	**(2)** Total non-links: $N_{\mathrm{NL}} = N_{\mathrm{TRP}} - N_{\mathrm{L}}$
	Totals	**(1)** Number of matches:[c] N_{M}	**(2)** Number of non-matches: $N_{\mathrm{NM}} = N_{\mathrm{TRP}} - N_{\mathrm{M}}$	**(1)** Total number of row pairs:[d] N_{TRP}

[a] Identified by merging paired IDs from the data set of true matches with the results of the matching exercise.
[b] The number of unique paired rows identified by the matching exercise.
[c] Assumed to be equal to the size of the smallest data set.
[d] Calculated according to the formula shown in equation 5.1.

Chapter 6

Linkage Results

This chapter presents the results of the linkage analysis performed for the Admissions Validation Study (AVS). The available data are summarised, with a brief discussion about their quality focused on the Hospital Episode Statistics (HES) data. Results from each of the four main linkage comparisons are presented in section 6.3. This is followed by descriptive results from the manual review of possible links, then by the results from the assessment of error (section 6.5). Deterministic linkage results, performed by the NHS Health and Social Care Information Centre (HSCIC), are presented, along with final results describing the saved HES data prior to data destruction. At the end, there is a summary of the chapter.

6.1 Available data

6.1.1 EPICure

The 1995 EPICure data had been previously split up to separate out confidential data (name, home address) from a pseudonymised, working data set. Unfortunately, this meant some data had been lost over the years. Hence, there were several available data sets of varying sizes, shown in table 6.1, containing different variables; none – and no combination – contained the complete data set of 4,004 births described by Costeloe et al [54]. After merging the available data and restricting the data to those births known to have occurred in England, full data were available from 668 subjects. This corresponds to the 666 babies described in the EPICure papers who were born in

Table 6.1: Available data sets from the EPICure (1995) study.

Data set name	N	Detail
'log95'	3,222	Brief data from the labour ward logbooks of all birth occurring in England (incomplete data set)
'CRF95'	811	Detailed data contained in the case review form for all babies admitted onto a neonatal intensive care unit.
'e1nhsno'	314	Contained only study ID and NHS number for surviving subjects
'e1data'	454	This was an incomplete administrative data set that contained a mixture of confidential and clinical information, that all proved to be duplicated within the other data sets.

England between 22 and 26 weeks gestation and subsequently admitted into neonatal intensive care [54, 55], plus a further two live births at 21 weeks gestation.

In 2006, the data were of a much higher quality. However, there were still two data sets available: the pseudonymised working data set, containing 4,103 records, and the administrative data set that contained 4,144 records; when the two were merged using the EPICure unique identifier, there were 4,145 rows. This was restricted to 3,376 by excluding records pertaining to women who had a termination of pregnancy (ToP). The data were then further restricted to exclude still births occuring below 24 weeks gestational age as these are not included in routine HES data collection; following this, there were 2750 records available for inclusion.

6.1.2 Hospital Episode Statistics

Hospital Episode Statistics (HES) data were supplied by the NHS HSCIC, with a separate file for each of the analysis years. There were a total of 575,509 records for 1995 and 631,499 for 2006. Births occurring in January or February were excluded from the 1995 data in order to match the time period of the EPICure study; 8,807 records with a missing date of birth were retained, meaning 486,705 records were used in the linkage analysis. There were no duplicate records in 1995; however, for 2006, there were 98 duplicate rows, meaning that 631,401 records were available for analysis.

6.2 Data quality

It was apparent upon receiving the HES data that they were of insufficient quality to provide the desired supplemental data for the EPICure cohorts. In the 1995 data set, socioeconomic data were completely absent, and fewer than 20% of the subjects had information on ethnicity. The 2006 data were better, with socioeconomic information available for over 50% of subjects, and ethnicity unavailable for only 157,781 (24.99%) records. Fifty seven percent of the population (364,902) were white, seven percent (44,681) Indian, Pakistani or Bengali, and 4.7% black (29,885).

As well as the lack of desired information, other requested variables were also missing, as described in section 5.3.1, and listed in appendix F. Particularly, there were no data on NHS number provided – this was most important for the matching process in the 2006 data set where the NHS numbers of EPICure 2 subjects *were* available.

6.2.1 Linkage variables

With respect to the variables chosen for data linkage, there were more data missing from HES in 2006 than 1995 (table 6.2). Most importantly, gestational age at birth was missing in 336,178 (53.23%) of the HES subjects in 2006, but only 164,006 (28.50%) in 1995; similarly, birth weight was missing in 288,014 (45.61%) of the 2006 subjects but only 152,641 (26.52%) subjects in 1995. In contrast, birth weight and gestational age were complete in the EPICure data, other than for seven subjects who were missing gestational age at birth in 2006. The 2006 linkage did benefit from the inclusion of postcode in the matching criteria, although this was present for only 340,939 (54.00%) of HES subjects; it was missing from just one record in the EPICure 2 data set.

There were an average of 47,225 births recorded per month in 1995, and 52,261 in 2006 – an increase of 10.70%; births were equally spread throughout the year. In the HES data in 1995, there were three categories for "sex" – *male*, *female* and *indeterminate*; however, only male and female were used in 2006; males outnumbered females in all data sets (table 6.3) with only nine observations missing from EPICure data, in 2006. Discharge date was missing from 16,912 (2.93%) HES observations in 1995 but was not thought to be very useful because the HES data describe "episodes" rather than complete duration of hospital stay; furthermore, it was unclear if the discharge dates represented discharge home from hospital, or simply transfer to a different

Table 6.2: Variables in each of the Hospital Episode Statistics (HES) and EPICure data sets that were used for matching in 1995 and 2006 and their levels of *missingness*. (**HES (1995)** = 575,509 (for the entire year); **EPICure** $n = 668$ (March – December); **EPICure 2** $n = 2,750$; **HES (2006)** $n = 631,401$.)

Variable	HES 1995 Missing (%)	EPICure (1995) Missing (%)	HES 2006 Missing (%)	EPICure-2 (2006) Missing (%)
Date of birth	8807 (1.53)	0 (0.00)	4265 (0.68)	0 (0.00)
GA at birth	164006 (28.50)	0 (0.00)	336178 (53.23)	7 (0.25)
Sex	2616 (0.45)	0 (0.00)	3202 (0.51)	9 (0.33)
Discharge date	16912 (2.94)	373 (55.84)	—	—
Date of death	571417 (99.29)	268 (40.12)	—	—
Birth weight	152641 (26.52)	0 (0.00)	288014 (45.61)	26 (0.95)
Birth order	250718 (43.56)	0 (0.00)	224632 (35.57)	0 (0.00)
Delivery method	168018 (29.19)	1 (0.15)	—	—
Number.of babies	152378 (26.48)	0 (0.00)	209455 (33.17)	0 (0.00)
Previous pregnancies number	—	—	618692 (97.97)	101 (3.67)
Ethnic category	462999 (80.45)	0 (0.00)	—	—
Postcode	—	—	290462 (46.00)	1 (0.04)
Mother's dob	—	—	273426 (43.30)	2750 (100.00)
Mother's age at delivery	214999 (37.36)	4 (0.60)	273430 (43.30)	8 (0.29)

hospital. Using "discharge date" and "discharge method" from the HES data, it was possible to construct date of death – however, this applied to under one percent of the 1995 population (table 6.2). For the EPICure population in 1995, discharge date was recorded for 373 of the 668 subjects (55.84%) and date of death was available for 268 (40.12%) subjects.

Most births HES in 1995 were recorded as following a singleton pregnancy (412,166, 71.62%), with only 10,965 (1.90%) from multiple order pregnancies, although a large proportion were missing data (26.48% – 152,378 records). Numbers were similar in 2006: 409,175 (64.80%) singleton pregnancies and 209,419 (33.17%) missing, with 12,807 (2.03%) following a multiple pregnancy. In conjunction with this, the majority of babies were first born, with only approximately 5,500 babies reported in the HES data for each year to have *not* been born first. However, there were again high levels of missing data (43.56% and 35.57% in 1995 and 2006, respectively). In the EPICure data, 562 (84.13%) and 2428 (88.29%) were first born babies in 1995 and 2006, respectively.

Table 6.3: Spread of values amongst matching variables used in the linkage exercises performed between EPICure and Hospital Episode Statistics (HES) data for 1995 and 2006. (**EPICure** $n = 668$ (March – December); **EPICure 2** $n = 2,750$; **HES (1995)** $= 575,509$ for the entire year; **HES (2006)** $n = 631,401$.)

Variable	**HES 1995**	**EPICure**	**HES 2006**	**EPICure 2**
Level	**N (%)**	**N (%)**	**N (%)**	**N (%)**
Month of birth				
Jan	45804 (7.96)	—	50192 (7.95)	231 (8.40)
Feb	43000 (7.47)	—	46817 (7.41)	209 (7.60)
Mar	47591 (8.27)	56 (8.38)	51436 (8.15)	234 (8.51)
Apr	46631 (8.10)	66 (9.88)	50261 (7.96)	240 (8.73)
May	49634 (8.62)	66 (9.88)	54118 (8.57)	218 (7.93)
Jun	49524 (8.61)	76 (11.38)	53277 (8.44)	264 (9.60)
Jul	50048 (8.70)	62 (9.28)	54016 (8.55)	225 (8.18)
Aug	49236 (8.56)	65 (9.73)	55377 (8.77)	210 (7.64)
Sep	49224 (8.55)	55 (8.23)	55725 (8.83)	263 (9.56)
Oct	48975 (8.51)	69 (10.33)	56435 (8.94)	238 (8.65)
Nov	45475 (7.90)	66 (9.88)	52284 (8.28)	205 (7.45)
Dec	41560 (7.22)	87 (13.02)	47198 (7.48)	213 (7.75)
Missing	8807 (1.53)	0 (0.00)	4265 (0.68)	0 (0.00)
GA at birth (weeks)				
<20	0 (0.00)	—	95 (0.02)	24 (0.87)
20	28 (0.00)	—	36 (0.01)	34 (1.24)
21	44 (0.01)	2 (0.30)	66 (0.01)	66 (2.40)
22	82 (0.01)	16 (2.40)	87 (0.01)	151 (5.49)
23	156 (0.03)	115 (17.22)	212 (0.03)	322 (11.71)
24	369 (0.06)	244 (36.53)	384 (0.06)	637 (23.16)
25	398 (0.07)	291 (43.56)	682 (0.11)	686 (24.95)
26-29	2519 (0.44)	—	6446 (1.02)	820 (29.82)
30-34	10727 (1.86)	—	7813 (1.24)	3 (0.11)
35-39	173806 (30.20)	—	129294 (20.48)	0 (0.00)
40+	223374 (38.81)	—	150170 (23.78)	0 (0.00)
Missing	164006 (28.50)	0 (0.00)	336116 (53.23)	7 (0.25)
Sex				
Female	278267 (48.35)	359 (53.74)	307080 (48.63)	1422 (51.71)
Male	292844 (50.88)	309 (46.26)	321119 (50.86)	1319 (47.96)
Indeterminate	1782 (0.31)	—	—	—
Missing	2616 (0.45)	0 (0.00)	3202 (0.51)	9 (0.33)
Birth weight (g)				
<500	2634 (0.46)	27 (4.04)	520 (0.08)	455 (16.55)
500-999	1896 (0.33)	638 (95.51)	1904 (0.30)	2117 (76.98)
1000-1499	3455 (0.60)	3 (0.45)	2833 (0.45)	149 (5.42)
1500-1999	6263 (1.09)	0 (0.00)	5913 (0.94)	3 (0.11)
2000-2499	19528 (3.39)	0 (0.00)	17477 (2.77)	0 (0.00)
2500-2999	72098 (12.53)	—	59906 (9.49)	—

Continued on next page...

Table 6.3: (Continued) Spread of values amongst matching variables used in the linkage exercises performed between EPICure and Hospital Episode Statistics (HES) data for 1995 and 2006. (**EPICure** $n = 668$ (March – December); **EPICure 2** $n = 2,750$; **HES (1995)** $= 575,509$ for the entire year; **HES (2006)** $n = 631,401$.)

Variable	HES 1995	EPICure	HES 2006	EPICure 2
Level	**N (%)**	**N (%)**	**N (%)**	**N (%)**
3000-3499	152613 (26.52)	—	121699 (19.27)	—
3500-3999	119605 (20.78)	—	96202 (15.24)	—
4000-4499	38222 (6.64)	—	31414 (4.98)	—
4500-4999	5860 (1.02)	—	4944 (0.78)	—
5000+	694 (0.12)	—	622 (0.10)	—
Missing	152641 (26.52)	0 (0.00)	287967 (45.61)	26 (0.95)
Birth order				
1	318895 (55.41)	562 (84.13)	401473 (63.58)	2428 (88.29)
>1	5896 (1.02)	106 (15.87)	5338 (0.85)	322 (11.71)
Missing	250718 (43.56)	0 (0.00)	224590 (35.57)	0 (0.00)
Maternal age (years)				
<16	940 (0.16)	3 (0.45)	682 (0.11)	12 (0.44)
16-19	21343 (3.71)	49 (7.34)	23494 (3.72)	219 (7.96)
20-29	190184 (33.05)	326 (48.80)	162499 (25.74)	1260 (45.82)
30-39	138764 (24.11)	276 (41.32)	158740 (25.14)	1116 (40.58)
40-49	9255 (1.61)	10 (1.50)	12583 (1.99)	130 (4.73)
50+	24 (0.00)	0 (0.00)	35 (0.01)	5 (0.18)
Missing	214999 (37.36)	4 (0.60)	273368 (43.30)	8 (0.29)

The majority of mothers, approximately 90% in all data sets, were aged between 20 and 40 years old – although approximately 40% of HES entries were missing data in both years (the proportions shown in table 6.3 differ as they report the percentage of the entire data set, rather than percentage of complete entries).

Ethnicity data were not used for linkage in 2006, but in 1995, greater than 80% of HES records contained missing data (462,999 records in total); of the remainder, 78,195 were white, 12,640 black, 10,043 Indian, Pakistani or Bengali and the remaining 11,632 were described as "other". Delivery method was also adapted for use in 1995: 340,483 (59.16%) babies were recorded as being born vaginally, 67,008 (11.64%) by Caesarean section (39,057 (58.29%) after the onset of labour) and for 168,018 (29.19%) babies, delivery method was not recorded. There were no missing records for ethnicity and only one record missing for delivery method in the EPICure data in 1995.

6.2.2 Data consistency

Issues were also identified relating to the internal consistency of the HES data. While data accuracy overall within the data sets may be extremely high,[132, 133] at the extremes of gestation with which this study was dealing there were obvious errors – for example, in 1995, 2,184 (82.92%) of 2,634 subjects with a recorded birth weight of less than 500 grams were described as having a gestational age in the range 35-45 weeks, as shown in table 6.4. Furthermore, there was a higher proportion of missing data for those recorded as being of a low gestational age in 1995: for instance, birth weight was missing from 14.3%, 11.4% at 20 and 21 weeks gestational age, from 6.2-7.5% between 22 and 25 weeks, and only 0.2% of those born between 35 and 39 weeks. By comparison, in 2006, there were problems across the gestational age range included in the data set, with levels of missing birth weight data running at 20.0% below 22 weeks, 18.3%, 13.5%, 8.8% and 9.2% at 22-25 weeks, respectively, and greater than 20% at all gestations from 30 weeks up (table 6.5). Similar issues do not affect the EPICure data: these data sets were extensively checked at the time of collection.[54, 55]

6.3 Main comparisons

As can be seen using equation 5.1, there were $477,898 \times 668 = 325,118,940$ potential comparison pairs in 1995, and $631,401 \times 2,750 = 1,736,352,750$ pairs in 2006. It was not possible (or desirable) to save all this information as the vast majority (i.e. the total number of row pairs minus the maximum number of matches) were false matches. Therefore, each linkage method required a preliminary review of the calculated weights in order to select appropriate cut-offs above which to retain linked or potentially linked data pairs (one each from the HES and EPICure data sets). Cut-off points were selected according to where a "reasonable" number of linked pairs was obtained.

Of the three methods (Fellegi-Sunter technique, using probabilities from both the guestimate and Dattani et al estimates, EpiLink approach, EM algorithm), the EM approach to calculating weights was by far the most time-consuming (requiring approximately 10x as much CPU time). Therefore, analysis predominantly focused on using the other two methods – including varying the weights between the "guestimate" probabilities and those derived from Dattani *et al.*

Table 6.4: Numbers of subjects (percentages of overall data set) according to birth weight (g) by gestational age (weeks), as recorded in the 1995 Hospital Episode Statistics data set.

Birth weight category	**Gestational age (weeks)**											
	20	**21**	**22**	**23**	**24**	**25**	**26-29**	**30-34**	**35-39**	**40+**	**Missing**	**Total**
<500	19 (0.003)	23 (0.004)	39 (0.007)	36 (0.006)	53 (0.009)	40 (0.007)	117 (0.020)	91 (0.016)	960 (0.167)	1224 (0.213)	32 (0.006)	2634 (0.458)
500-999	3 (0.001)	7 (0.001)	29 (0.005)	103 (0.018)	273 (0.047)	292 (0.051)	864 (0.150)	181 (0.031)	36 (0.006)	19 (0.003)	89 (0.015)	1896 (0.329)
1000-1499	0 (0.000)	1 (0.000)	0 (0.000)	1 (0.000)	6 (0.001)	20 (0.003)	1179 (0.205)	1590 (0.276)	326 (0.057)	180 (0.031)	152 (0.026)	3455 (0.600)
1500-1999	0 (0.000)	2 (0.000)	1 (0.000)	0 (0.000)	4 (0.001)	2 (0.000)	141 (0.025)	3785 (0.658)	2010 (0.349)	94 (0.016)	224 (0.039)	6263 (1.088)
2000-2499	0 (0.000)	0 (0.000)	0 (0.000)	1 (0.000)	0 (0.000)	2 (0.000)	26 (0.005)	3367 (0.585)	13684 (2.378)	1831 (0.318)	617 (0.107)	19528 (3.393)
2500-2999	0 (0.000)	0 (0.000)	1 (0.000)	0 (0.000)	0 (0.000)	5 (0.001)	24 (0.004)	1053 (0.183)	47425 (8.241)	21521 (3.739)	2069 (0.360)	72098 (12.528)
3000-3499	0 (0.000)	5 (0.001)	2 (0.000)	3 (0.001)	6 (0.001)	2 (0.000)	30 (0.005)	334 (0.058)	67091 (11.658)	80740 (14.029)	4400 (0.765)	152613 (26.518)
3500-3999	2 (0.000)	1 (0.000)	4 (0.001)	1 (0.000)	3 (0.001)	5 (0.001)	13 (0.002)	117 (0.020)	33709 (5.857)	82270 (14.295)	3480 (0.605)	119605 (20.782)
4000-4499	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	5 (0.001)	32 (0.006)	7202 (1.251)	29943 (5.203)	1040 (0.181)	38222 (6.641)
4500-4999	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	2 (0.000)	4 (0.001)	916 (0.159)	4763 (0.828)	175 (0.030)	5860 (1.018)
5000+	0 (0.000)	0 (0.000)	0 (0.000)	1 (0.000)	1 (0.000)	0 (0.000)	8 (0.001)	4 (0.001)	132 (0.023)	520 (0.090)	28 (0.005)	694 (0.121)
Missing	4 (0.001)	5 (0.001)	6 (0.001)	10 (0.002)	23 (0.004)	30 (0.005)	110 (0.019)	169 (0.029)	315 (0.055)	269 (0.047)	151700 (26.359)	152641 (26.523)
Total	28 (0.005)	44 (0.008)	82 (0.014)	156 (0.027)	369 (0.064)	398 (0.069)	2519 (0.438)	10727 (1.864)	173806 (30.200)	223374 (38.813)	164006 (28.498)	575509 (100.000)

Table 6.5: Numbers of subjects (percentage of overall data set) according to birth weight (g) by gestational age (weeks), as recorded in the 2006 Hospital Episode Statistics data set.

Birth weight category	**Gestational age (weeks)**													
	< 20	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27-29**	**30-34**	**35-39**	**40+**	**Missing**	**Total**
<500	16 (0.003)	17 (0.003)	28 (0.004)	28 (0.004)	21 (0.003)	35 (0.006)	25 (0.004)	15 (0.002)	28 (0.004)	14 (0.002)	64 (0.010)	86 (0.014)	143 (0.023)	520 (0.082)
500-999	10 (0.002)	1 (0.000)	3 (0.000)	19 (0.003)	101 (0.016)	218 (0.035)	226 (0.036)	213 (0.034)	345 (0.055)	86 (0.014)	11 (0.002)	9 (0.001)	662 (0.105)	1904 (0.302)
1000-1499	13 (0.002)	11 (0.002)	14 (0.002)	4 (0.001)	5 (0.001)	6 (0.001)	14 (0.002)	41 (0.006)	685 (0.108)	934 (0.148)	92 (0.015)	23 (0.004)	991 (0.157)	2833 (0.449)
1500-1999	0 (0.000)	0 (0.000)	8 (0.001)	17 (0.003)	22 (0.003)	13 (0.002)	31 (0.005)	4 (0.001)	86 (0.014)	2355 (0.373)	1231 (0.195)	70 (0.011)	2076 (0.329)	5913 (0.936)
2000-2499	1 (0.000)	0 (0.000)	1 (0.000)	4 (0.001)	15 (0.002)	37 (0.006)	89 (0.014)	64 (0.010)	82 (0.013)	2038 (0.323)	8289 (1.313)	1015 (0.161)	5842 (0.925)	17477 (2.768)
2500-2999	1 (0.000)	2 (0.000)	1 (0.000)	0 (0.000)	6 (0.001)	13 (0.002)	137 (0.022)	216 (0.034)	667 (0.106)	568 (0.090)	27750 (4.395)	11464 (1.816)	19081 (3.022)	59906 (9.488)
3000-3499	3 (0.000)	2 (0.000)	2 (0.000)	0 (0.000)	3 (0.000)	6 (0.001)	78 (0.012)	222 (0.035)	1580 (0.250)	152 (0.024)	38932 (6.166)	41831 (6.625)	38888 (6.159)	121699 (19.274)
3500-3999	1 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	1 (0.000)	14 (0.002)	87 (0.014)	1261 (0.200)	61 (0.010)	20293 (3.214)	43579 (6.902)	30905 (4.895)	96202 (15.236)
4000-4499	4 (0.001)	0 (0.000)	0 (0.000)	1 (0.000)	0 (0.000)	1 (0.000)	7 (0.001)	23 (0.004)	405 (0.064)	10 (0.002)	4516 (0.715)	16275 (2.578)	10172 (1.611)	31414 (4.975)
4500-4999	2 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	1 (0.000)	0 (0.000)	60 (0.010)	1 (0.000)	630 (0.100)	2684 (0.425)	1566 (0.248)	4944 (0.783)
5000+	12 (0.002)	1 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	2 (0.000)	0 (0.000)	1 (0.000)	3 (0.000)	4 (0.001)	96 (0.015)	292 (0.046)	211 (0.033)	622 (0.099)
Missing	32 (0.005)	2 (0.000)	9 (0.001)	14 (0.002)	39 (0.006)	52 (0.008)	60 (0.010)	87 (0.014)	271 (0.043)	1590 (0.252)	27390 (4.338)	32842 (5.201)	225579 (35.727)	287967 (45.608)
Total	95 (0.015)	36 (0.006)	66 (0.010)	87 (0.014)	212 (0.034)	384 (0.061)	682 (0.108)	973 (0.154)	5473 (0.867)	7813 (1.237)	129294 (20.477)	150170 (23.784)	336116 (53.233)	(100.000)

6.3.1 Baseline estimated values

In 1995, the maximum weight of a linked pair was 42.3, and a weight of 15 was chosen for the cut-off value; above this, 2,093 unique pairs were identified, representing 537 EPIcure IDs and 1,846 HES IDs. The number of matches dropped markedly above a cut-off of 17, to 792 unique pairs with 365 unique EPICure IDs and 692 unique HES IDs (table 6.7). This drop-off can be seen in the density graph of weights (i.e. the area under the curve represents the number of pairs of linked records at each weight; the analysis was coded as "fs.D") shown in figure 6.1a, and in the number of unique records linked from each data set seen in figure 6.2a. Above a weight of 30, the number of linked pairs was the same as the number of IDs from each data set – i.e. there were 86 uniquely matched pairs with no repeated IDs from either EPICure or HES data sets.

The maximum weight in the 2006 analysis was 54.51, and a cut-off value of 10 was initially chosen for the weights: graphs of the data for 2006 are shown in figures 6.1b (density graph) and 6.2b (unique IDs). Note that for both figures 6.1 and 6.2, axes vary in scale between the years. Above this cut-off value, there were 44,719 unique pairs of records identified, which represented 2,729 individual EPICure 2 IDs and 36,025 HES IDs. A large decrease in the number of linkages was then seen above a cut-off value of 12, the number of linkages was greatly reduced - to 2,459 pairs overall, and 1,569 and 1,811 individual EPICure 2 and HES IDs, respectively (table 6.6).

6.3.2 Dattani estimates

Using the estimated probabilities obtained from Dattani et al [132] – listed in table 5.2 – meant the maximum weights obtained were higher: in 1995, it was 65.7, whereas in 2006 it was 71.57. The initial cut-off values chosen were, respectively, 35 and 15, yet this still provided extremely high numbers of potential matches, particularly in 2006 where 53,413 pairs were potentially linked. These values were chosen because there was a sharp attenuation in both data sets at a value at or just above the cut-off. In 1995, this resulted in a relatively constant decrease in the number of EPICure IDs with increasing weight throughout the weight range encompassed by the retained data, whereas the number of potentially linked HES IDs decreased drastically from 16,385 to just 3,540 at a weight of 36. In 2006, there was a large decrease in the number of linked pairs from a value of 19, with the number of HES IDs matched dropping from 32,051 at

Table 6.6: Table of the number of pairs matched in 2006 from each data set for differing cutoffs in the value of the weight calculated by the Fellegi-Sunter (guestimate) method of data linkage.

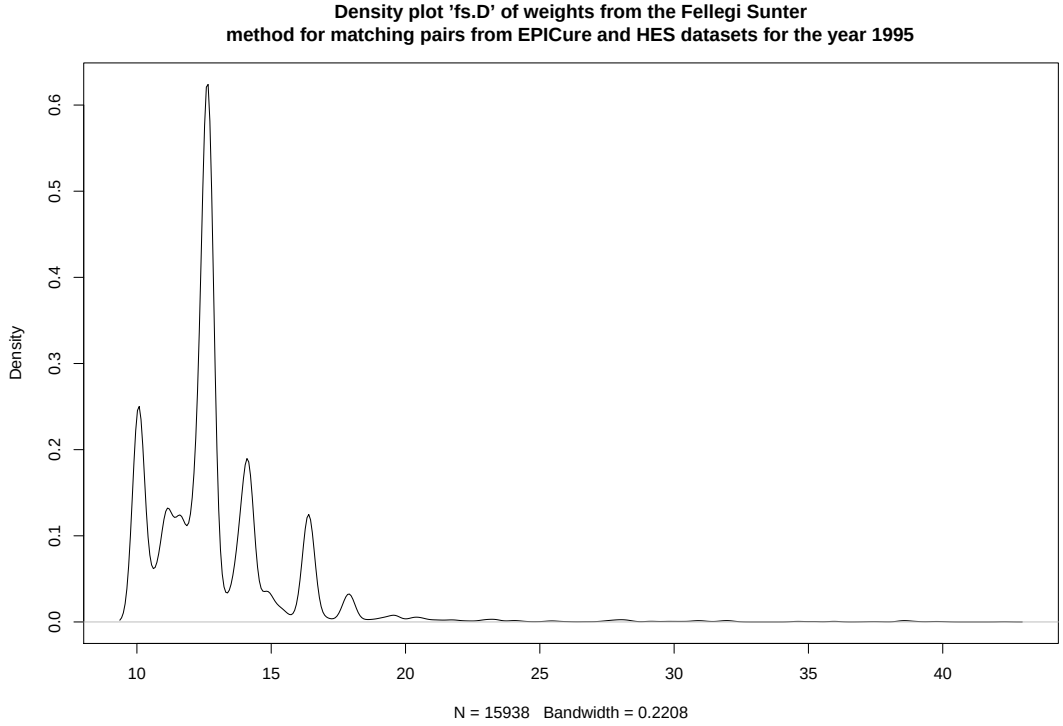
Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
10	44719	2729	99.24	36025	5.71
11	40363	2701	98.22	34741	5.50
12	2459	1569	57.05	1811	0.29
13	2445	1567	56.98	1806	0.29
16	2130	1455	52.91	1608	0.25
17	2082	1435	52.18	1576	0.25
19	2072	1430	52.00	1572	0.25
20	2004	1394	50.69	1526	0.24
21	1969	1385	50.36	1517	0.24
22	1895	1362	49.53	1489	0.24
23	1886	1360	49.45	1488	0.24
24	1515	1128	41.02	1189	0.19
26	1499	1121	40.76	1181	0.19
27	1475	1119	40.69	1179	0.19
28	1474	1118	40.65	1178	0.19
29	1460	1108	40.29	1168	0.18
30	1411	1086	39.49	1149	0.18
31	1168	917	33.35	970	0.15
32	1165	915	33.27	968	0.15
33	1127	889	32.33	934	0.15
34	994	788	28.65	836	0.13
35	972	777	28.25	828	0.13
36	970	775	28.18	826	0.13
37	726	578	21.02	614	0.10
38	724	576	20.95	612	0.10
39	710	565	20.55	601	0.10
41	675	539	19.60	571	0.09
42	559	446	16.22	476	0.08
43	494	430	15.64	461	0.07
44	266	233	8.47	259	0.04
45	265	233	8.47	258	0.04
46	249	218	7.93	242	0.04
53	126	117	4.25	122	0.02
54	102	93	3.38	98	0.02
55	0	0	0.00	0	0.00

Table 6.7: Table of the number of pairs in 1995 matched from each data set for differing cutoffs in the value of the weight calculated by the Fellegi-Sunter (guestimate) method of data linkage.

Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
15	2093	537	80.39	1846	0.38
16	1939	528	79.04	1726	0.35
17	792	365	54.64	692	0.14
18	467	302	45.21	401	0.08
19	435	285	42.66	380	0.08
20	335	256	38.32	294	0.06
21	270	216	32.34	237	0.05
22	229	200	29.94	208	0.04
23	202	182	27.25	193	0.04
24	175	166	24.85	167	0.03
25	158	150	22.46	152	0.03
26	145	138	20.66	142	0.03
27	140	133	19.91	137	0.03
28	112	110	16.47	109	0.02
29	97	96	14.37	96	0.02
30	86	86	12.87	86	0.02
31	67	67	10.03	67	0.01
32	50	50	7.49	50	0.01
34	47	47	7.04	47	0.01
35	41	41	6.14	41	0.01
37	31	31	4.64	31	0.01
38	26	26	3.89	26	0.01
39	9	9	1.35	9	0.00
40	4	4	0.60	4	0.00
42	2	2	0.30	2	0.00
43	0	0	0.00	0	0.00

Figure 6.1: Density distribution of weights from the stochastic linkage analyses using guestimate probabilities. Axes are not to the same scale.

(a) 1995



(b) 2006

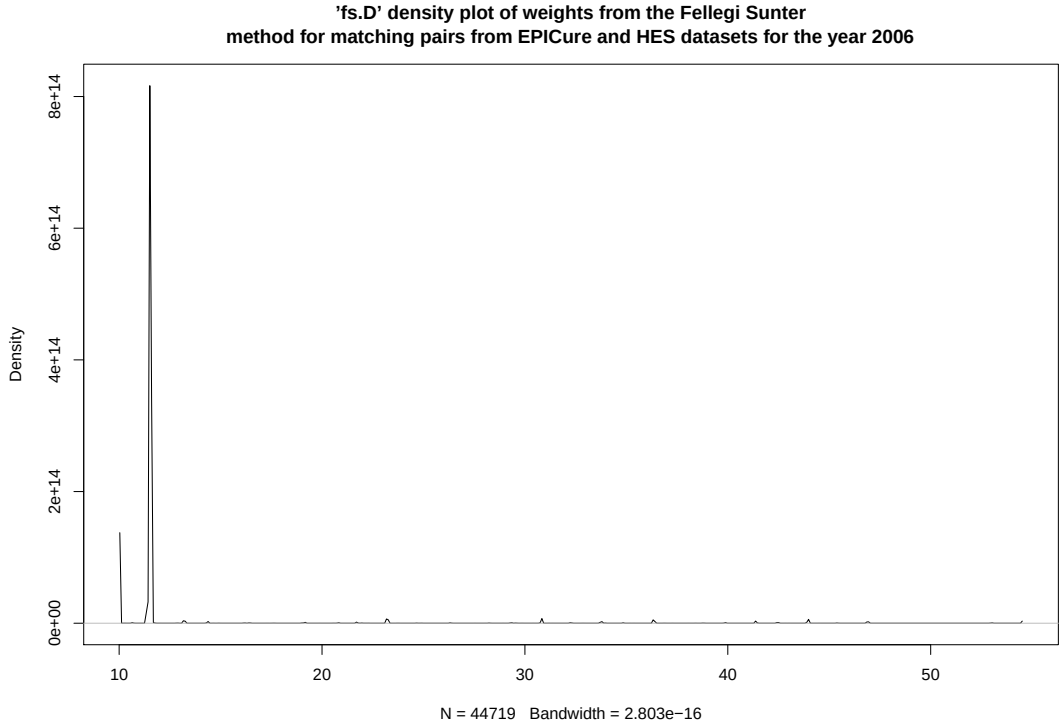
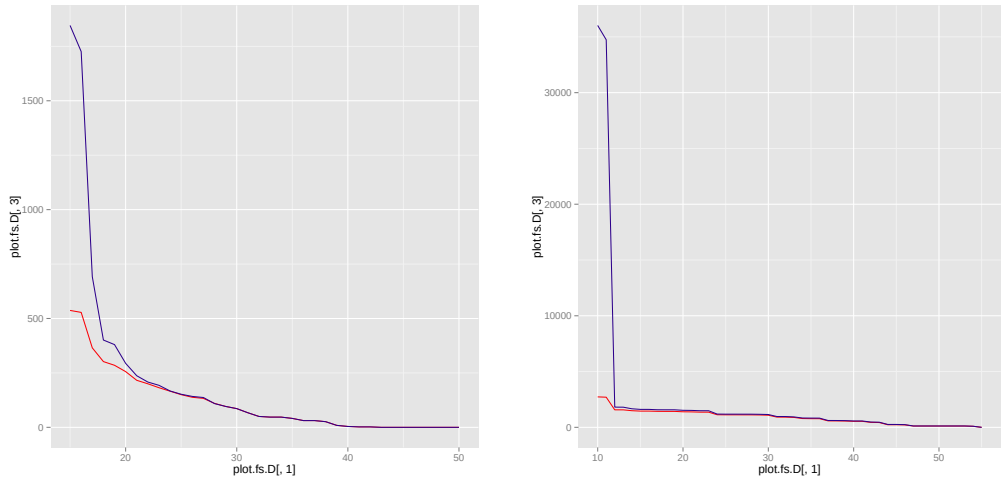


Figure 6.2: Numbers of individual matches according to weight from each of the Hospital Episode Statistics (blue line) and EPICure (red line) data sets in the stochastic linkage analysis using "guestimate" probabilities. "Weight" is on the x-axis, number of matches on the y-axis; axes are not to the same scale.

(a) 1995 **(b)** 2006



a weight of 18 to 3,129 with a weight of 19 or greater. This change corresponds to the largest peak for linkages seen on the 2006 density graph (figure 6.3b and also the sharp fall off seen in HES linkages figure 6.4b (blue line). By comparison, the peaks seen in the 1995 density graph (figure 6.3a) correspond to the biggest decrease immediately after the cut-off, reflected in the decrease in unique HES IDs linked (figure 6.4a; again, it should be noted that the axes from the two graphs use different scales). Full details of the number of pairs and unique IDs linked at each weight are presented in tables 6.8 (for 1995) and 6.9 (2006).

6.3.3 Contiero algorithm

The Contiero algorithm, unlike the classical stochastic method described above, assigns weights in the range $0 \leq w \leq 1$.[143, 199] For both years, a value of 0.35 was chosen as the cut-off weight above which to retain identified pairs for further review. This resulted in many more linked pairs being retained in 1995 than in 2006 – 45,349 compared to 6,323. Not surprisingly, this resulted in a much higher attenuation in 1995 than 2006; conversely, the spread of weights was much better in 2006 as can be seen in the density

Table 6.8: Table of the number of pairs matched in 1995 from each data set for differing cutoffs in the value of the weight calculated by the Fellegi-Sunter method of data linkage using estimates based on Dattani et al.[132]

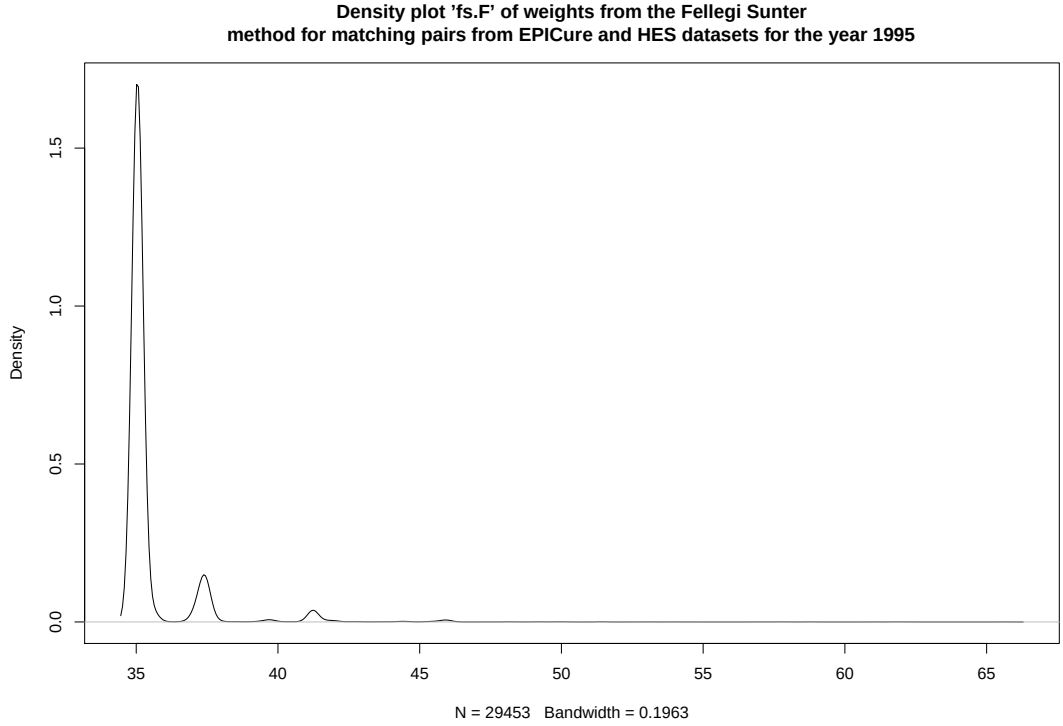
Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
35	29453	453	67.81	16385	3.37
36	3674	409	61.23	3540	0.73
37	3649	400	59.88	3518	0.72
38	1064	299	44.76	1049	0.22
39	1044	290	43.41	1029	0.21
40	920	277	41.47	915	0.19
41	905	270	40.42	900	0.18
42	313	198	29.64	311	0.06
43	254	169	25.30	252	0.05
44	240	160	23.95	240	0.05
45	205	147	22.01	205	0.04
46	87	85	12.72	87	0.02
47	85	83	12.43	85	0.02
48	80	78	11.68	80	0.02
49	69	69	10.33	69	0.01
50	51	51	7.63	51	0.01
51	49	49	7.34	49	0.01
52	40	40	5.99	40	0.01
53	38	38	5.69	38	0.01
54	33	33	4.94	33	0.01
55	29	29	4.34	29	0.01
56	29	29	4.34	29	0.01
57	27	27	4.04	27	0.01
58	18	18	2.69	18	0
59	13	13	1.95	13	0
61	13	13	1.95	13	0
62	5	5	0.75	5	0
63	3	3	0.45	3	0
65	3	3	0.45	3	0
66	0	0	0	0	0

Table 6.9: Table of the number of pairs matched in 2006 from each data set for differing cutoffs in the value of the weight calculated by the Fellegi-Sunter method of data linkage using estimates based on Dattani et al.[132]

Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
15	53413	2709	98.51	40104	6.35
16	38608	2672	97.16	32232	5.10
17	38020	2662	96.80	32114	5.09
18	37857	2658	96.65	32051	5.08
19	4497	1897	68.98	3129	0.50
20	4397	1885	68.55	3078	0.49
21	3614	1741	63.31	2489	0.39
22	3605	1741	63.31	2489	0.39
23	3582	1736	63.13	2469	0.39
24	2362	1512	54.98	1814	0.29
25	2358	1511	54.95	1812	0.29
26	2279	1488	54.11	1776	0.28
27	2200	1449	52.69	1717	0.27
28	2172	1432	52.07	1701	0.27
29	2159	1431	52.04	1700	0.27
30	1440	1089	39.60	1144	0.18
31	1439	1089	39.60	1144	0.18
32	1427	1082	39.35	1137	0.18
33	1404	1073	39.02	1129	0.18
34	1169	923	33.56	980	0.16
35	1162	920	33.45	974	0.15
36	1118	899	32.69	933	0.15
37	1116	899	32.69	933	0.15
38	930	748	27.20	781	0.12
39	821	661	24.04	692	0.11
40	817	658	23.93	690	0.11
41	772	624	22.69	653	0.10
42	540	457	16.62	488	0.08
43	431	368	13.38	398	0.06
44	431	368	13.38	398	0.06
45	385	338	12.29	362	0.06
46	300	253	9.20	277	0.04
47	300	253	9.20	277	0.04
48	278	231	8.40	255	0.04
49	242	206	7.49	224	0.04
50	242	206	7.49	224	0.04
51	141	114	4.15	127	0.02
52	83	70	2.55	75	0.01
53	38	37	1.35	38	0.01
56	38	37	1.35	38	0.01
57	26	26	0.95	26	0
60	26	26	0.95	26	0
61	2	2	0.07	2	0
63	2	2	0.07	2	0
64	0	0	0	0	0

Figure 6.3: Density distribution of weights from the stochastic linkage analyses using probabilities based on Dattani et al.[132] Axes are not to the same scale.

(a) 1995 baseline guestimates



(b) 2006 baseline guestimates

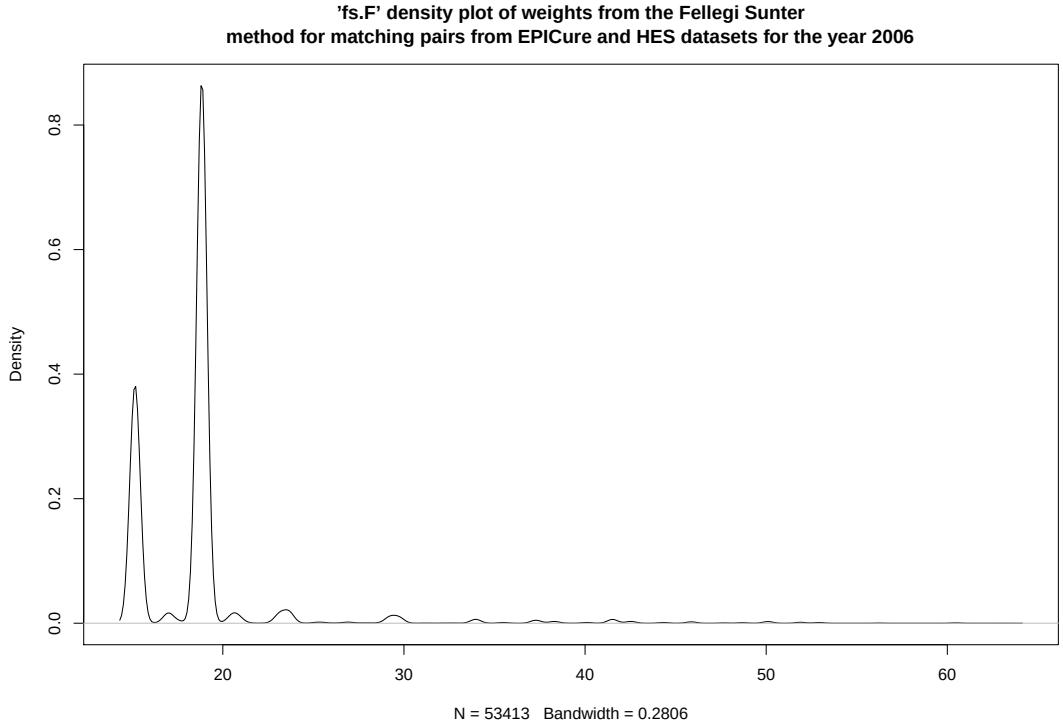
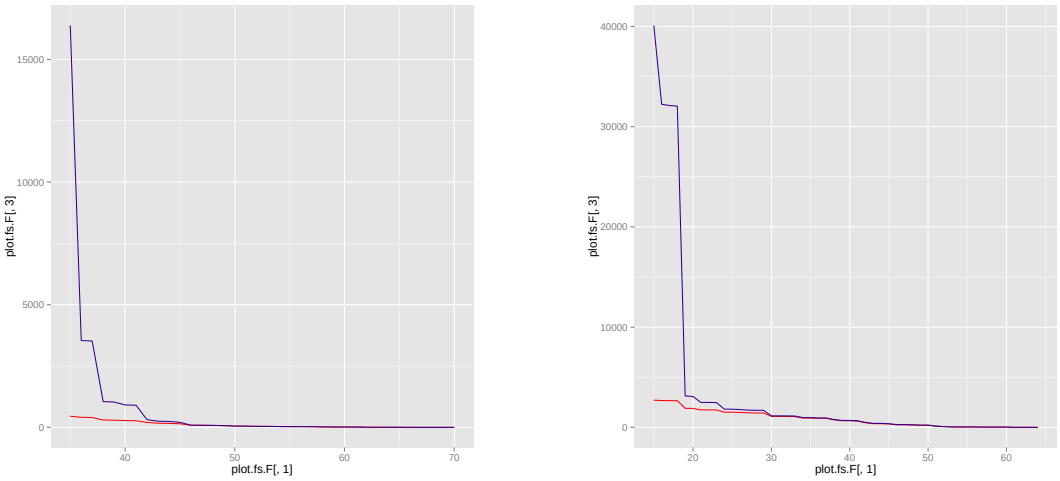


Figure 6.4: Numbers of individual matches according to weight from each of the Hospital Episode Statistics (blue line) and EPICure (red line) data sets in the stochastic linkage analysis using probabilities based on Dattani et al.[132] "Weight" is on the x-axis, number of matches on the y-axis; axes are not to the same scale.

(a) 1995 **(b)** 2006



graphs (figure 6.5). This was reflected by the maximum weight obtained in each of the analyses: 0.9494 in 2006 but only 0.8678 in 1995.

Interestingly, convergence in the numbers of matched IDs from each of the data sets (i.e. where the number of potential links for each true match approaches one) occurred around a weight of 0.45 in both epochs. This can be clearly seen in figures 6.6a and 6.6b which show the number of unique IDs identified from each data set in each epoch. However, it was only above a value of 0.75 that exclusively unique matches were identified in 1995 (a total of 20 – as shown in table 6.10), and in 2006 there was no obvious point at which unique matches could be identified: even at a weight of 0.90 there was some ambiguity, with 122 pairs identified but representing only 116 unique EPICure 2 IDs; there were no matches above a weight of 0.95 (table 6.11).

6.3.4 Estimation-Maximisation likelihood algorithm

The EM algorithm, unlike the previous analyses, did not require any prior inputs and calculated absolute weights based solely on the data sets available for matching. The maximum weight in 1995 was 65.7, with a maximum of 71.57 in 2006. Within the

Table 6.10: Table of the number of pairs matched in 1995 from each data set for differing cutoffs in the value of the weight calculated by the EpiLink (Contiero) method of data linkage.

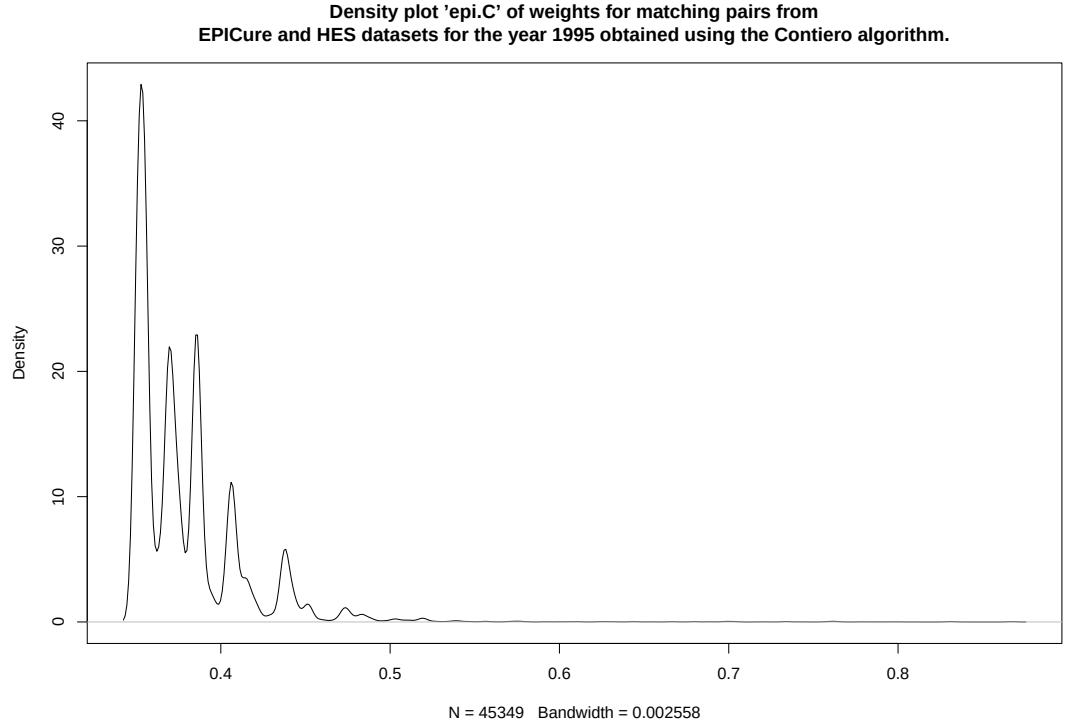
Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
0.35	45349	662	99.10	38163	7.84
0.40	9329	612	91.62	8533	1.75
0.45	1670	421	63.02	1541	0.32
0.50	492	279	41.77	461	0.09
0.55	213	193	28.89	209	0.04
0.60	157	147	22.01	153	0.03
0.65	117	111	16.62	114	0.02
0.70	78	74	11.08	78	0.02
0.75	51	51	7.63	51	0.01
0.80	20	20	2.99	20	0.00
0.85	8	8	1.20	8	0.00
0.90	0	0	0.00	0	0.00
0.95	0	0	0.00	0	0.00
1.00	0	0	0.00	0	0.00

Table 6.11: Table of the number of pairs matched in 2006 from each data set for differing cutoffs in the value of the weight calculated by the EpiLink (Contiero) method of data linkage.

Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
0.35	6323	2162	78.62	5608	0.89
0.40	2746	1542	56.07	2405	0.38
0.45	1573	1225	44.55	1328	0.21
0.50	1413	1135	41.27	1198	0.19
0.55	1180	965	35.09	1018	0.16
0.60	1125	922	33.53	975	0.15
0.65	928	774	28.15	805	0.13
0.70	623	514	18.69	545	0.09
0.75	538	447	16.25	478	0.08
0.80	264	237	8.62	264	0.04
0.85	137	129	4.69	137	0.02
0.90	122	116	4.22	122	0.02
0.95	0	0	0.00	0	0.00

Figure 6.5: Density distribution of weights from the linkage analyse using Contiero's EpiLink algorithm.[143, 199] Axes are not to the same scale.

(a) 1995



(b) 2006

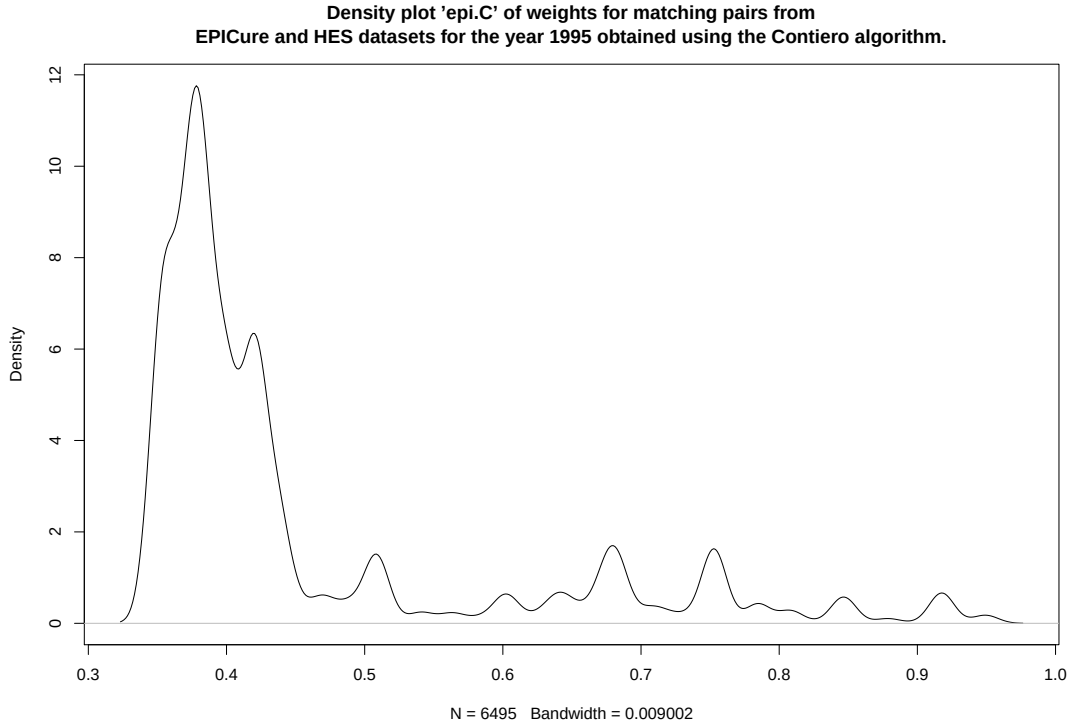
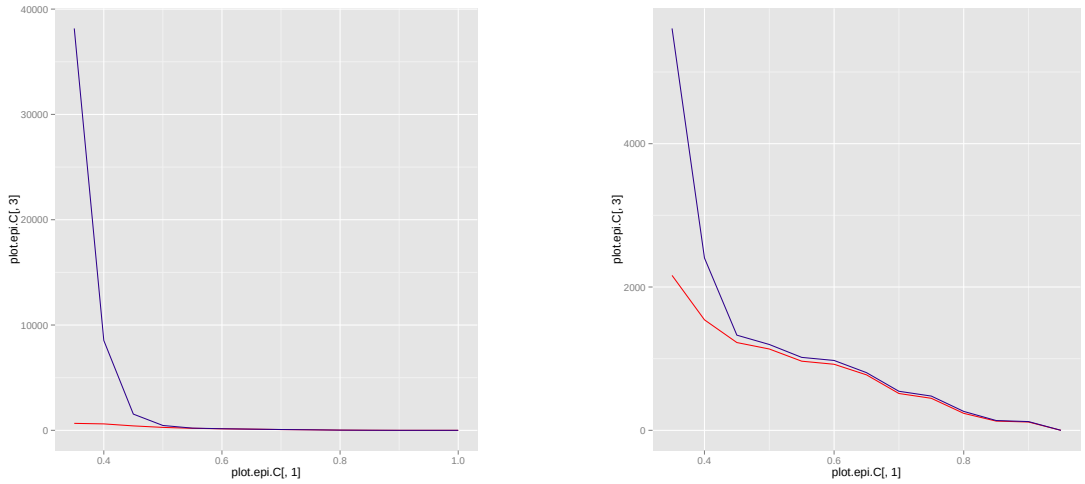


Figure 6.6: Numbers of individual matches according to weight from each of the Hospital Episode Statistics (blue line) and EPICure (red line) data sets in the linkage analysis performed using Contiero's EpiLink algorithm.[143, 199] "Weight" is on the x-axis, number of matches on the y-axis; axes are not to the same scale.

(a) 1995 **(b)** 2006



range of saved weights – a cut-off weight of 10 was used for both analyses – there was an essentially bimodal distribution for 2006 with peaks around approximate values of 10-15 and 20, and the remainder of the curve being almost flat (figure 6.7b). In contrast, the distribution in 1995 (figure 6.7a) was much more unimodal with a pronounced right-skew.

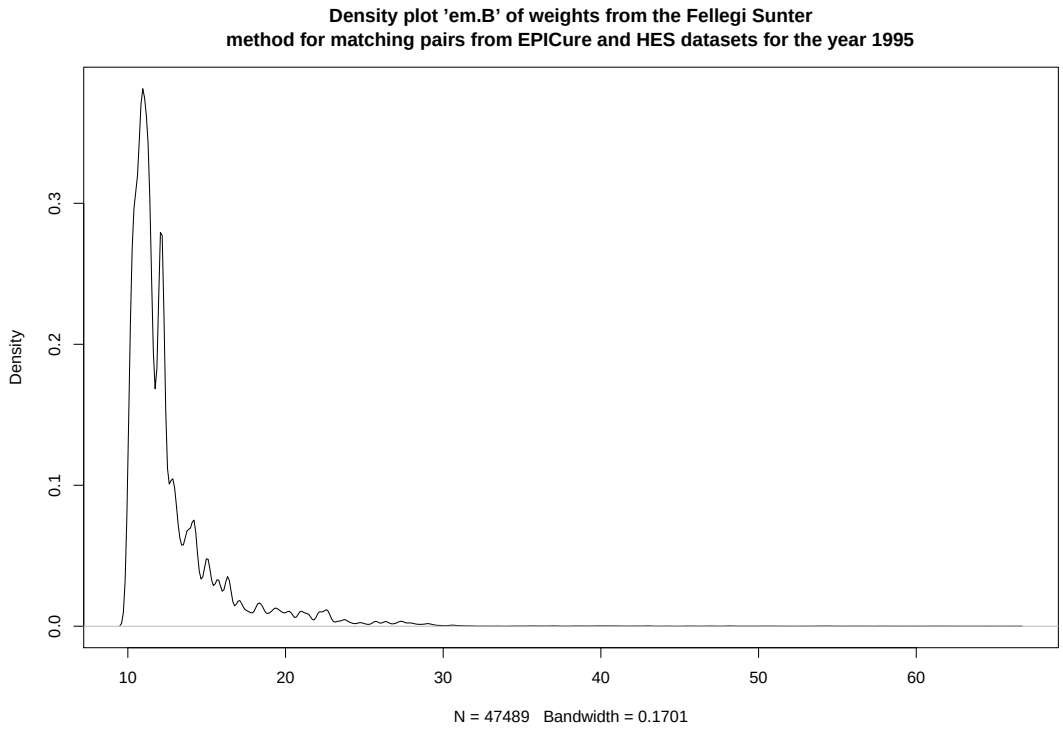
These findings are reflected in the number of pairs linked for each year: the rate of attenuation was much more steady in 1995 than for 2006, where there was a large drop in the number of linked pairs around a weight of 20 (see figures 6.8a and 6.8b to compare the number of linked IDs in 1995 and 2006). Even so, as displayed in table 6.12, in 1995 it was only above a weight of 43 that pairs were uniquely matched, and in 2006, only two unique pairs were identified – above a weight of 70, as shown in table 6.13.

Table 6.12: Table of the number of pairs matched from each data in 1995 set for differing cutoffs in the value of the weight calculated by the Estimation-Maximisation method of data linkage.

Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
25	644	232	34.73	502	0.10
26	541	212	31.74	421	0.09
27	429	189	28.29	348	0.07
28	294	175	26.20	241	0.05
29	231	155	23.20	201	0.04
30	177	139	20.81	161	0.03
31	153	132	19.76	147	0.03
32	140	123	18.41	135	0.03
33	136	120	17.96	132	0.03
34	134	118	17.66	130	0.03
35	129	114	17.07	125	0.03
36	121	108	16.17	117	0.02
37	114	101	15.12	110	0.02
38	107	96	14.37	103	0.02
39	99	90	13.47	95	0.02
40	90	82	12.28	86	0.02
41	80	77	11.53	79	0.02
42	74	73	10.93	73	0.01
43	66	66	9.88	66	0.01
44	61	61	9.13	61	0.01
45	58	58	8.68	58	0.01
46	53	53	7.93	53	0.01
47	48	48	7.19	48	0.01
48	43	43	6.44	43	0.01
49	37	37	5.54	37	0.01
50	34	34	5.09	34	0.01

Figure 6.7: Density distribution of weights from the linkage analyses using the Estimation-Maximisation algorithm

(a) 1995 baseline guestimates (note: this graph is mistakenly labelled "Fellegi Sunter"; it is actually from the EM analysis).



(b) 2006 baseline guestimates

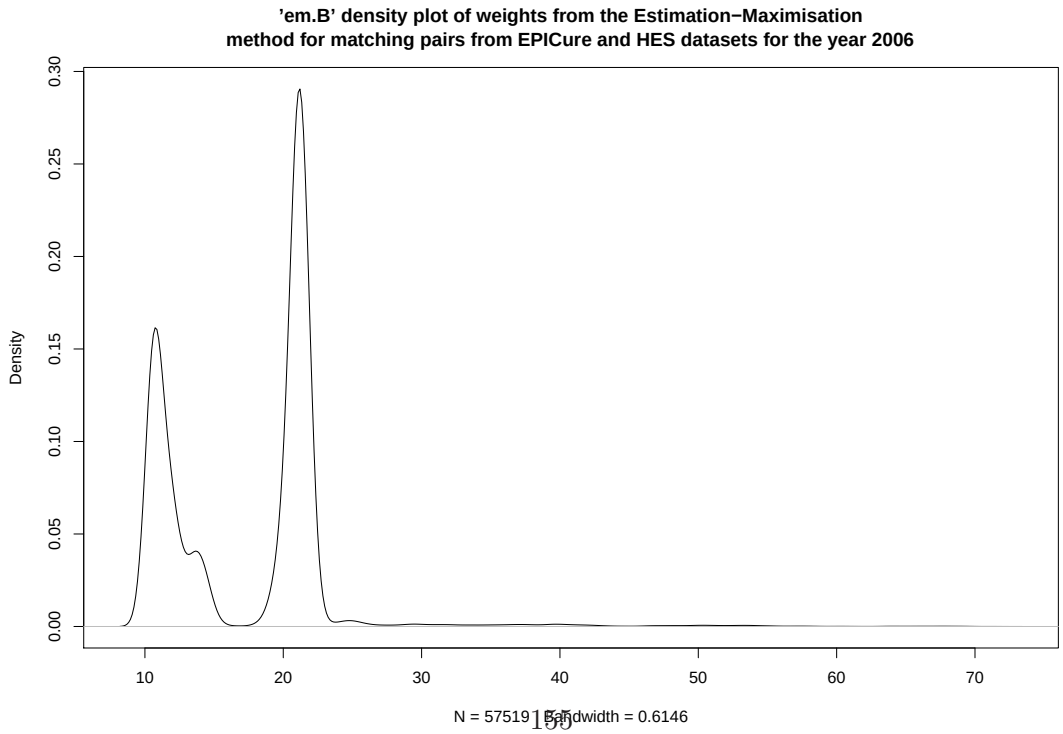
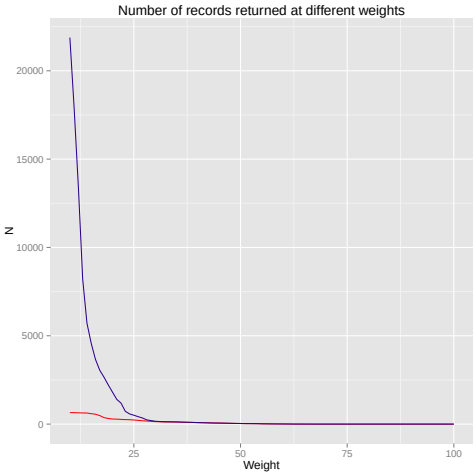


Figure 6.8: Numbers of individual matches according to weight from each of the Hospital Episode Statistics (blue line) and EPICure (red line) data sets in the linkage analysis based on the Estimation-Maximisation algorithm. "Weight" is on the x-axis, number of matches on the y-axis; axes are not all to the same scale.

(a) 1995



(b) 2006

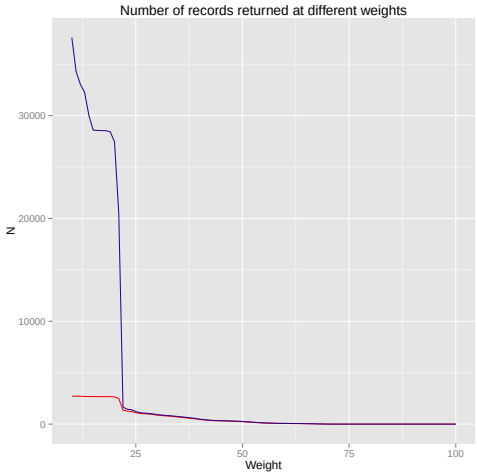


Table 6.13: Table of the number of pairs matched from each data set in 2006 for differing cutoffs in the value of the weight calculated by the Estimation-Maximisation method of data linkage.

Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
10	57519	2712	98.62	37589	5.95
11	45110	2712	98.62	34287	5.43
12	40419	2706	98.40	33051	5.23
13	37885	2686	97.67	32260	5.11
14	35164	2683	97.56	30017	4.75
15	33606	2678	97.38	28588	4.53
16	33570	2677	97.35	28558	4.52
17	33554	2677	97.35	28542	4.52
18	33540	2677	97.35	28529	4.52
19	33391	2677	97.35	28428	4.50
20	31450	2644	96.15	27453	4.35
21	22960	2483	90.29	20527	3.25
22	1981	1362	49.53	1645	0.26
23	1766	1265	46.00	1450	0.23
24	1687	1225	44.55	1395	0.22
25	1473	1113	40.47	1219	0.19
26	1340	1036	37.67	1106	0.18
27	1292	1009	36.69	1071	0.17

Continued on next page...

Table 6.13: Number of pairs at different cut-offs obtained from the Estimation-Maximisation algorithm in 2006 (Continued).

Cut off weight	N pairs	N EPICure	% EPICure	N HES	% HES
28	1247	980	35.64	1036	0.16
29	1201	939	34.15	993	0.16
30	1113	871	31.67	927	0.15
31	1056	835	30.36	883	0.14
32	999	799	29.05	845	0.13
33	953	771	28.04	816	0.13
34	902	729	26.51	776	0.12
35	856	686	24.95	734	0.12
36	806	645	23.45	695	0.11
37	747	599	21.78	647	0.10
38	685	550	20.00	597	0.09
39	638	513	18.65	553	0.09
40	555	449	16.33	483	0.08
41	499	408	14.84	438	0.07
42	455	374	13.60	400	0.06
43	414	339	12.33	359	0.06
44	400	325	11.82	345	0.05
45	388	315	11.45	336	0.05
46	382	310	11.27	330	0.05
47	356	289	10.51	307	0.05
48	335	277	10.07	296	0.05
49	304	259	9.42	279	0.04
50	277	239	8.69	259	0.04
51	232	207	7.53	225	0.04
52	206	184	6.69	202	0.03
53	173	155	5.64	172	0.03
54	137	131	4.76	137	0.02
55	117	113	4.11	117	0.02
56	100	96	3.49	100	0.02
57	89	85	3.09	89	0.01
58	77	73	2.65	77	0.01
59	70	66	2.40	70	0.01
60	65	61	2.22	65	0.01
61	61	57	2.07	61	0.01
62	58	54	1.96	58	0.01
63	56	52	1.89	56	0.01
64	48	45	1.64	48	0.01
65	41	38	1.38	41	0.01
66	36	33	1.20	36	0.01
67	27	25	0.91	27	0.00
68	15	14	0.51	15	0.00
69	9	8	0.29	9	0.00
70	2	2	0.07	2	0.00

6.4 Manual review of linked pairs

With none of the linkage methods achieving complete linkage of the EPICure data within a "reasonable" range of weights – i.e. the range above a cut-off weight chosen to allow realistic manual review of the linked pairs – the objective of obtaining supplementary data became unobtainable. Instead, it was decided to concatenate results from the different linkage analyses into one file for review: the 1,820 pairs identified across the four analyses in 1995 resulted in a total of 1,070 unique pairs, and 8,913 pairs in 2006 in 4,378 unique pairs.

The data from 1995 were then reviewed manually a total of four times. After the first review, which compared the HES with EPICure data, there were 935 pairs remaining, and after the second, 433 – which consisted of 431 unique subjects from EPICure and 427 unique HES subjects. The third review reversed the comparison, using the HES ID as the base and providing a choice of which EPICure IDs matched; this reduced the data to 425 unique HES and 423 unique EPICure IDs, following which the remaining unconfirmed matches were removed. This was because among the remaining unconfirmed links, the data conflicted in multiple variables. For example, one EPICure ID was matched with two possible rows in the HES data; the first potential match conflicted on birth weight, gestational age and sex, whereas the second conflicted on date of birth, sex and outcome (HES: unknown, but discharged on the date of birth; EPICure: died at 14 days of age). Thus, a total of 422 matches were confirmed between the EPICure and HES data in 1995, 63.17% of the potential 668 maximum potential matches.

For 2006, a similar process of review was undertaken. After the first review, 3,935 rows remained, consisting of 2,689 unique EPICure 2 and 3,865 unique HES subjects. This was then reduced to 1,708 rows (1,691 and 1,688 EPICure 2 and HES IDs, respectively) during the next round of review. Each review used the EPICure ID as a baseline for comparison, as there were fewer unique EPICure than HES IDs throughout. The size of the data set was reduced in a final review to 1670 rows, with 1,666 unique EPICure 2 and 1,670 unique HES IDs. There were insufficient data available to discriminate among the four remaining EPICure 2 IDs, which were each paired with two HES IDs: key data such as date of birth, sex, gestational age and birth weight all matched. Hence, it was not clear if the excess HES records represented duplicate entries

for a single subject, whether there were incorrect data (e.g. perhaps a set of records may represent twin births, with birth order miss-coded in one entry), or whether it was merely a result of an insufficient number of variables for performing the matching.

Discarding the four unconfirmed links from the 2006 data meant that overall there were a total of 1,662 confirmed of a maximum 2,750 possible matches – at 60.40%, slightly lower than the match rate in 1995.

6.5 Assessment of error

Using the data sets of linked pairs that were identified as true matches, it was possible to assess how well each analysis had identified matches. Sensitivity and specificity as well as positive and negative predictive values were therefore calculated, and are presented in tables 6.14 and 6.15 for the 1995 and 2006 analyses, respectively.

Across both epochs, the stochastic, classical Fellegi-Sunter analysis using the baseline "guestimates" provided the most accurate results, identifying 402 pairings correctly in 1995, and 1740 in 2006. It also had the highest sensitivity and specificity for each time period – although it only identified 63.27% of subjects in 2006, and 60.18% in 1995.

Table 6.14: Error measures of results obtained using different methods for linkage between the HES and EPICure data sets in 1995. EM: Estimation-Maximisation; FS: Fellegi-Sunter.

Linkage algorithm	cutoff	True matches	PPV	NPV	Sensitivity	Specificity
EM	10.00	238	0.005012	0.999999	0.356287	0.999855
EpiLink (Contiero)	0.35	387	0.008534	0.999999	0.579341	0.999862
FS (baseline model)	15.00	402	0.192069	0.999999	0.601796	0.999995
FS (Dattani estimates)	35.00	244	0.008284	0.999999	0.365269	0.999910

6.6 Deterministic linkage by the NHS HSCIC

The NHS HSCIC were unable to perform any linkage with the 1995 data due to the lack of identifiers (particularly, NHS number, but also postcode) in the EPICure data. They had great difficulty with the linkage in 2006, reporting that 1,567 subjects could not be matched at all – 1,336 of those because EPICure 2 did not have the NHS

Table 6.15: Error measures of results obtained using different methods for linkage between the HES and EPICure data sets in 2006.

Linkage algorithm	cutoff	True matches	PPV	NPV	Sensitivity	Specificity
Estimation-Maximisation	10	1408	0.0245	1.0000	0.5120	1.0000
EpiLink (Contiero)	0.35	1501	0.2374	1.0000	0.5458	1.0000
Fellegi-Sunter (baseline model)	10	1740	0.0389	1.0000	0.6327	1.0000
Fellegi-Sunter (Dattani estimates)	15	1665	0.0312	1.0000	0.6055	1.0000

numbers recorded. Ultimately, the NHS HSCIC returned one file containing those unmatched IDs and another file containing only 47 rows with two variables each, labelled "EXTRACT_HESID" and "OLD_EXTRACT_HESID".

It was unclear what exactly these IDs represented, nor what had happened with the other, matched IDs. I did not pursue the other linked pairs with the HSCIC, however, as there was a limited time to conduct the analyses and the original aim of using the data had already been nullified. This latter reason, in turn, was primarily due to the lack of data from 1995. Further reasons were the poor quality of the HES data and the reported lack of complete matching between the 2006 data sets. Combined, these factors meant that it was not possible to explore sociodemographic differences between 1995 and 2006 that may potentially have contributed to the change in admission rates seen for extremely premature babies in the EPICure cohorts.

6.7 Saved HES data

During the 10 months of the EPICure study in 1995, from 1st March to 31st December, there were 867 births recorded in HES with a gestational age of 25 weeks or lower. These were merged with the 422 "true" matches identified in the probabilistic linkage; there were 300 matches, leaving 567 subjects for whom no further investigation was possible.

By comparison, in the calendar year of 2006 there were 2,569 records identified of births at less than 27 completed weeks gestational age; 34 of these were duplicate records. Consequently, 2,535 births in total were recorded in HES as being of less than 27 weeks GA. These were compared with the results of the probabilistic matching: 1,670 records had been saved, of which 8 rows from the HES data had insufficient information

to match correctly with one of the corresponding four EPICure records; combining the data with these records excluded resulted in 932 matching rows, leaving 1603 for whom further review was not possible. The four undecided pairs (each with two choices from the HES data) mentioned above were then merged back with the EPICure data to gain further information; even after this step was successfully accomplished, there was insufficient information available to confirm which were the correct matches.

6.8 Answering the original question

Although it had already been decided in advance of record linkage that the original question (i.e. whether the apparent 44% increase in admissions to neonatal care between 1995 and 2006 was true) could not be answered due to the lack of additional data that were available (see section 6.2), this did not prevent a cautious investigation of crude changes between 1995 and 2006 that might be seen in the HES data.

Crucially, it was not possible to precisely distinguish the same populations in HES as were available in the EPICure data. In 1995, recorded data from EPICure were only available for babies of less than 26 completed weeks of gestation who were *admitted into neonatal intensive care*. In 2006, data on *all* births were collected in EPICure 2. By comparison, in HES in both years, there were data from live births of less than 24 weeks and for *all* births at 24 weeks gestation and above. Most importantly – for both study epochs – it was not possible within the HES data to distinguish live births who died in the delivery room from those who were admitted into neonatal intensive care but died on the same calendar date.

Table 6.16 shows how the data potentially changed between the two study epochs. There were 867 births recorded in HES in 1995 of < 26 completed weeks gestational age; 213 of these were still births. Examining this in relation to corresponding data for 2006 (i.e. also of less than 26 completed weeks gestational age and of a corresponding time period) shows there was a 37.0% increase in recorded births. For recorded live births, there was an 42.8% increase. Data are also shown in table 6.16 for three other populations contained within HES. The first is the "true" population: this contains data for those subjects identified by the linkage exercise after the clerical review. Second is the "confirmed" population, representing those that were reported as below 26 (and,

for 2006, also below 27) weeks who were confirmed by the linkage exercise; in contrast, the final group contains those from the reported group who were *not* identified.

6.9 Data destruction

The HES data were not investigated further, and were instead destroyed in their entirety. This was because they were of insufficient value – i.e. they could not be used for further investigations, primarily due to the problems with missing data and concerns over the accuracy of the remainder – to warrant further time expenditure on analysing and verifying them. Destruction was carried out using the `srm` facility provided in the `secure-delete` package in Debian,[200] which both removes data and 'wipes' the disc according to the algorithm developed by Gutmann.[189] This was confirmed to the NIGB by email in May 2014.

6.10 Chapter summary

The results from the Admissions Validation Study, which aimed to pair EPICure data with records from Hospital Episode Statistics in order to investigate an observed increase in admissions to neonatal intensive care in England, were presented in this chapter. The data sets received from the NHS HSCIC are described, including detail on both requested variables that were completely missing as well as levels of missingness within variables for which data were provided. There were high levels of missingness in the variables considered most important – socioeconomic status, which was completely absent from the 1995 data, and ethnicity, which was poorly completed.

Results from the four probabilistic analyses conducted for each epoch are then presented in sections 6.3.1 to 6.3.4. Overall, identification of true matches by the matching algorithms was poor. Combining all of the results from a given time-period, the number of matches achieved with HES data was approximately sixty percent of the numbers of births recorded by EPICure in both 1995 and 2006; of these 60% of records identified overall, the individual linkage analyses only identified between 35.63% – 63.27% each. Positive predictive value was also extremely low: under 5% in three-quarters of the analyses, and below 25% in all of them. Both negative predictive value and specificity were extremely high (tables 6.14 and 6.15).

Table 6.16: Changes in the number of births in Hospital Episode Statistics (HES) data between 1995 and 2006: reported, "true", "confirmed" and "misclassified" data.

HES data set [a]	1995 [b]				2006 (< 26 weeks) [c]				Percentage change [d]	2006 (< 27 weeks) [e]			
	Live births	Still births	Not known	Total	Live births	Still births	Not known	Total		Live births	Still births	Not known	Total
Reported	621	213	33	867	887	121	180	1188	37%	1856	201	278	2535
"True"	396	16	10	422	699	127	187	1013	140%	1158	213	291	1662
"Confirmed"	282	13	5	300	412	81	75	568	89%	684	134	114	932
"Misclassified"	339	200	28	567	475	40	105	620	9%	1172	67	364	1603

[a] For each year, data sets were created based upon : a) gestational age as reported in the original HES data; b) only the "true" data identified by the data linkage exercise (i.e. contained in both HES and EPICure); c) HES data "confirmed" by the "true" data; and, d) "misclassified" data, which are those reported by HES but that were not identified as "true" during data linkage.

[b] In 1995, data were available from March 1st – December 31st for babies of < 26 completed weeks gestational age.

[c] Comparison data sets from 2006 were created to include babies born between 1st March and 31st December at less than 26 weeks gestational age.

[d] The total percentage increase in all births is presented.

[e] The complete data sets from 2006 include births of < 27 completed weeks gestational age from the entire year.

The chapter continued in section 6.6 with a brief outline of what happened in the deterministic linkage. This was carried out by the NHS HSCIC – who were unable to perform any linkage for 1995, and achieved only partial success in 2006. A crude review of how the original question ("was the 44% increase in admissions to neonatal intensive care that was seen between the EPICure studies of 1995 and 2006 true?") might have been answered was then presented, before confirming that all identifiable data were securely disposed of.

Chapter 7

Discussion of data linkage exercise

This chapter discusses the results of the Admissions Validation Study. Structure broadly follows the STROBE guidelines,[148] but also takes into consideration points raised by a proposed guideline for data linkage studies.[146] The key results from probabilistic linkage studies between data from Hospital Episode Statistics (HES) and the two EPICure cohorts of 1995 and 2006 are described. I then discuss the limitations of the data linkage, considering both the data and the methods separately before more general issues of bias, confounding and chance. Next, in light of the limitations, the results are interpreted, then their generalisability considered. The chapter concludes with a brief summary.

7.1 Key results

HES data were hypothesised to be a suitable data source with which to investigate the apparent 44% increase neonatal admissions between 22 and 25 completed weeks gestational age that was seen between the two EPICure studies, from 1995 to 2006. This was found to not be true. In 1995, data on 575,509 HES subjects were available, of whom 486,705 were included in a data linkage exercise pairing records with 668 subjects from EPICure; in 2006, 631,401 records of an available 631,499 from HES were linked with 2,750 subjects from EPICure 2. Overall, approximately 60% of available EPICure records were successfully linked using a combination of probabilistic methods in each

study epoch.

Three specific methods were utilised, corresponding to the stochastic method described by Fellegi and Sunter, the EpiLink method described by Conterio, and the Estimation-Maximisation (EM) algorithm. Of these, the Fellegi and Sunter technique using best-guess estimates ("guestimates") of matching probabilities was the most successful method for data linkage in 1995, having a sensitivity of 0.601796, specificity of 0.999995 and positive predictive value (PPV) of 0.192069; negative predictive value (NPV) for all algorithms was 0.999999. There was no clear "best technique" in 2006. The PPV was only 0.0389 for the baseline Fellegi and Sunter guestimate, although the sensitivity was the highest of the assessed techniques at 0.6327. In contrast, the EpiLink method had a PPV of 0.2374 and a sensitivity of 0.5458. There was no discrimination between any of the techniques in 2006 in terms of NPV or specificity.

Deterministic linkage performed by the NHS Health and Social Care Information Centre was unsuccessful. Examination of the original HES data sources in isolation demonstrated that there was an increase between 1995 and 2006 of 37% in the overall number of reported births, and 42.8% in live births in a population similar to that of the EPICure studies: less than 26 weeks gestational age and born between March 1st and December 31st. This suggests that the 44% increase in admissions to neonatal intensive care that was seen in the EPICure data might be real. However, there were insufficient other data (ethnicity, socioeconomic status) to permit detailed investigation.

7.2 Limitations

There are many issues with this study that impact the conclusions. These relate to both the data and the methods that were employed in the record linkage. By combining these two issues, it is possible to then review the study overall, and examine issues of bias, chance and confounding.

7.2.1 Data considerations

Data quality may be affected by variations in population coverage, completeness of individual variables collected, and by the accuracy of the data recording. In this study, there were clear differences in the quality of the data sets that were available for each of these areas.

Hospital Episode Statistics are a routine data set that have been collected from secondary care sources with primarily non-clinical motives since 1989.[122] With respect to birth data, they are incomplete. Births in non-NHS locations (private hospitals or birthing centres, or at residential locations) may not be collected, and there is marked variation in reporting by different health care providers (be they regions or hospital trusts).[110, 132, 201] In contrast, the EPICure data were specific cohort studies run in collaboration with national confidential enquiries (CESDI and CEMACH).[54, 55] It is thought that data were collected on all relevant births that occurred during the study periods.[54, 55]

The completeness of individual variables also differs between the data sources. Data entry for HES may be conducted by midwives immediately after delivery via point-of-care systems or separately by clinical coders; reporting practices have changed over time.[201] High levels of *missingness* were seen in almost all of the variables used for record linkage (table 6.2); many variables contained a complete absence of data (appendix F). Similarly high levels of missingness have also been reported elsewhere for HES maternity data in 2006 [132] and more recent years [133, 202]. For the EPICure cohorts, data were *only* collected about specific births by those directly involved in care under the responsibility of a delegated EPICure contact (usually a doctor) at each perinatal centre in England.[54, 55] Data returns with missing data were individually investigated by the principal researchers.[54, 55] Consequently, few data are missing.

The EPICure data are also more likely to be accurate than the HES data - again, due to the manner in which the data sets were collected. This was demonstrated by inconsistencies between gestational age and birth weight categories in the HES data available for this study, as shown in tables 6.4 and 6.5, while the EPICure data are known to correlate well.[54, 55] Problems with variable quality have been identified previously in Hospital Episode Statistics, [203–205] as well as other routine data sets.[206] In linkage studies particularly, decreasing gestational age is associated with fewer matches and poorer quality data.[205, 207, 208]

Thus, overall, the combined differences between the EPICure and HES data sets severely limited the capacity for accurate data linkage, and prevented further meaningful investigations. However, the issues encountered are not unique to this study.

7.2.2 Methodological considerations

Irrespective of the data quality concerns – perhaps even because of them – there are interesting lessons to be learned from the record linkage. This used three different techniques, including, for one of them, the use of two different sets of estimates. Four different linkage analyses were therefore carried out for each year of available data.

Fellegi and Sunter analysis

The stochastic linkage method in R that was based on Fellegi and Sunter's classic description [141] allows M and U probabilities to be specified in advance, else defaults are provided based on calculated frequencies of the variables' response values.[143] Two sets of estimates were used: baseline "guestimates" and "Dattani estimates" from work by Dattani et al.[132]

For the guestimate analyses, U-probabilities were proportional to the frequency of responses per variable. For example, chance agreement with date of birth is likely to occur once in approximately 365 records, thus resulting in a probability of 0.00274. M-probabilities were based upon how likely data were to be entered correctly; for instance, a high level of confidence was expressed that "sex" would be right the majority of the time (M=0.999 – see table 5.2) whereas date of birth would be correct only 90% of the time (as it is a more complex variable to enter so more susceptible to typing errors). The values were thought to be reasonable starting points; at the time they were chosen, the data linkage was not expected to be so time-consuming, hence I thought it would be possible to vary these in subsequent sensitivity analyses. As it turned out, this wasn't possible. However, the "guestimate" analyses identified more pairs than any of the others in both 1995 and 2006, so clearly the chosen values were reasonable. Results for the chosen values may also be compared with those from the Dattani estimates (discussed below). Related factors that weren't considered during study design were the resultant weights: for those variables where the M and U probabilities were the same, the weight for both matched and non-matched pairs was zero. But no difference in weight means no distinction is made between matched and non-matched pairs – essentially meaning that the variable is not considered during matching! This was not a fatal design flaw as there were sufficient variables used in each of the matching exercises that the impact wasn't noticed. However, by better selecting probability estimates, it

may have been possible to achieve a better spread of weights – and consequently increase discrimination between linked pairs.

In the second set of analyses, using the Dattani estimates, values were calculated using data from a previous matching exercise performed between HES and registry/NN4B data.[132] Where data were not present (e.g. because a variable wasn't reported on previously), they were supplemented by the estimates used for the "guestimate" analyses. This was necessary for discharge date, mothers age at delivery and date of birth, and the number of previous pregnancies that she had had. A major criticism of this is that Dattani's work was based on matching of maternity HES data from 2006.[132] Here, the estimates were used for analyses covering both 1995 and 2006. Consequently, the estimates may not have been relevant for the earlier time period.

It is worth examining this from a different angle. First, it was fortunate that there were *any* estimates. Dattani's paper was only published in spring 2011,[132] shortly before this analysis was planned (see section 4.6.1). Using these estimates for the 2006 data therefore seemed reasonable, and in the absence of any other evidence, the application to other data from the same source (albeit collected 11 years earlier) not unreasonable. Secondly, estimates are just that: approximations to what the true value may be. We cannot know this with certainty prior to linkage, nor which would be the best – or best combination of – probability estimates. The ideal way to investigate this would be to repeat the exercise numerous times, on each occasion adjusting just a single estimate (i.e. for one variable at a time). This might enable optimal values for best linkage results to be obtained – but would be a highly time-consuming process to iterate through a virtually infinite set of possible estimates for each variable. This means it is useful to have a second set of estimates to permit comparison between linkage analysis results, even if it is not possible to distinguish effects of the individual estimates.

One consideration here is that only some of the probability estimates were changed in the Dattani analyses. Most importantly, the two variables that previously had equal weights (delivery method and number of babies) were included among the variables for whom information was available from the previous study.[132] This would have resulted in better discrimination for the summary weights calculated per pair, and thus a better spread of values. This certainly appears to be the case when comparing figures 6.1 and

6.3, more so for 2006 than 1995 – although it is difficult to assess due to the variation in scale on the x-axes.

Estimation-Maximisation algorithm

The EM algorithm, unlike the previous analyses, did not require any prior inputs and calculated weights based solely on the available data. This was not fully appreciated in advance of conducting the analyses. Consequently, the analysis function that was created to analyse the data according to the work flow, described in section 5.3.2, will have had a high likelihood of introducing errors into the linkage process. This is due to the way in which each completed EPICure data set was examined in relation to HES data for just a single day at a time in the corresponding year; results were then concatenated to ensure all potential links were assessed. However, this method will have resulted in different starting points *each day* for the EM algorithm, thus potentially causing errors to be introduced. This is because the same cut-off weight applied throughout the whole analysis, despite different weights being calculated for each daily block.

Of all the probabilistic methods that were used, the EM algorithm was the most resource-intensive in terms of processing time. With the available computer resources, analyses took approximately 10–14 days. This should be compared with hours (14–18) to days (2–3) for the Fellegi and Sunter, and EpiLink/Contiero analyses. Unfortunately, precise records of the time usage for the individual methods were not kept. This, in combination with the previously stated problem regarding calculation of weights and the fact that Estimation-Maximisation identified the fewest number of records among all the linkage analyses (table 6.14 and 6.15), lead to the EM algorithm being the least preferred of the linkage methods utilised.

EpiLink (Contiero) approach

There was also the potential for a similar error to that which affected the EM analyses to occur in the EpiLink analyses.[143] This was because the Contiero algorithm bases estimates of weights on the frequency of responses and estimated error rates – shown in equation 5.7.[143, 199] Such an issue was avoided in the analyses by specifying frequencies based on rates in the entire data set, and by setting an error rate in advance.

Results from the EpiLink analyses were reasonable. In the 1995 analysis, the second highest number of "true" matches were identified, with a sensitivity of approximately 58% (table 6.14). In 2006, it did not perform so well, although PPV was just under 24%, compared with under 4% for all the other methods. Interestingly, the threshold that is specified represents the similarity of the matched records.[199] This means that the results shown for numbers returned at differing thresholds in tables 6.10 (1995) and 6.11 (2006) reflect the accuracy with which data has been recorded in the data sources – although it is not possible to say in *which* data source any inaccuracies may have occurred.

7.2.3 Bias

In the current scenario of performing record linkage between two data sources, the population under consideration is the entire set of possible links. Thus, the population consists of two groups: true matches and the true non-matches (where 'match' corresponds to a real, actual match whether identified or not by the linkage process). Bias relates to the information available for each of these groups.

Selection bias

Knowing the population being studied means it is possible to decide the extent of any selection bias. Are there any differences in how the group of true matches was selected for inclusion, in comparison to how the true non-matches were selected? If the answer is yes, how might this have impacted the data linkage?

A number of issues that related to this have already been described in section 7.2.1. These primarily relate to the HES data, which may have biased the population coverage – for example, if data were less well reported in some regions or for some hospitals than in the EPICure studies. The result of this would have been that some matches were not identified when they could have been had the data been present in HES.

It is also possible that selection bias occurred due to the population of births selected for linkage: those at the extremes of viability, specifically including (and for 1995, *only* including) those who were admitted into neonatal intensive care units; these babies are almost exclusively born in hospital. EPICure data do not report any births out of hospital in 1995 [54] and, in 2006, 62 of 3133 (2.0%) of births were not in a hospital; 38 of these were live born.[55] Only 2.4% of all births were at home in 2011; this number

had previously increased from a low in the 1980s through to 2008-09, from which it had started falling again.[209] These are not important differences: performing a comparison test of the two proportions [210] demonstrates a p-value of 0.255; therefore, any impact could be considered as minimal. However, there is marked variation in the home birth rate by region and maternal age,[209] and the home birth rate has not always included births that take place in establishments outside of an NHS hospital, although many locations now contribute data to HES.[132] Thus, there is more uncertainty around the impact of this potential bias.

Information bias

Bias is more likely to arise as a consequence of the data consistency issues within HES, and from errors in methodological design. The inconsistencies between gestational age and birth weight presented in this thesis are likely to be the tip of an iceberg. Similar work performing linkage with maternity data for England and Wales has shown low rates of discordance between sources;[111, 131–133, 211, 212] however, data quality issues are more likely to be an issue for those born in unusual circumstances such as those who are extremely premature. Any such errors are likely to apply equally across the gestational age ranges included in the two study epochs, thus meaning any misclassification was non-differential, and therefore likely to bias the linkage towards non-identification of true matches.

This contrasts with the differential misclassification that is likely to have occurred during the EM analysis. Because weights were calculated by an algorithm from the available measures, and HES records were grouped by date at the level of a single day, birth records for each single day would have been exposed to a different set of weights from every other day. It is unknown what the effect of this may be - whether greater or fewer records would be identified. However, it is probably more likely that this produced an underestimate (i.e. a nullification of effect), as there were several dates noted during the analysis when no convergence of the algorithm was achieved – indicating that it was not possible to calculate weights for the available data and hence there were no matches.

Another problem, again within the analysis stage, arises from combining the results prior to manual review. The consequent time reduction for manual review was probably great, but means it is likely contamination occurred between the different

linkage methods. This is because any remaining linkage pairs with HES or EPICure IDs corresponding to those in an identified "true" match were removed, thus meaning identification of a match from one analysis could influence the choice of match arising from another.

7.2.4 Confounding

It is hard to know whether there was any confounding, although this can never be excluded. In this study, the exposure can be defined as those subjects with a recorded gestational age within the pre-stated range (<26 weeks for 1995, and <27 for 2006); the outcome is successful linkage. Viewed in this fashion, there is little scope for confounding: it might occur if birth weight were closely correlated with gestational age – but it has already been discussed that this is not the case. There were no other variables in the HES data that would be expected to show a strong correlation with extreme prematurity.

7.2.5 Chance

The remaining factor that may affect the analyses is random error. Given that the purpose of probabilistic linkage is to assign a weight from which a threshold may then be chosen to identify the most "true" matches, random error is unlikely to be of great importance. This is because manual intervention is required, if not for clerical review, at least for selection of a threshold. This consequently provides a counterbalance to random error: an acceptable level of error is determined by the number of records to review.

To facilitate a shorter review period, linked pairs from all the comparisons were merged into one file. This will have changed the probability of a match being identified, but it is not known whether this would have caused an effect or what that effect might have been.

7.3 Interpretation

Given the many concerns just detailed, there *are* some interpretations that can be made, and the study was *not* an unmitigated failure. Although it was not possible to investigate changes in socioeconomic factors or ethnicity over time using the HES data,

it is extremely interesting to note that there were similar increases to the 44% rise in admissions seen in EPICure. Moreover, the increases were approximately the same: around 40% of all births and of live births only in those reported in HES. Increases were even larger when considering only the population reported in HES *and* identified as "true" in the linkage analyses: 140% increase in all births, and 76.5% in live births only (figures in table 6.16). This provides some confidence that the EPICure findings are true – although leaves us without any obvious possibility for further investigation of potential reasons why.

Indeed, the conclusion that HES data are an extremely poor source for information about those born extremely prematurely is one of the most important results. There is no evidence that HES data improved between 1995 and 2006 for this group of babies, either. This is the first study to look at such a specific population: previous studies have focused on linkage for the entire gestational age range, thus errors at the extremes are dissipated.

What about methodological interpretation? Firstly, the methods took a lot longer than had originally been anticipated; this was facilitated with an extension to the permission granted by the National Information Governance Board for the HES data (see section 4.8.1). However, further investigation of linked data was not possible due to the need for data destruction, and procedures to obtain a further extension were disproportionately onerous to the information potentially gained. The research governance requirements were not facilitated, either, by transitions between agencies: data permissions were initially obtained from two agencies (the Northern and Yorkshire REC and the Ethics and Confidentiality Committee at the NIGB), but the data were then obtained from a third (the NHS Information Centre); during the project, all three transitioned their roles to new agencies (North East - York REC; Confidentiality Advisory Group of the Health Research Authority; and the NHS HSCIC) while the project was being undertaken.

Bureaucratic difficulties are not novel, but they are cumbersome and not infrequently described in the literature.[213, 214] One recent UK study investigated the effect of delays in three clinical trials: the involvement of multiple agencies in governance, alongside multiple trial locations and external time limitations (e.g. time-limited funding), created delays that had important impacts. These ranged from delays in recruitment and/or data collection through to delays in the production of results and

final reporting.[213] Other studies have previously found similar delays attributable to a cumbersome governance process, and recommended changes.[214] Restructuring the organisations involved in governance and health systems does little to help the problem.[215]

More specifically, there are lessons about the analyses. The coding could have been better. They may also have been quicker or easier using a different software: there are programs specifically designed for data linkage (like Contiero's *EpiLink* [199]) although then comparisons of different techniques may not have been possible. Estimates for the Fellegi and Sunter analyses could have been chosen in a different way – either with different rationales, leading to different estimates or, better, in the fashion described earlier, changing just one estimate at a time. It might have been better to run some pilot analyses, although in fact this was almost done by default due to the way data was processed on a *per diem* basis necessitating review at regular intervals.

7.4 Generalisability

How useful are the findings from this study? One the one hand, the findings in relation to the primary objective – to confirm whether there was an increase in births between 1995 and 2006 – are extremely important, as they suggest that extremely premature birth is becoming more frequent and build on the observations of the EPICure study.[55] On the other hand, specific data produced by this study are likely to be of little use: the overarching lesson being that routine data sources are not useful for investigations of extreme prematurity. Rather, from the data linkage, a great deal was learned about *process* – both setting up the study (acquiring the HES data and the research governance frameworks in England) and conducting it (the individual statistical techniques, and how to improve statistical coding). These lessons are highly applicable to other research settings.

7.5 Conclusions

This chapter has discussed the key findings from the Admissions Validation Study. It commenced with a brief summary of the key findings, before considering the limitations of the study in section 7.2. Problems with both the data and the individual techniques

were highlighted, and overall concerns relating to bias, confounding and chance were discussed. Interpretation of the findings was presented in section 7.3. This commented on specific findings from the study as well as lessons that could be learned from the overall conduct of the study. The increased applicability of these latter findings, compared to specific detail, was then noted.

In summary, the most robust finding was that HES data are an extremely poor source for information about those born extremely prematurely, with no improvements in data quality seen over time. Increases in extreme prematurity were seen that were in the same direction and of a commensurate value to those seen in the EPICure studies.

Part III

Obstetric antecedents of extreme prematurity

Chapter 8

Methods for the obstetric antecedents analysis

The relationship of antenatal and perinatal factors to the chances of the baby being born in a good condition and to perinatal death were investigated using data from the EPICure 2 cohort. The investigation sought to assess whether perinatal decision-making is solely responsible for improved short-term outcome, or whether there were additional, independent benefits conveyed by obstetric interventions.

In this chapter, the methods of this investigation are described. First, the data set that is used in the study is briefly outlined, and inclusion and exclusion criteria and their rationales are explained. Following this, specific attention is paid to detail about the choice of variables used in the study as exposures, outcomes and potential confounders. The populations included in the analyses are explained after this. Then, the statistical methods employed are outlined, followed by the final section which discusses the sensitivity analyses that were performed.

8.1 EPICure 2

EPICure 2 data were available on all babies known to have been born between 22 and 26 weeks completed gestational age to English mothers in England during 2006. Full details of data collection and the resultant data are provided in section 4.2.

8.2 Inclusion and exclusion criteria

The study population was restricted to mothers with singleton pregnancies who delivered between 22^{+0} and 26^{+6} weeks gestational age where the fetus was considered to be alive at admission to hospital, and at either the start of monitoring of the labour or the point at which it was decided to perform Caesarean section. The reason for this was that obstetric interventions that are performed with the aim of optimising the fetus in order to promote neonatal well-being, such as the administration of antenatal steroids, are only valid options for women who arrive at hospital and are assessed as having a live fetus. It is moot if the fetus is assessed *prior* to admission as being alive, as there is no possibility of intervening, likewise if the fetus is no longer alive at admission, there is no point in providing therapy that aims to improve outcome.

Women with multiple pregnancies were excluded, as were women who had terminations of pregnancy (ToP). This was because multiple pregnancies may have different outcomes compared to singleton pregnancies, thus complicating explanations of the outcome, and for ToPs, there is no expectation (or desire) that the fetus will survive.

8.3 Choice of variables

For this study, data collected from questions on the perinatal EPICure 2 "PN:E2" form were used to investigate the effects of antenatal maternal steroid and tocolytic administration, and of Caesarean delivery, on condition at birth and on death prior to admission into a neonatal intensive care unit.

8.3.1 Exposures

Three exposures were chosen: maternal antenatal administration of steroids, antenatal tocolytic therapy, and delivery of the fetus by Caesarean section. Antenatal steroid administration was categorised into three levels - none, partial course (if a patient received the last dose less than 24 hours prior to delivery), and full course (if the time interval to delivery was greater than one day). Tocolysis, although initially analysed by type of drug administered, was re-categorised in a binary fashion due to inadequate spread of data. Mode of delivery was categorised as either vaginal or Caesarean, with labour type coded separately.

These three variables were chosen for the following reasons:

Antenatal steroids: These are associated with an improved outcome in premature babies – although few studies have large numbers of extremely premature babies included. Therefore, the aim was to find out if antenatal steroids were effective in the population of singleton babies born between 22^{+0} and 26^{+6} weeks gestational age.

Tocolysis: There is debate over the utility of tocolysis due to concerns about prolonging fetal exposure to an adverse *in utero* environment.[216] Hence, use of tocolysis was chosen as an exposure to explore its effects on the two outcomes, perinatal death and condition at birth.

Caesarean delivery: Operative delivery in a situation where the fetus is non-viable is contraindicated due to the associated maternal effects. Due to a lack of evidence, many obstetricians are also reluctant to perform Caesarean section at extremes of prematurity when fetal prognosis is unknown. An analysis using mode of delivery as an exposure was performed in order to fill this gap in knowledge.

8.3.2 Outcomes

Two binary outcomes were investigated: "birth in a good condition", and "perinatal death". The former was defined by the presence of a heart rate above 100 beats per minute (bpm) at 5 minutes after birth. Condition of the newborn at birth using this definition was noted to be related to longer term outcome in the first EPICure cohort, hence being chosen as an outcome.[54, 68] The latter outcome included all deaths occurring during labour or in the delivery room. This definition of "perinatal death" is different, notably, from the usual definition of the "perinatal mortality rate"[217] as the outcomes for this study were specifically designed to capture the impact of obstetric (rather than neonatal) interventions on the newborn.

8.3.3 Potential confounders

Variables available for consideration as potential confounders related to the condition of the mother and pregnancy antenatally, fetal and pregnancy factors around the time

of delivery, obstetric management factors, questions related to antenatal counselling and decision-making, and health service factors.

Antenatally, these variables could be broken down into: demographic data such as maternal age in years, ethnicity (white, black, Indian/Pakistani/Bengali, other), body size and smoking status (current or non-smoker at first pregnancy booking appointment); maternal medical complications (diabetes either before or during pregnancy, hypertension or epilepsy); obstetric complications (prolonged premature rupture of membranes, abruption, antepartum haemorrhage after 20 weeks of gestation, pre-eclampsia or cervical incompetence warranting placement of a cervical suture); and fetal complications (intrauterine-growth restriction and/or oligohydramnios). Fetal sex (male or female; those originally categorised as "indeterminate" had been re-classified by the principle study investigators) was considered in all analyses.

Gestational age of the infant, determined by the earliest available ultrasound scan, was included as a categorical variable (per week) in the statistical analyses for ease of presentation. Binary variables were created for labour type (spontaneous or none/induced) and presentation of the baby at delivery (cephalic or non-cephalic). The presence or suspicion of chorioamnionitis at any time was included as an antenatal risk factor, with maternal antibiotic administration prior to labour classified into treatment, prophylaxis or not prescribed.

Provision of antenatal counselling was divided into the actual provision (Was there counselling by a senior obstetrician? Was there paediatric counselling?) and the content of the discussion. This included whether a decision to *not* perform Caesarean section in cases of fetal distress was made; whether or not withholding care was discussed; and whether the parents expressed any choice about resuscitation and provision of neonatal intensive care (provide full care for any live birth; withhold intensive care; assess and provide care at paediatric discretion; or, no choice expressed).

Health service factors available for inclusion were restricted to whether the mother had been transferred antentally (assumed to be true if the birth hospital was different from the intended place of delivery at booking and the indicated reason for a difference was *in utero* transfer), and the level of neonatal care available at the delivering hospital. Maternal socio-economic status was based upon the Index of Multiple Deprivation 2007, using main residential postcode at the time of delivery.

8.4 Populations studied

When considering the effect of antenatal interventions, it is important to ensure that the appropriate population is selected for inclusion, otherwise the results may be biased. The population of women with a fetus that is alive at both the time of admission to hospital and at the time when active monitoring of labour is commenced would appear to be the appropriate population, as discussed in section 8.2. This is evidenced by editorial support [218] for the same population selection criteria being used in a study relating place of birth to survival that was also conducted using the EPICure 2 cohort and that I was involved in.[70]

However, it can also be argued that this is *not* the appropriate population to use, as some women will be in established labour at presentation and progress so rapidly to delivery that there is no time for antenatal intervention. For example, if it is assumed that births from these pregnancies will always have a poorer outcome (women who are not monitored in the final few days of pregnancy forming a more risky population than women who are monitored), including these women in a comparison of women who receive steroids to those who don't could bias the results such that an impact from steroids is exaggerated. Conversely, selecting too small a population – e.g. if it were possible to select only those women who had been in hospital for a minimum of two days – the sample may become biased to exclude those with a poorer outcome, thus decreasing the size of any effect seen. Both of these scenarios assume that there is an improved outcome following antenatal steroid administration.

In order to understand this better, possible pathways may be drawn, connecting events in the order in which they are likely to occur; examples are shown in figure 8.1. The individual pathways may be combined into a more complex, overall diagram depicting directions of flow for individual subjects. This technique – which combines prior knowledge (regarding biological processes of health and disease) with key descriptive facts about the study cohort – was followed in order to determine the most relevant populations in which to investigate the obstetric antecedents chosen as exposures.

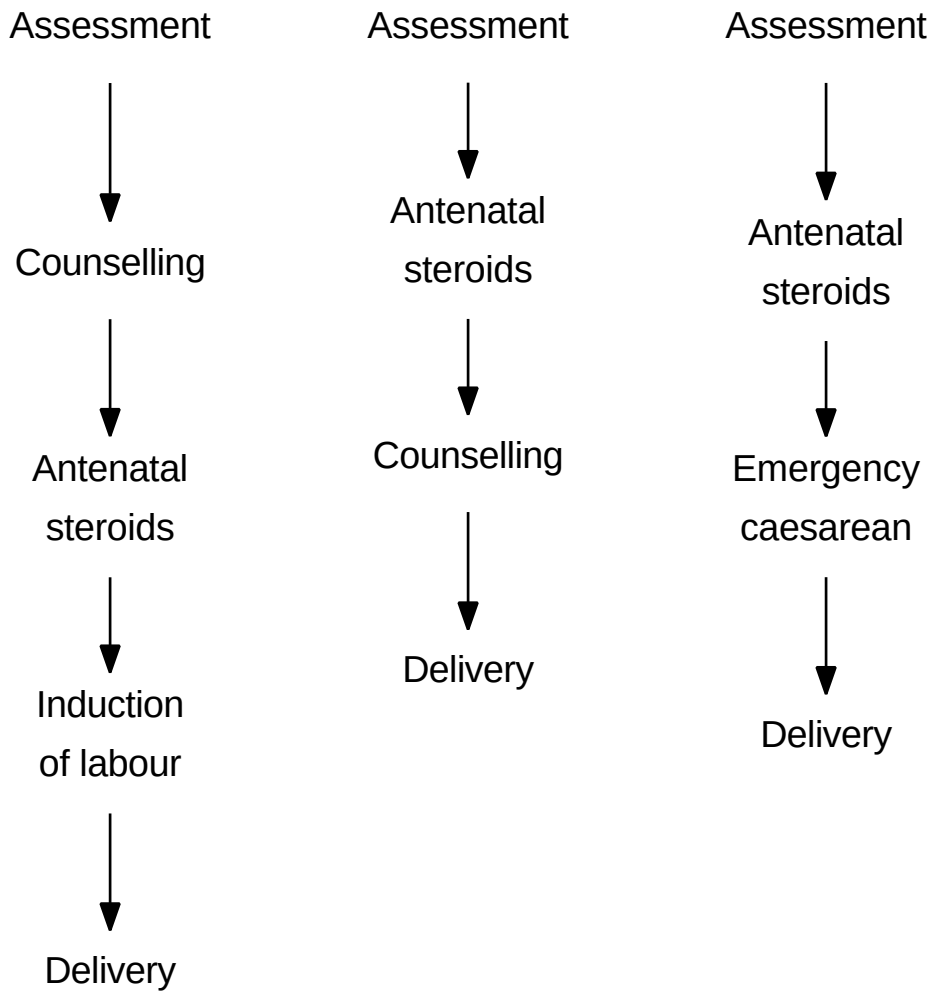
First, it was assumed that all women were equal at entry into the study (i.e. those women with a live fetus at both admission and the commencement of active monitoring of labour), other than the specific information that was available (such as demographics,

Figure 8.1: Pathways of potential scenarios for women presenting to hospital with a pregnancy-related complication before 27 weeks gestational age.

(a) Scenario 1: a mother with pre-eclampsia who requires induced or operative delivery, is first counselled about the planned delivery, then provided with steroids.

(b) Scenario 2: a mother with threatened preterm labour who received steroids on arrival, then later receives counselling prior to spontaneous delivery.

(c) Scenario 3: a mother presents with antepartum haemorrhage, receives steroids but then proceeds to have an abruption requiring Caesarean delivery.



background factors, and complications of the pregnancy). This formed the baseline population.

Next, the population of women who had received counselling was considered. Using this population made the assumption that, if the women had adequate time to receive counselling, there would also have been sufficient time to provide antenatal interventions such as steroids or tocolysis. A comparative population – the group of women who had received an antenatal intervention – could also be used in this way: the assumption being that if they were able to receive treatment there should also have been time for counselling and discussion about the management. Furthermore, in the latter group, there clearly *was* time for intervention, hence the use (or not) of other interventions may also be studied.

8.5 Data preparation

Variables were intially categorised into fixed or background, pregnancy, obstetric clinical management, counselling and delivery related factors (table 8.1). Data within and between different groups were cross-tabulated to explore relationships; this aided the construction of diagrams of potential causal pathways detailed in section 8.4. Missing data, where they formed a large proportion of the data, were accounted for by recoding variables to show whether a response was received or not, and by using adjusted populations in sensitivity analyses.

Table 8.1: Variables used in the analyses of obstetric antecedents of prematurity, divided according to relative time period.

Time Period	Variables
Background/fixed factors	Maternal factors: ethnicity, age, smoking status, past medical history, past obstetric history, body mass index, socio-economic status.
	Other factors: fetal sex, gestational age at birth, gestational age at booking
Pregnancy-related factors	Obstetric complications, fetal compromise, (suspected) chorioamnionitis
Obstetric management factors	Antenatal steroids, tocolysis, inutero transfer, maternal antibiotics
Delivery factors	Mode of delivery, NICU level, fetal presentation, labour type
Outcome measures	Heart rate > 100 at 5 minutes
	Perinatal death

8.6 Statistical methods

The primary endpoints for all analyses were the presence of a heart rate greater than 100bpm at five minutes after birth, or perinatal death. The effects on these outcomes of three specific factors relating to obstetric management were evaluated: the administration of antenatal steroids, the use of tocolysis and the mode of delivery (Caesarean section compared to a baseline of vaginal delivery). Each exposure-outcome pair was investigated in a separate analysis, with remaining variables considered as potential confounders; consequently, there were six primary regression analyses. All investigations were conducted using R.[219]

Following the detailed descriptive analysis of the overall data set, causal diagrams were used to aid in the identification of potential confounders for each of the planned analyses. An example diagram is shown in figure 8.2, demonstrating potential confounders of the relationships between delivery type (vaginal or Caesarean) and the two outcomes, perinatal death and birth in a good condition.

For each analysis, the unadjusted odds ratio (OR) of the exposure on the outcome was calculated, followed by the OR for each of the identified potential confounding variables in association with both the exposure and the outcome. Univariable analysis was performed using the Mantel-Haenszel approach (or univariable logistic regression if the exposure had greater than one level) to determine which factors had an important effect. Evidence for effect modification by variables was assessed using a chi-squared test for homogeneity or likelihood ratio test for each method, respectively; for these, a p-value of < 0.1 was accepted.

Next, multivariable analysis was employed using logistic regression with a forward step-wise approach. This involved introducing variables identified in the univariate analysis as having a confounding effect into the model one by one, each time testing whether it altered the relationship between the exposure and the outcome before deciding whether to retain that variable or not. This commenced with the variable that exhibited the strongest confounding effect, and proceeding in decreasing order of the confounding effect. A Wald test for association was used to assess the adjusted effect of individual factors on the outcome, with the likelihood ratio (LRT) test used to assess the relative importance between nested models. The significance level was set at $p < 0.05$. This was followed by assessment for possible interaction terms in the final

Figure 8.2: Causal diagram showing potential confounders of the relationship between an exposure of mode of delivery and two possible outcomes: perinatal death, or birth in a "good" condition (defined as the presence of a heart rate > 100 bpm at 5 minutes of age.

model. Models were developed such that all biological/clinical variables were included prior to variables related to antenatal counselling.

8.7 Sensitivity analyses

As well as using the baseline population of all women with a live fetus at presentation and the start of active monitoring during labour, the effect of steroids was investigated in the population of women who were recorded as having received counselling. Using this population ensured there was an adequate spread of subjects between those who did and did not receive the intervention. It also meant the assumption about missing data – that participants who were missing responses to the counselling had not had the relevant discussion – was more likely to be correct. For the analyses using tocolysis and mode of delivery as exposures, sensitivity analyses were performed using the group of women who had received steroids. This was based on the assumption that there was an active intention for the babies of women who had received steroids to survive.

8.8 Summary

The investigation of obstetric antecedents of prematurity focuses on three interventions that may impact the immediate outcome of the fetus at birth. These are the administration of antenatal steroids, the use of tocolysis to extend the length of the pregnancy, and performing a Caesarean section to facilitate delivery. Described in this chapter were the methods by which the investigation was carried out. The study population, taken from the 2006 EPICure 2 cohort, comprised mothers of singleton pregnancies where the fetus was alive at admission to hospital and at the point at which active monitoring of the labour began; women who had terminations of pregnancy were excluded. The choice of variables was then described in section 8.3 – particularly, of the exposures and of the outcomes, perinatal death (defined as death during labour or in the delivery room) and birth in a good condition (babies with a heart rate > 100 at 5 minutes of age).

In section 8.4, the population available for study was discussed in further detail, with an emphasis on the construction of potential causal diagrams using a combination of prior knowledge and the evidence supplied by the data. The use of these diagrams

in identifying potential confounders of the relationship between exposure and outcome for each of the analyses was highlighted using an example (figure 8.2). Next, there was a description of the statistical methods – specifically, the techniques used were the Mantel-Haenszel technique, and multivariate logistic regression analysis – that were used in the investigations. The chapter finished with an explanation of the sensitivity analyses that were carried out.

Chapter 9

Results of the obstetric antecedents of prematurity investigation

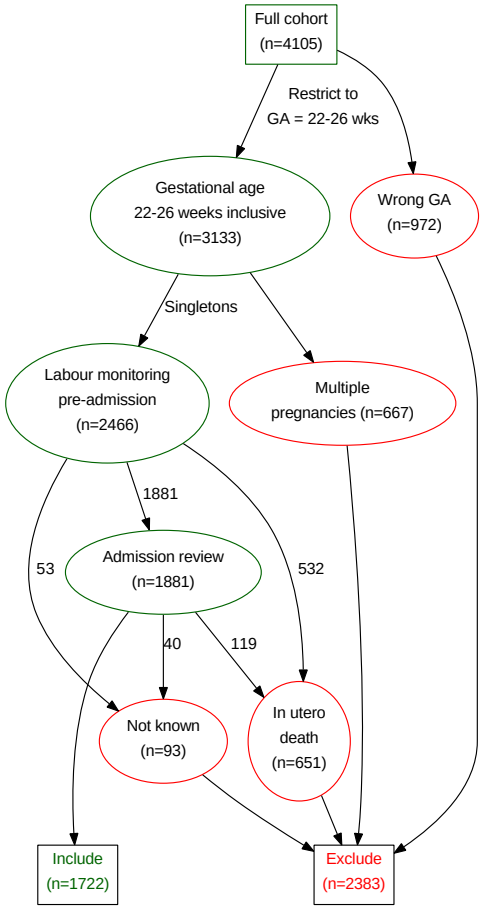
This chapter presents the results from the investigation into obstetric antecedents of prematurity – specifically, looking at the impact on condition at birth and perinatal death of antenatal steroids, tocolysis and Caesarean delivery. The first part describes the data that are used, and outlines the pathways that were developed to guide subsequent analyses. Each of the exposures is then discussed in its own section, presenting results for both of the outcomes as well as those of the relevant sensitivity analyses. There is a brief synopsis at the end, with further discussion taking place in chapter 14.

9.1 Data description

A total of 2,466 singleton pregnancies were delivered between 22^{+0} and 26^{+6} weeks gestation in England in 2006. Among these, there were 532 fetal deaths prior to maternal admission to hospital; a further 159 were alive at admission but died prior to the onset of active monitoring during labour, and 53 had insufficient data to confirm whether they met the inclusion criteria or not. Thus 1,722 women with a fetus known to be alive both at admission to hospital and at commencement of labour monitoring were included, as shown in figure 9.1.

Some form of medical therapy was administered to 1,395 women – therapies included

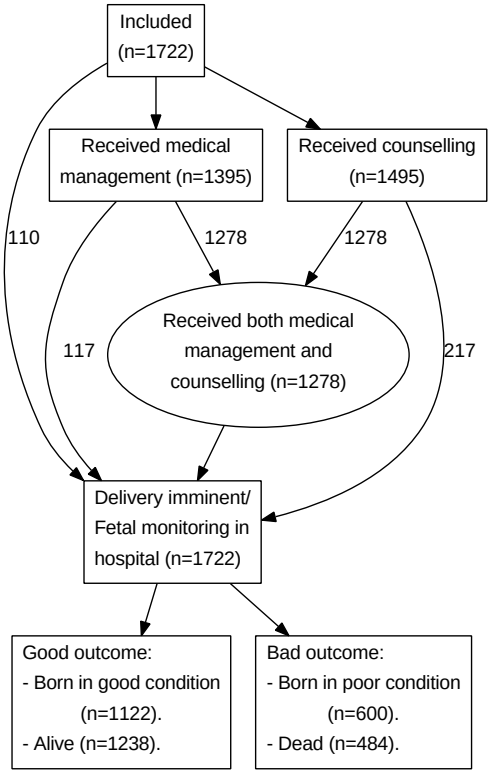
Figure 9.1: Flow diagram showing inclusion and exclusion numbers for participant mother-baby pairs from the EPICure 2 cohort of babies born between 22 and 26 completed weeks gestation in England in 2006. GA: gestational age.



antibiotics, tocolysis, antenatal steroids and *in utero* transfer; of these women, 1,278 were also recorded as having received counselling (figure 9.2). A total of 1,213 women received steroids and 437 tocolysis, but only 406 women received both, meaning that there were 31 women who received tocolysis without receiving steroids as well. The explanatory diagram that was developed is shown in figure 9.3; like the other diagrams, it assumes that administration of therapies is ordered.

Overall, there were very few data items missing: aside from data related to counselling, only three variables had fewer than 95% of values completed: gestational age at booking (90.1%), maternal body mass index (86.0%) and parity (45.0%). Of the five

Figure 9.2: Flow diagram showing inclusion numbers, main pathways and principal outcomes for an assessment of the impact of obstetric interventions on condition at birth and perinatal death in the EPICure 2 cohort of babies born between 22 and 26 completed weeks gestation in England in 2006.



questions asked relating to possible counselling, 1,495 responses provided some indication that counselling had occurred, including 351 cases where a decision was made not to perform emergency Caesarean section in cases of fetal distress, 465 women with whom withholding care was discussed, and 727 who expressed a choice regarding provision of resuscitation – and, potentially, intensive care – to a live born baby. Additionally, it is known that 1,287 women were counselled by a senior obstetrician and 1,246 by a paediatrician (see figure 9.4).

Figure 9.3: Flow diagram showing detail about clinical factors measured in the EPICure 2 cohort of babies born between 22 and 26 completed weeks gestation in England in 2006.

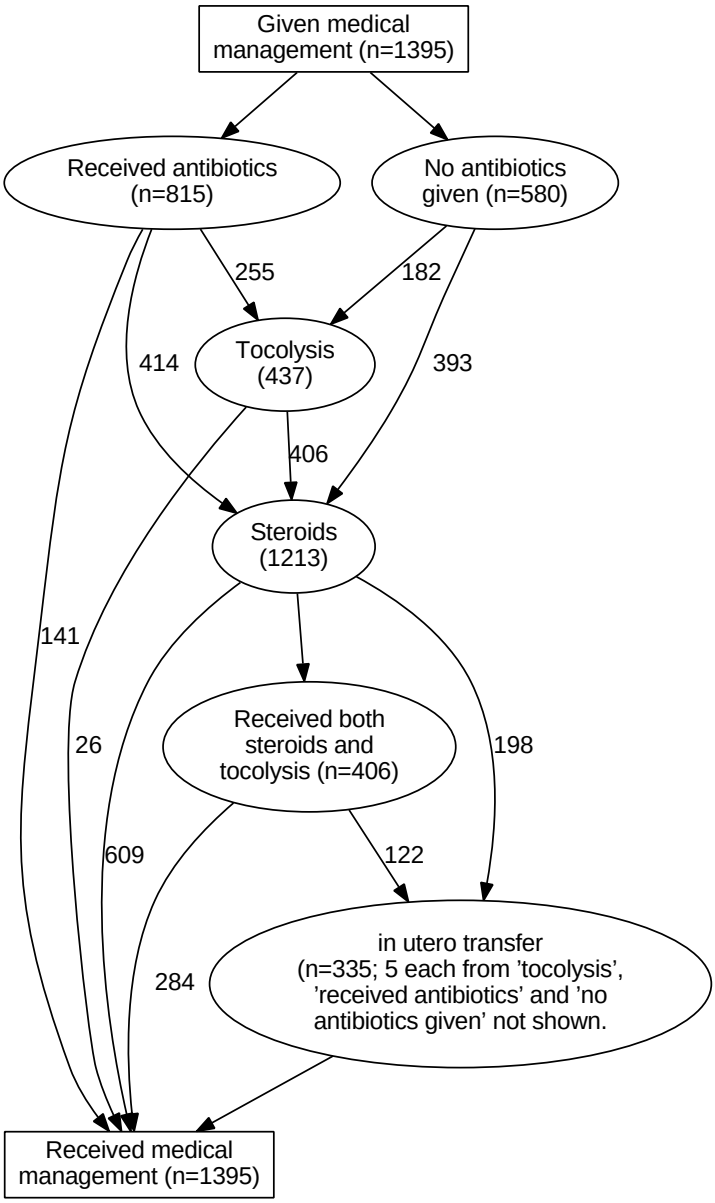


Figure 9.4: Flow diagram showing detail about counselling factors measured in the EPICure 2 cohort of babies born between 22 and 26 completed weeks gestation in England in 2006.

9. RESULTS OF THE OBSTETRIC ANTECEDENTS OF PREMATURITY INVESTIGATION

The pathway shown in figure 9.5 describes the final set of variables, relating to events that may happen when delivery is imminent: if the mother or fetus are unwell, labour may be induced or Caesarean delivery performed; similarly, in cases of spontaneous labour, operative delivery may still be preferred. In either case, resuscitation of the newborn may be attempted and therefore impact on either of the two outcomes used in this study: heart rate at 5 minutes of age, or death in the delivery room. However, as it lies on the causal pathway, it is clear that it is incorrect to make any adjustment for whether resuscitation was carried out.[151]

Four hundred and eighty-four (28.1%) babies died during labour or in the delivery room, of whom 28 were *not* in a poor condition at five minutes of age; overall, 600 (34.8%) babies had a heart rate below 100 beats per minute five minutes after delivery. Each of the three exposures was associated with birth in good condition: unadjusted odds ratio (OR) for a partial course of antenatal steroids was 5.08 (95% confidence interval (CI): 3.84 – 6.72); for a full course 7.24 (95% CI: 5.60 – 9.36) and the trend was statistically significant (Chi-squared test for trend $p< 0.001$); OR for tocolysis was 2.24 (95% CI: 1.75 – 2.90) and for delivery by Caesarean section 4.21 (95% CI: 3.06 – 5.80). Perinatal death was also associated with each outcome. Administration of steroids (partial course: 0.13, 95% CI: 0.10 – 0.18; full course: 0.09, 95% CI: 0.07 – 0.12), tocolysis (0.33, 95% CI: 0.24 – 0.44) and mode of delivery (Caesarean delivery: 0.16, 95% CI: 0.10 – 0.24) were all associated with lowered unadjusted odds ratio.

Both gestational age (in weeks) and fetal sex demonstrated important associations with condition at birth and perinatal death. Both outcomes were also associated with placental abruption, pre-eclampsia, *in utero* transfer, spontaneous labour, non-cephalic presentation and delivery in a centre with a level 3 neonatal intensive care unit. There were further important associations between all the counselling variables and both outcomes. Complete univariate associations are shown in table 9.1.

Comparing the main population with the sub-populations used in the sensitivity analyses demonstrated important differences in distributions for key variables in the steroids-only population. A greater proportion of women received tocolysis (33.47% compared with 25.41% in the overall population, χ^2 p-value < 0.01) and Caesarean delivery (25.72% v. 20.73%, $p< 0.01$), and the gestational age range at which deliveries occured was shifted to older age groups (table 9.2). There were no differences in terms of fetal sex or in the decision to not perform Caesarean section in the presence of

fetal distress. In the population that had evidence of receiving counselling, the use of antenatal steroids showed an important difference, with a greater proportion receiving steroids than in the overall population. There was no difference in the use of tocolysis or surgical delivery, however, nor in fetal sex, gestational age profile or either of the outcomes. Data for key variables are presented in table 9.2.

Figure 9.5: Flow diagram showing detail about delivery measured in the EPICure 2 cohort of babies born between 22 and 26 completed weeks gestation in England in 2006. Not all numbers add to 1,722 due to missing data in some of the variables.

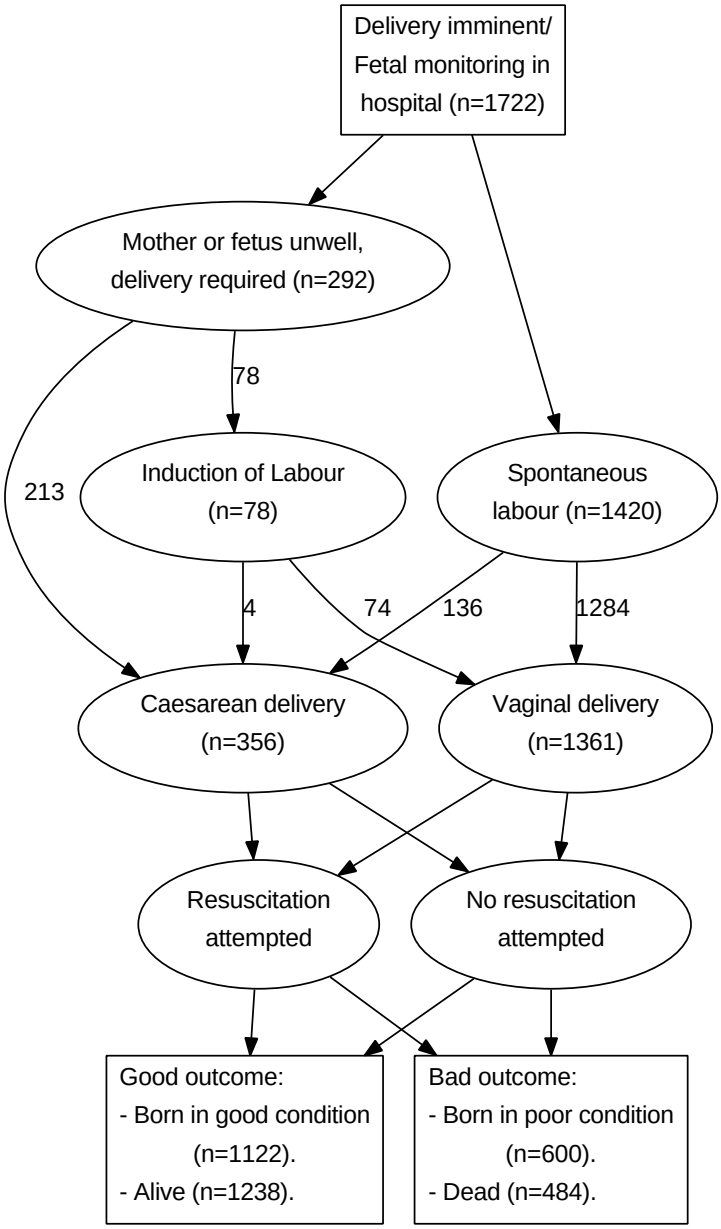


Table 9.1: Descriptive statistics of explanatory variables in relation to presence of a heart rate greater than 100 at five minutes of age ("good condition") or to death during labour or in the delivery room ("perinatal death") in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or decision to perform Caesarean section.

Variable (Levels)	**N**	**(%)**	**Born in good condition**	**(HR %)**	**OR**	**(95% CI)**	**P-value**	**Perinatal death**	**(%)**	**OR**	**(95% CI)**	**P-value**
Antenatal steroid course[a] *N=1701*												
none	488	28.3	172	35.2	1.00	(—)	—	301	61.7	1.00	—	(—)
partial	452	26.2	332	73.5	5.08	(3.84 – 6.72)	< 0.001	79	17.5	0.13	(0.10 – 0.18)	< 0.001
full	761	44.2	607	79.8	7.24	(5.60 – 9.36)	< 0.001	99	13.0	0.09	(0.07 – 0.12)	< 0.001
Tocolysis *N=1720*												
no	1283	74.5	782	61.0	1.00	(—)	—	423	33.0	1.00	(—)	—
yes	437	25.4	340	77.8	2.25	(1.75 – 2.89)	< 0.001	61	14.0	0.33	(0.25 – 0.44)	< 0.001
Delivery by Caesarean *N=1717*												
no	1361	79.0	814	59.8	1.00	—	(—)	456	33.5	1.00	—	(—)
yes	356	20.7	307	86.2	4.21	(3.06 – 5.80)	< 0.001	27	7.6	0.16	(0.11 – 0.24)	< 0.001
Gestational age (weeks) *N=1722*												
22	204	11.8	22	10.8	0.02	(0.01 – 0.03)	< 0.001	190	93.1	297.14	(145.89 – 605.20)	< 0.001
23	281	16.3	127	45.2	0.12	(0.09 – 0.18)	< 0.001	141	50.2	22.05	(13.16 – 36.94)	< 0.001
24	366	21.3	243	66.4	0.30	(0.21 – 0.42)	< 0.001	89	24.3	7.03	(4.19 – 11.81)	< 0.001
25	436	25.3	352	80.7	0.63	(0.44 – 0.91)	0.02	45	10.3	2.52	(1.45 – 4.38)	< 0.001
26	435	25.3	378	86.9	1.00	(—)	—	19	4.4	1.00	(—)	—
Fetal sex *N=1719*												
Female	804	46.7	549	68.3	1.00	(—)	—	201	25.0	1.00	(—)	—
Male	915	53.1	572	62.5	0.77	(0.63 – 0.95)	0.01	280	30.6	1.32	(1.07 – 1.64)	0.01
Ethnicity *N=1722*												
White	1118	64.9	725	64.8	1.00	(—)	—	312	27.9	1.00	(—)	—

Continued on next page...

Table 9.1: (Continued) Descriptive statistics of explanatory variables in relation to condition at birth or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or decision to perform Caesarean section.

Variable (Levels)	N	(%)	Born in good condition	(HR %)	OR	(95% CI)	P-value	Perinatal death	(%)	OR	(95% CI)	P-value
Black	334	19.4	225	67.4	1.12	(0.86 – 1.45)	0.43	96	28.7	1.04	(0.79 – 1.37)	0.82
Indian	183	10.6	114	62.3	0.90	(0.65 – 1.24)	0.56	53	29.0	1.05	(0.75 – 1.49)	0.84
Other	87	5.1	58	66.7	1.08	(0.68 – 1.72)	0.82	23	26.4	0.93	(0.57 – 1.52)	0.86
Maternal age (categories)	*N=1718*											
<20 years	157	9.1	98	62.4	1.00	(—)	—	50	31.8	1.00	(—)	—
20-24 years	344	20.0	214	62.2	0.99	(0.67 – 1.46)	1.00	96	27.9	0.83	(0.55 – 1.25)	0.43
25-29 years	467	27.1	305	65.3	1.13	(0.78 – 1.65)	0.58	129	27.6	0.82	(0.55 – 1.21)	0.36
30-34 years	388	22.5	253	65.2	1.13	(0.77 – 1.66)	0.61	116	29.9	0.91	(0.61 – 1.36)	0.73
35-39 years	278	16.1	193	69.4	1.37	(0.91 – 2.06)	0.17	73	26.3	0.76	(0.50 – 1.17)	0.26
>40 years	84	4.9	58	69.0	1.34	(0.76 – 2.36)	0.38	19	22.6	0.63	(0.34 – 1.15)	0.17
Gestational age at booking	*N=1551*											
<10 weeks	463	26.9	295	63.7	1.00	(—)	—	136	29.4	1.00	(—)	—
10-12 weeks	487	28.3	314	64.5	1.03	(0.79 – 1.35)	0.86	142	29.2	0.99	(0.75 – 1.31)	1.00
13-19 weeks	487	28.3	320	65.7	1.09	(0.84 – 1.42)	0.57	137	28.1	0.94	(0.71 – 1.25)	0.73
20+ weeks	114	6.6	80	70.2	1.34	(0.86 – 2.09)	0.24	26	22.8	0.71	(0.44 – 1.15)	0.20
Maternal smoking	*N=1722*											
no	1303	75.7	851	65.3	1.00	(—)	—	365	28.0	1.00	(—)	—
yes	419	24.3	271	64.7	0.97	(0.77 – 1.22)	0.86	119	28.4	1.02	(0.80 – 1.30)	0.93
Maternal body mass index (BMI)	*N=1481*											
<18.5 (underweight)	56	3.3	31	55.4	1.00	(—)	—	17	30.4	1.00	(—)	—
18.5-24.9 (appropriate)	658	38.2	421	64.0	1.43	(0.83 – 2.48)	0.25	189	28.7	0.92	(0.51 – 1.67)	0.92
>=25 (overweight)	767	44.5	497	64.8	1.48	(0.86 – 2.57)	0.2	227	29.6	0.96	(0.53 – 1.74)	1.00
Pre-pregnancy diabetes, type I or II	*N=1709*											
no/miss	1690	98.1	1104	65.3	1.00	(—)	—	476	28.2	1.00	(—)	—

Continued on next page. . .

Table 9.1: (Continued) Descriptive statistics of explanatory variables in relation to condition at birth or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or decision to perform Caesarean section.

Variable (Levels)	N	(%)	Born in good condition	(HR %)	OR	(95% CI)	P-value	Perinatal death	(%)	OR	(95% CI)	P-value
yes	19	1.1	14	73.7	1.49	(0.53 – 4.15)	0.60	3	15.8	0.48	(0.14 – 1.65)	0.35
Non-insulin dependent gestational diabetes *N=1709*												
no/miss	1702	98.8	1114	65.5	1.00	(—)	—	477	28.0	1.00	(—)	—
yes	7	0.4	4	57.1	0.70	(0.16 – 3.15)	0.95	2	28.6	1.03	(0.2 – 5.31)	1.00
Essential hypertension, on treatment at booking *N=1709*												
no/miss	1656	96.2	1080	65.2	1.00	(—)	—	467	28.2	1.00	(—)	—
yes	53	3.1	38	71.7	1.35	(0.74 – 2.48)	0.41	12	22.6	0.75	(0.39 – 1.43)	0.46
Epilepsy *N=1709*												
no/miss	1697	98.5	1110	65.4	1.00	(—)	—	476	28.0	1.00	(—)	—
yes	12	0.7	8	66.7	1.06	(0.32 – 3.53)	1.00	3	25.0	0.86	(0.23 – 3.17)	1.00
Other important medical history *N=1709*												
no/miss	1694	98.4	1107	65.3	1.00	(—)	—	477	28.2	1.00	(—)	—
yes	15	0.9	11	73.3	1.46	(0.46 – 4.60)	0.71	2	13.3	0.39	(0.09 – 1.75)	0.33
Any maternal medical problems *N=1711*												
no	1610	93.5	1046	65.0	1.00	(—)	—	458	28.4	1.00	(—)	—
yes	101	5.9	72	71.3	1.34	(0.86 – 2.08)	0.24	21	20.8	0.66	(0.40 – 1.08)	0.12
Parity (categorical) *N=774*												
Primip	365	21.2	244	66.8	1.00	(—)	—	98	26.8	1.00	(—)	—
G1	344	20.0	215	62.5	0.83	(0.61 – 1.13)	0.26	90	26.2	0.97	(0.69 – 1.35)	0.9
2+ prev. babies	65	3.8	45	69.2	1.12	(0.63 – 1.97)	0.82	16	24.6	0.89	(0.48 – 1.64)	0.82
Primiparous (first registrable birth) *N=1701*												
no	858	49.8	560	65.3	1.00	(—)	—	228	26.6	1.00	(—)	—
yes	843	49.0	554	65.7	1.02	(0.84 – 1.25)	0.89	246	29.2	1.14	(0.92 – 1.41)	0.25
Primigravid (first pregnancy) *N=1712*												
No	1180	68.5	765	64.8	1.00	(—)	—	334	28.3	1.00	(—)	—

Continued on next page. . .

Table 9.1: (Continued) Descriptive statistics of explanatory variables in relation to condition at birth or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or decision to perform Caesarean section.

Variable (Levels)	N	(%)	Born in good condition	(HR %)	OR	(95% CI)	P-value	Perinatal death	(%)	OR	(95% CI)	P-value
Yes	532	30.9	353	66.4	1.07	(0.86 – 1.33)	0.58	146	27.4	0.96	(0.76 – 1.2)	0.76
Previous Caesarean *N=1677*												
no	1445	83.9	951	65.8	1.00	(—)	—	402	27.8	1.00	(—)	—
yes	232	13.5	154	66.4	1.03	(0.76 – 1.38)	0.92	58	25.0	0.86	(0.63 – 1.19)	0.42
Previous premature birth *N=1687*												
no	1420	82.5	941	66.3	1.00	(—)	—	395	27.8	1.00	(—)	—
yes	267	15.5	168	62.9	0.86	(0.66 – 1.13)	0.32	71	26.6	0.94	(0.70 – 1.26)	0.74
IMD national quintiles *N=1700*												
Least	615	35.7	392	63.7	1.00	(—)	—	189	30.7	1.00	(—)	—
2	582	33.8	378	64.9	1.05	(0.83 – 1.34)	0.71	169	29.0	0.92	(0.72 – 1.18)	0.56
3	389	22.6	261	67.1	1.16	(0.89 – 1.52)	0.31	101	26.0	0.79	(0.59 – 1.05)	0.12
4	112	6.5	75	67.0	1.15	(0.75 – 1.77)	0.58	21	18.8	0.52	(0.31 – 0.86)	0.01
Most	2	0.1	2	100.0	∞	(0.00 – ∞)	0.74	0	0.0	0.00	(0.00 – ∞)	0.86
IMD study quintiles *N=1700*												
Least	532	30.9	336	63.2	1.00	(—)	—	163	30.6	1.00	(—)	—
2	506	29.4	326	64.4	1.06	(0.82 – 1.36)	0.72	151	29.8	0.96	(0.74 – 1.26)	0.83
3	412	23.9	278	67.5	1.21	(0.92 – 1.59)	0.19	111	26.9	0.83	(0.63 – 1.11)	0.24
4	214	12.4	146	68.2	1.25	(0.89 – 1.76)	0.22	49	22.9	0.67	(0.47 – 0.97)	0.04
Most	36	2.1	22	61.1	0.92	(0.46 – 1.83)	0.95	6	16.7	0.45	(0.18 – 1.11)	0.11
Any counselling *N=1722*												
No	227	13.2	132	58.1	1.00	(—)	—	67	29.5	1.00	(—)	—
Yes	1495	86.8	990	66.2	1.41	(1.06 – 1.88)	0.02	417	27.9	0.92	(0.68 – 1.26)	0.67
Were any options discussed *N=1722*												
No	773	44.9	558	72.2	1.00	(—)	—	149	19.3	1.00	(—)	—
Yes	949	55.1	564	59.4	0.56	(0.46 – 0.69)	< 0.001	335	35.3	2.28	(1.83 – 2.86)	< 0.001

Continued on next page. . .

Table 9.1: (Continued) Descriptive statistics of explanatory variables in relation to condition at birth or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or decision to perform Caesarean section.

Variable (Levels)	N	(%)	Born in good condition	(HR %)	OR	(95% CI)	P-value	Perinatal death	(%)	OR	(95% CI)	P-value
Paediatric counselling *N=1633*												
no	387	22.5	216	55.8	1.00	(—)	—	136	35.1	1.00	(—)	—
yes	1246	72.4	846	67.9	1.67	(1.33 – 2.11)	< 0.001	326	26.2	0.65	(0.51 – 0.83)	< 0.001
Choice made about newborn care (resus/IC) *N=1235*												
none expressed	508	29.5	336	66.1	1.00	(—)	—	131	25.8	1.00	(—)	—
full NIC	391	22.7	309	79.0	1.93	(1.42 – 2.62)	< 0.001	44	11.3	0.36	(0.25 – 0.53)	< 0.001
withhold IC	45	2.6	5	11.1	0.06	(0.02 – 0.17)	< 0.001	43	95.6	61.87	(14.78 – 258.98)	< 0.001
paed discretion	291	16.9	139	47.8	0.47	(0.35 – 0.63)	< 0.001	140	48.1	2.67	(1.97 – 3.62)	< 0.001
Decision not to perform C/S for fetal distress *N=1158*												
No	807	46.9	578	71.6	1.00	(—)	—	169	20.9	1.00	(—)	—
Yes	351	20.4	175	49.9	0.39	(0.30 – 0.51)	< 0.001	158	45.0	3.09	(2.36 – 4.05)	< 0.001
Withholding care discussed *N=1178*												
No	713	41.4	539	75.6	1.00	(—)	—	106	14.9	1.00	(—)	—
Yes	465	27.0	218	46.9	0.28	(0.22 – 0.37)	< 0.001	239	51.4	6.06	(4.60 – 7.97)	< 0.001
Senior obstetric counselling *N=1722*												
No	435	25.3	270	62.1	1.00	(—)	—	120	27.6	1.00	(—)	—
Yes	1287	74.7	852	66.2	1.2	(0.96 – 1.50)	0.13	364	28.3	(1.04	0.81) – 1.32	0.83
PROM >24 hrs *N=1710*												
no	1211	70.3	793	65.5	1.00	(—)	—	344	28.4	1.00	(—)	—
yes	499	29.0	324	64.9	0.98	(0.78 – 1.21)	0.87	136	27.3	0.94	(0.75 – 1.19)	0.67
Abruption *N=1710*												
no	1577	91.6	1043	66.1	1.00	(—)	—	428	27.1	1.00	(—)	—
yes	133	7.7	74	55.6	0.64	(0.45 – 0.92)	0.02	52	39.1	1.72	(1.20 – 2.48)	< 0.001
APH >20 weeks *N=1710*												
no	1406	81.6	916	65.1	1.00	(—)	—	393	28	1.00	(—)	—

Continued on next page. . .

Table 9.1: (Continued) Descriptive statistics of explanatory variables in relation to condition at birth or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or decision to perform Caesarean section.

Variable (Levels)	N	(%)	Born in good condition	(HR %)	OR	(95% CI)	P-value	Perinatal death	(%)	OR	(95% CI)	P-value
yes	304	17.7	201	66.1	1.04	(0.80 – 1.36)	0.80	87	28.6	1.03	(0.79 – 1.36)	0.87
Pre-eclampsia *N=1710*												
no	1586	92.1	1014	63.9	1.00	(—)	—	465	29.3	1.00	(—)	—
yes	124	7.2	103	83.1	2.77	(1.71 – 4.47)	< 0.001	15	12.1	0.33	(0.19 – 0.58)	< 0.001
Cervical suture *N=1710*												
no	1613	93.7	1049	65.0	1.00	(—)	—	460	28.5	1.00	(—)	—
yes	97	5.6	68	70.1	1.26	(0.81 – 1.97)	0.36	20	20.6	0.65	(0.39 – 1.08)	0.12
Obstetric complication in pregnancy *N=1712*												
no	723	42	454	62.8	1.00	(—)	—	219	30.3	1.00	(—)	—
yes	989	57.4	663	67.0	1.21	(0.99 – 1.47)	0.08	261	26.4	0.83	(0.67 – 1.02)	0.09
Fetal complication *N=1695*												
none	1545	89.7	1005	65	1.00	(—)	—	435	28.2	1.00	(—)	—
iugr	91	5.3	68	74.7	1.59	(0.98 – 2.58)	0.08	19	20.9	0.67	(0.40 – 1.13)	0.17
oligo	38	2.2	22	57.9	0.74	(0.38 – 1.42)	0.46	13	34.2	1.33	(0.67 – 2.62)	0.52
both	21	1.2	15	71.4	1.34	(0.52 – 3.48)	0.70	6	28.6	1.02	(0.39 – 2.65)	1.00
Chorioamnionitis - definite or suspected *N=1683*												
no	1285	74.6	836	65.1	1.00	(—)	—	355	27.6	1.00	(—)	—
yes	398	23.1	262	65.8	1.03	(0.82 – 1.31)	0.82	119	29.9	1.12	(0.87 – 1.43)	0.41
Labour type (binary) *N=1712*												
none	292	17.0	233	79.8	1.00	(—)	—	48	16.4	1.00	(—)	—
spont	1420	82.5	884	62.3	0.42	(0.31 – 0.57)	< 0.001	434	30.6	2.24	(1.61 – 3.11)	< 0.001
Tocolysis (by type) *N=1708*												
None	1283	74.5	782	61.0	1.00	(—)	—	423	33.0	1.00	(—)	—
Atosiban	143	8.3	118	82.5	3.02	(1.94 – 4.72)	< 0.001	14	9.8	0.22	(0.13 – 0.39)	< 0.001
Ritodrine	7	0.4	6	85.7	3.84	(0.46 – 32.02)	0.34	1	14.3	0.34	(0.04 – 2.82)	0.52

Continued on next page. . .

Table 9.1: (Continued) Descriptive statistics of explanatory variables in relation to condition at birth or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or decision to perform Caesarean section.

Variable (Levels)	N	(%)	Born in good condition	(HR %)	OR	(95% CI)	P-value	Perinatal death	(%)	OR	(95% CI)	P-value
Indometacin	27	1.6	24	88.9	5.13	(1.54 – 17.11)	0.01	2	7.4	0.16	(0.04 – 0.69)	0.01
Nifedipine	189	11	140	74.1	1.83	(1.30 – 2.58)	< 0.001	33	17.5	0.43	(0.29 – 0.64)	< 0.001
Other	19	1.1	16	84.2	3.42	(0.99 – 11.79)	0.07	1	5.3	0.11	(0.02 – 0.85)	0.02
More than one	40	2.3	31	77.5	2.21	(1.04 – 4.67)	0.05	6	15.0	0.36	(0.15 – 0.86)	0.03
Antenatal (in utero) transfer *N=1705*												
no	1370	79.6	844	61.6	1.00	(—)	—	434	31.7	1.00	(—)	—
yes	335	19.5	276	82.4	2.92	(2.16 – 3.94)	< 0.001	39	11.6	0.28	(0.20 – 0.40)	< 0.001
Maternal antibiotics in labour *N=1696*												
none	881	51.2	581	65.9	1.00	(—)	—	248	28.1	1.00	(—)	—
proph	468	27.2	309	66.0	1.00	(0.79 – 1.27)	1.00	122	26.1	0.90	(0.70 – 1.16)	0.45
treatment	347	20.2	217	62.5	0.86	(0.67 – 1.12)	0.29	108	31.1	1.15	(0.88 – 1.51)	0.33
Presentation *N=1643*												
Cephalic	915	53.1	657	71.8	1.00	(—)	—	199	21.7	1.00	(—)	—
Breech	653	37.9	379	58.0	0.54	(0.44 – 0.67)	< 0.001	234	35.8	2.01	(1.61 – 2.51)	< 0.001
Other	75	4.4	47	62.7	0.66	(0.40 – 1.08)	0.12	21	28	1.40	(0.83 – 2.37)	0.27
Unit NHS category *N=1693*												
1	147	8.5	70	47.6	0.35	(0.25 – 0.50)	< 0.001	69	46.9	3.31	(2.31 – 4.74)	< 0.001
2	583	33.9	351	60.2	0.58	(0.47 – 0.73)	< 0.001	201	34.5	1.97	(1.56 – 2.48)	< 0.001
3	963	55.9	695	72.2	1.00	(—)	—	203	21.1	1.00	(—)	—

[a] Chi-squared test for trend: < 0.001 ; [b] Chi-squared test for trend: 0.049 ; [c] Chi-squared test for trend: 0.528 ;

[d] Level 1: Special Care Baby Unit; Level 2: Local Neonatal Unit; Level 3: Network Intensive Care Unit.

Table 9.2: Characteristics of key variables from the EPICure 2 cohort study of extremely preterm births in England in 2005. Mother-baby pairs were divided into three populations: the full population consisted of babies who were alive at admission to hospital and the commencement of labour monitoring or the decision to perform caesarean section (C/S). The two sub-populations studied in sensitivity analyses comprised mothers who received counselling, and those who received steroids.

Variable	"Complete" population		"Counselled" population			"Steroids only" population		
	N	%	N	%	P-value [a]	N	%	P-value [b]
Antenatal steroids	1701	—	1484	—	—	1213	—	—
None	488	28.69	358	24.12		0	0.00	
Partial	452	26.57	400	26.95	0.01	452	37.26	< 0.01
Full	761	44.74	726	48.92		761	62.74	
Tocolysis	1720	—	1495	—	—	1213	—	—
No	1283	74.59	1085	72.58	0.21	807	66.53	< 0.01
Yes	437	25.41	410	27.42		406	33.47	
Delivery mode	1717	—	1495	—	—	1213	—	—
Vaginal	1361	79.27	1170	78.26	0.51	901	74.28	< 0.01
Caesarean	356	20.73	325	21.74		312	25.72	
Gestational age	1722	—	1495	—	—	1213	—	—
22 weeks	204	11.85	165	11.04		19	1.57	
23 weeks	281	16.32	252	16.86		139	11.46	
24 weeks	366	21.25	322	21.54	0.90	303	24.98	< 0.01
25 weeks	436	25.32	391	26.15		374	30.83	
26 weeks	435	25.26	365	24.41		378	31.16	
Fetal sex	1719	—	1492	—	—	1213	—	—
Female	804	46.77	699	46.85	0.99	586	48.31	0.43
Male	915	53.23	793	53.15		627	51.69	
HR>100 5 minutes after birth	1722	—	1495	—	—	1213	—	—
No	600	34.84	505	33.78	0.55	274	22.59	< 0.01
Yes	1122	65.16	990	66.22		939	77.41	
Perinatal death	1722	—	1495	—	—	1213	—	—
No	1238	71.89	1078	72.11	0.92	1035	85.33	< 0.01
Yes	484	28.11	417	27.89		178	14.67	

Continued on next page...

Table 9.2: (Continued.) Characteristics of key variables from the EPICure 2 cohort study of extremely preterm births in England in 2005.

Variable	"Complete" population		'Counselled' population			"Steroids only" population		
	N	%	N	%	P-value [a]	N	%	P-value [b]
Counselled by...	1722	—	1495	—	—	1213	—	—
No one/not counselled	227	13.18	0	0.00		87	7.17	
Junior obstetrician only	168	9.76	168	11.24		94	7.75	
Senior obstetrician only	165	9.58	165	11.04	< 0.01	96	7.91	< 0.01
Paediatrician and junior obstetrician	418	24.27	418	27.96		353	29.10	
Paediatrician and senior obstetrician	744	43.21	744	49.77		583	48.06	
Decision to not perform C/S	1158	—	1017	—	—	820	—	—
No	807	69.69	666	65.49	0.04	571	69.63	1.00
Yes	351	30.31	351	34.51		249	30.37	
Resuscitation choice expressed	1235	—	1130	—	—	881	—	—
No choice expressed	508	41.13	403	35.66		332	37.68	
Full intensive care	391	31.66	391	34.60	0.06	346	39.27	< 0.01
Withhold intensive care	45	3.64	45	3.98		9	1.02	
At paediatric discretion	291	23.56	291	25.75		194	22.02	
Discussion about withholding care	1178	—	1048	—	—	816	—	—
Not discussed	713	60.53	583	55.63	0.02	545	66.79	0.01
Discussed	465	39.47	465	44.37		271	33.21	

[a] Chi-squared test comparing distributions between the "counselled" and "complete" populations.

[b] Chi-squared test comparing distributions between the "steroids only" and "complete" populations..

9.2 Antenatal steroids

Adjusting analyses for single risk factors demonstrated that gestational age had the largest impact on the relationship between antenatal steroids and both outcomes. For condition at birth, gestational age reduced the ORs to 2.53 (95% CI: 1.85 – 3.45) and 2.09 (95% CI: 1.51 – 2.90) for complete and partial courses of steroids, respectively

Table 9.3: Odds ratios of the effect of antenatal steroids on the presence of a heart rate greater than 100 at minutes of age in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or delivery to perform Caesarean section.

Model	Good condition at birth					
		Partial course		Full course		P-value [a]
	N	Effect	(95% CI)	Effect	(95% CI)	
Confounding only models						
Baseline	1701	5.08	(3.84 – 6.72)	7.24	(5.60 – 9.36)	—
Baseline + GA	1701	2.09	(1.51 – 2.90)	2.53	(1.85 – 3.45)	< 0.001
'Clinical' model [b]	1602	1.42	(0.99 – 2.05)	1.64	(1.14 – 2.36)	0.013
Complete model [c]	1602	1.38	(0.93 – 2.03)	1.82	(1.22 – 2.71)	< 0.001
Models with effect modification by mode of delivery [b c]						
Clinical model: Vaginal delivery	1602	1.61	(1.09 – 2.37)	1.66	(1.13 – 2.44)	0.018
Clinical model: Caesarean delivery		1.31	(0.46 – 3.69)	0.48	(0.17 – 1.39)	
Complete model: vaginal delivery		1.84	(1.20 – 2.82)	1.63	(1.08 – 2.47)	0.021
Complete model: Caesarean delivery		1.23	(0.42 – 3.61)	0.42	(0.14 – 1.26)	

[a] Likelihood ratio test p-value, comparing current to next simplest model described.

[b] The clinical model for condition at birth was adjusted for: gestational age, tocolysis, mode of delivery, NICU level at birth, *in utero* transfer, presentation at delivery.

[c] The complete models were adjusted for the same factors as the clinical models, plus: who the parents were counselled by, what choice was expressed for resuscitation, and whether a decision was made to not perform Caesarean section in the presence of fetal distress.

(table 9.3), and for perinatal death, the ORs were 0.39 (95% CI: 0.27 – 0.56) for partial and 0.35 (95% CI: 0.25 – 0.50) for a full course of steroids (table 9.4). Gestational age was therefore included as the first potential confounder in the multivariable analysis, followed by (in order) use of tocolysis, mode of delivery, NICU level at birth, *in utero* transfer, presence of spontaneous labour, and presentation at birth. Who the woman was counselled by, whether withholding care was discussed or the parents expressed a choice regarding resuscitation options, and the decision whether to perform Caesarean in case of fetal distress were also included as potential confounders. Additionally, there was evidence of effect modification by mode of delivery (LRT p< 0.001 in the univariate analysis for condition at birth, p=0.008 for perinatal death) – but not from any other factors – hence this was entered into the model last.

Of the potential clinical confounding factors, all except spontaneous/induced labour provided important contributions to the final model for birth in good condition, result-

ing in fully adjusted ORs of 1.42 (95% CI: 0.99 – 2.04) for a partial course and 1.64 (95% CI: 1.16 – 2.35) for a complete course of steroids. The only counselling variable not affecting the final model related to whether or not withholding care was discussed. Even after the other counselling variables were included, a complete course of steroids retained a strong effect, with a fully-adjusted OR of 1.82 (95% CI: 1.22 – 2.71). Almost the same factors were influential on the odds of perinatal death, except for delivery presentation which did not have an effect and presence of a spontaneous labour which did. Results adjusted for clinical variables show a change in the odds ratio to 0.47 (95% CI: 0.32 – 0.70) for a partial course and 0.45 (95% CI: 0.30 – 0.67) for a full course of steroids. After adjustment for counselling factors, the odds ratios were 0.47 (95% CI: 0.31 – 0.73) and 0.37 (95% CI: 0.23 – 0.58) for partial and full courses, respectively.

When taking into account the potential interaction with mode of delivery, there was no evidence of birth in an improved condition after administration of steroids antenatally for those babies born by Caesarean section, but a strong benefit to babies born vaginally (LRT p=0.018) . Those who received a partial course of steroids demonstrated an improved odds ratio of 1.61 (95% CI: 1.09 – 2.37) which increased further to 1.84 (95% CI: 1.20 – 2.82) after the inclusion of counselling variables, as shown in table 9.3. In contrast, the OR for those born after a complete course of steroids decreased slightly from 1.66 (95% CI: 1.13 – 2.44) to 1.63 (95% CI: 1.08 – 2.47). These findings were reflected for perinatal death; full results are in table 9.4,

Sensitivity analysis

A sensitivity analysis was performed in the population who received counselling. The same potentially confounding variables were included for both outcomes as were considered in the primary analysis; however, whether the mother was in labour or not did not contribute to the model with condition at birth as the outcome, and NICU level at the hospital of birth did not impact on the effect of steroids on either outcome. As per the main analysis, the biggest confounding impact on the effect of steroids was seen by gestational age, although there was no evidence of differing effects by gestational age (likelihood ratio test p=0.44 for condition at birth and p=0.46 for perinatal death). After adjustment for all clinical variables, both partial and full courses of steroids had important associations with condition at birth (1.76, 95% CI: 1.18 – 2.62, for partial and 2.18, 95% CI: 1.47 – 3.23, for full courses) and with perinatal death (0.45, 95% CI:

Table 9.4: Odds ratios of the effect of antenatal steroids on perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or delivery to perform Caesarean section.

Model	Perinatal death					
		Partial course		Full course		P-value [a]
	N	Effect	(95% CI)	Effect	(95% CI)	
Confounding only models						
Baseline	1701	0.13	(0.10 – 0.18)	0.09	(0.07 – 0.12)	—
Baseline + GA	1701	0.39	(0.27 – 0.56)	0.35	(0.25 – 0.50)	< 0.001
'Clinical' model [b]	1656	0.47	(0.32 – 0.70)	0.45	(0.30 – 0.67)	0.215
Complete model [c]	1656	0.47	(0.31 – 0.74)	0.37	0.23 – 0.58)	< 0.001
Models with effect modification by mode of delivery [b c]						
Clinical model: Vaginal delivery	1656	0.44	(0.29 – 0.67)	0.42	(0.28 – 0.63)	0.10
Clinical model: Caesarean delivery		0.72	(0.21 – 2.50)	1.63	(0.44 – 6.09)	
Complete model: vaginal delivery		0.34	(0.21 – 0.55)	0.41	(0.26 – 0.64)	0.083
Complete model: Caesarean delivery		0.93	(0.24 – 3.63)	2.06	(0.49 – 8.65)	

[a] Likelihood ratio test p-value, comparing current to next simplest model described.

[b] The clinical model for perinatal death was adjusted for: gestational age, tocolysis, mode of delivery, NICU level at birth, *in utero* transfer, and presence of a spontaneous labour.

[c] The complete models were adjusted for the same factors as the clinical models, plus: who the parents were counselled by, what choice was expressed for resuscitation, and whether a decision was made to not perform Caesarean section in the presence of fetal distress.

0.29 – 0.69, and 0.36, 95% CI: 0.23 – 0.56, for partial and full courses respectively). This was attenuated by counselling for both outcomes for a partial course of steroids, but not for a full course (full adjusted ORs 1.56, 95% CI: 1.02 – 2.37 and 0.53, 95% CI: 0.33 – 0.86 for condition at birth and perinatal death, respectively; full results in table 9.5). As for the primary analysis, there was evidence of interaction with mode of delivery, with no effect seen in those born by Caesarean section, but marked effects seen in babies born vaginally for both partial and full courses of steroids.

9.3 Tocolysis

Women who were treated with tocolytics (n=437) during the study period predominantly received atosiban (n=143, 32.7%) or nifedipine (n=189, 43.3%). Other drugs were given to 53 women (12.1%), 40 (9.2%) received multiple drugs, and 12 women (2.7%) were recorded as having tocolysis without the type of tocolytic being recorded. As the study was not powered to examine individual tocolytic drug effects, subjects

Table 9.5: Odds ratios of the effect of partial or full course of steroids on the presence of a heart rate greater than 100 at minutes of age or to perinatal death in the population of women who received counselling in England in 2006.

Model	N	Partial course		Full course		P-value
		OR	(95% CI)	OR	(95% CI)	
Condition at birth						
Baseline	1484	5.43	(3.98 – 7.42)	8.32	(6.25 – 11.08)	< 0.001
Baseline + GA [a]	1484	2.05	(1.41 – 2.98)	2.65	(1.85 – 3.80)	< 0.001
Clinical [b]	1423	1.76	(1.18 – 2.62)	2.18	(1.47 – 3.23)	< 0.001
Complete [c]	1423	1.56	(1.02 – 2.37)	2.18	(1.43 – 3.33)	< 0.001
Models with effect modification by mode of delivery:						
Clinical (vaginal delivery) [b]	1423	2.11	(1.38 – 3.23)	2.23	(1.45 – 3.41)	0.023
Clinical (Caesarean delivery) [b]		0.53	(0.17 – 1.65)	1.47	(0.48 – 4.46)	
Complete (vaginal delivery) [c]	1423	1.88	(1.20 – 2.94)	2.27	(1.44 – 3.57)	0.029
Complete (Caesarean delivery) [c]		0.45	(0.14 – 1.46)	1.33	(0.42 – 4.20)	
Perinatal death						
Baseline	1484	0.12	(0.08 – 0.16	0.07	(0.06 – 0.10)	< 0.001
Baseline + GA [a]	1484	0.38	(0.25 – 0.58)	0.32	(0.22 – 0.48)	< 0.001
Clinical [b]	1423	0.45	(0.29 – 0.69)	0.36	(0.23 – 0.56)	< 0.001
Complete [c]	1423	0.53	(0.33 – 0.86)	0.35	(0.22 – 0.58)	< 0.001
Models with effect modification by mode of delivery:						
Clinical (vaginal delivery) [b]	1423	0.38	(0.24 – 0.60)	0.33	(0.20 – 0.52)	0.056
Clinical (Caesarean delivery) [b]		2.48	(0.47 – 12.97)	1.04	(0.21 – 5.16)	
Complete (vaginal delivery) [c]	1423	0.44	(0.26 – 0.73)	0.31	(0.18 – 0.52)	0.033
Complete (Caesarean delivery) [c]		4.91	(0.63 – 38.35)	2.09	(0.28 – 15.47)	

[a] GA: gestational age

[b] The clinical models were additionally adjusted for: mode of delivery, *in utero* transfer, maternal pre-eclampsia and placental abruption.

[c] The complete models were adjusted for the same factors as the clinical models, plus: who the parents were counselled by, whether a decision was made not to perform Caesarean section in cases of fetal distress, what choice was expressed for resuscitation, and whether withholding care was discussed.

were re-classified as either having received tocolysis or not. Overall, tocolysis was associated with a more frequent good outcome (unadjusted OR 2.25, 95% CI: 1.75 – 2.90).

Further analysis identified gestational age, antenatal steroids, level of NICU care available at birth, *in utero* transfer and placental abruption as potential clinical confounders of the relationships with condition at birth or of perinatal death. Of the listed factors, all except labour type had an important confounding effect between the use of tocolysis and condition at birth, reducing the odds ratios to 1.37 (95% CI: 1.01 – 1.87). After adding in variables representing who provided counselling (if any was provided), parental choice regarding neonatal resuscitation, and decision to not perform Caesarean section for fetal distress, the odds ratio was 1.45 (95% CI: 1.05 – 2.00). The inclusion of whether or not withholding care was discussed did not affect the model, nor was there any evidence of effect modification. The findings for perinatal death were very similar, with the final model including identical variables and demonstrating an effect by tocolysis of 0.48 (95% CI: 0.32 – 0.73). Results are shown in table 9.6.

Sensitivity analysis

The effect of tocolysis as a combined group was also studied in a population including only babies who received steroids. The same factors plus maternal pre-eclamptic toxaemia (PET) were considered as potential confounders; however, there was no impact from differing amounts of steroids (partial or full courses), *in utero* transfer or labour type on the effect of tocolysis on condition at birth, with the clinical model showing an odds ratio of 1.42 (95% CI: 1.04 – 1.96) and the complete model an OR of 1.47 (95% CI: 1.06 – 2.04). For perinatal death, gestational age and placental abruption were the clinical factors identified as playing an important confounding role, leading to an adjusted OR of 0.49 (95% CI: 0.32 – 0.73). Of the counselling factors, all except whether withholding care was discussed made an important contribution to the final model (OR 0.44 (95% CI: 0.28 – 0.67)); there was no evidence of interaction. Results are tabulated alongside the complete population in table 9.6.

Table 9.6: Odds ratios of the effect of tocolysis on the presence of a heart rate greater than 100 at minutes of age or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or delivery to perform Caesarean section, and the restricted population of those babies born to women who received antenatal steroids.

Model	Complete population				Steroids-only population			
	N	OR	95% CI	P-value [a]	N	OR	95% CI	P-value [a]
Condition at birth								
Baseline	1720	2.25	(1.75 – 2.90)	< 0.0001	1213	1.44	(1.08 – 1.95)	0.0138
Baseline + GA [b]	1720	1.64	(1.24 – 2.18)	< 0.0001	1213	1.53	(1.12 – 2.09)	< 0.0001
Clinical [c]	1655	1.37	(1.01 – 1.87)	< 0.0001	1205	1.42	(1.04 – 1.96)	< 0.0001
Counselling [d]	1655	1.45	(1.05 – 2.00)	< 0.0001	1205	1.47	(1.06 – 2.04)	< 0.0001
Perinatal death								
Baseline	1720	0.33	(0.24 – 0.44)	< 0.0001	1213	0.53	(0.36 – 0.76)	0.0005
Baseline + GA [b]	1720	0.46	(0.32 – 0.65)	< 0.0001	1213	0.47	(0.31 – 0.70)	< 0.0001
Clinical [c]	1655	0.58	(0.39 – 0.84)	< 0.0001	1211	0.49	(0.32 – 0.73)	< 0.0001
Counselling [d]	1655	0.48	(0.32 – 0.73)	< 0.0001	1211	0.43	(0.28 – 0.66)	< 0.0001

[a] Likelihood ratio test p-value comparing model against next simplest model.

[b] GA: gestational age

[c] The clinical models for condition at birth and perinatal death werew adjusted for: provision of antenatal steroids, NICU level at birth, *in utero* transfer and presence of placental abruption.

[d] In addition to the factors adjusted for in the clinical models, the complete models were also adjusted for who counselling was provided by, whether and what choice was expressed for resuscitation, and whether a decision was made to *not* perform Caesarean section in the presence of fetal distress.

9.4 Mode of delivery

The mode of delivery analysis comparing Caesarean section to a baseline of vaginal delivery identified gestational age as the most important associated factor. This caused a change in the OR for condition at birth from 4.21 (95% CI: 3.06 – 5.80) to 1.63 (95% CI: 1.15 – 2.33), and for perinatal death from 0.16 (95% CI: 0.11 – 0.24) to 0.67 (95% CI: 0.42 – 1.05). However, both relationships showed strong evidence of effect modification by gestational age (likelihood ratio test p=0.017 for the relationship with condition at birth and p=0.016 for that with perinatal death). The principle reason for this is that there are very few Caesarean sections performed at low gestational ages (1, 6 and 42 at 22, 23 and 24 weeks gestation, respectively), meaning there were insufficient numbers to develop statistical significance. The 22-24 week categories were therefore collapsed into a single group, following which there was evidence of an association between Caesarean delivery and birth in a good condition at $\leq$24 and 25 weeks (odds ratios 3.89 (95% CI:

2.00 – 7.58) and 1.85 (95% CI: 1.02 – 3.34)), but not at 26 weeks gestation (OR 1.22, 95% CI: 0.69 – 2.18). For perinatal death, a difference was only shown for those born at 24 weeks or below (0.24, 95% CI: 0.12 – 0.50), with no association at either 25 or 26 weeks (0.79, 95% CI: 0.39 – 1.63, and 0.63, 95% CI: 0.24 – 1.70, respectively).

The other important factor to consider as an effect modifier is the presence or not of labour: only 64 women underwent Caesarean section following spontaneous onset of labour, compared with 292 who were not in labour. Including this in the model had an important effect, removing the association with condition at birth for women in spontaneous labour, but showing a strong association in women who were not in labour at both 24 weeks and below (OR 13.5, 95% CI: 3.50 – 52.08), and 25 weeks (4.13, 95% CI: 1.35 – 12.62). These findings were little changed after the addition of the remaining clinical variables (22.96, 95% CI: 5.19 – 101.52 at 24 weeks and below, and 8.24, 95% CI: 2.49 – 27.33 at 25 weeks), or of the counselling variables (odds ratios 12.68, 95% CI: 2.79 – 57.60 for 24 weeks and below, and for 25 weeks, 4.94, 95% CI: 1.44 – 16.90). There were similar findings for perinatal death, with no evidence of an effect from Caesarean delivery after spontaneous labour and marked evidence for babies born to women not in labour at 24 weeks gestation or lower (OR after adjustment for all clinical and counselling variables was 0.03 with a 95% CI: 0.01 – 0.21) and at 25 weeks (fully adjusted OR 0.13, 95% CI: 0.03 – 0.55), but no effect for those born at 26 weeks (adjusted OR 0.12, 95% CI: 0.01 – 1.63). All results comparing Caesarean section with vaginal delivery are shown in table 9.7.

Sensitivity analysis

The findings for Caesarean delivery were tested by conducting a sensitivity analysis in the steroids-only population. As in the earlier analyses, the most important confounding effect on both outcomes came from gestational age, along with strong evidence of interaction by labour type; there was, however, no effect modification by gestational age evident. For the analysis with condition at birth as the outcome, *in utero* transfer, presentation at delivery, the level of neonatal care available at the birth hospital, maternal pre-eclampsia and placental abruption were additionally identified as factors impacting the relationship, giving adjusted ORs of 8.67 (95% CI: 3.47 – 21.70) and 1.22 (95% CI: 0.71 – 2.09) for women who were not in labour and those in spontaneous labour, respectively. The ORs were further impacted by counselling; specifically, after

Table 9.7: Odds ratios of the effect of delivery by caesaerean section on the presence of a heart rate greater than 100 at minutes of age or to perinatal death in the complete population of babies born in England in 2006 who were known to be alive at admission to hospital and at the start of labour monitoring or delivery to perform Caesarean section.

Model	N	≤24 weeks		25 weeks		26 weeks		P-value [a]
		OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	
Condition at birth								
Baseline	1717	3.89	(2.00 – 7.58)	1.85	(1.02 – 3.34)	1.22	(0.69 – 2.18)	0.031
Baseline: no labour	1712	13.5	(3.50 – 52.08)	4.13	(1.35 – 12.62)	0.98	(0.11 – 8.41)	0.018
Baseline: spontaneous labour		1.79	(0.76 – 4.24)	1.95	(0.79 – 4.79)	0.75	(0.35 – 1.58)	
Clinical: no labour [b]	1680	22.96	(5.19 – 101.52)	8.24	(2.49 – 27.33)	2.22	(0.24 – 20.16)	0.001
Clinical: spontaneous labour [b]		1.18	(0.47 – 2.96)	2.01	(0.80 – 5.08)	0.71	(0.33 – 1.54)	
Complete: no labour [c]	1680	12.67	(2.79 – 57.60)	4.94	(1.44 – 16.90)	1.56	(0.16 – 14.81)	< 0.001
Complete: spontaneous labour [c]		0.95	(0.38 – 2.40)	1.58	(0.62 – 4.03)	0.70	(0.32 – 1.52)	
Perinatal death								
Baseline	1717	0.24	(0.12 – 0.50)	0.79	(0.39 – 1.63)	0.63	(0.24 – 1.70)	0.055
Baseline: no labour	1712	0.05	(0.01 – 0.23)	0.20	(0.06 – 0.67)	0.39	(0.04 – 3.76)	0.019
Baseline: spontaneous labour		0.56	(0.23 – 1.35)	0.82	(0.27 – 2.44)	0.65	(0.14 – 2.97)	
Clinical: no labour [b]	1680	0.01	(0.00 – 0.09)	0.06	(0.02 – 0.25)	0.09	(0.01 – 1.03)	< 0.001
Clinical: spontaneous labour [b]		0.96	(0.36 – 2.55)	0.77	(0.25 – 2.43)	0.63	(0.13 – 3.00)	
Complete: no labour [c]	1680	0.03	(0.01 – 0.21)	0.13	(0.03 – 0.55)	0.12	(0.01 – 1.63)	< 0.001
Complete: spontaneous labour [c]		1.36	(0.50 – 3.70)	1.07	(0.33 – 3.46)	0.63	(0.13 – 3.00)	

[a] Likelihood ratio test comparing model against next simplest model

[b] The clinical models were additionally adjusted for: antenatal steroids, *in utero* transfer, maternal pre-eclampsia and placental abruption.

[c] The complete models were adjusted for the same factors as the clinical models, plus: who the parents were counselled by, whether a decision was made not to perform Caesarean section in cases of fetal distress, what choice was expressed for resuscitation, and whether withholding care was discussed.

adjustment for who the subject was counselled by, whether a decision was made not to perform Caesarean section in the presence of fetal distress, and whether or not a choice was expressed about resuscitation, the OR for women not in labour became 5.55 (95% CI: 2.13 – 14.47) while that for women in labour remained not significant (table 9.8).

In addition to gestational age and labour type, confounders identified for the effect of Caesarean delivery on perinatal death were the length of course of antenatal steroids (partial or full), *in utero* transfer, presentation at delivery, maternal pre-eclampsia and placental abruption. Inclusion of these demonstrated reduced the odds of death for babies born to women not in labour to 0.05 (95% CI: 0.02 – 0.17), with no impact shown of Caesarean delivery for women who were in spontaneous labour (OR 0.90, 95% CI: 0.45 – 1.81). Following the inclusion of all counselling variables, there was little change, with the fully adjusted odds ratio for women not in labour of 0.10 (95% CI: 0.03 – 0.34) while the OR remained insignificant for those in spontaneous labour.

9.5 Summary of results

In the EPICure 2 cohort in 2006, 1,722 women were identified with singleton pregnancies where the fetus was alive both at admission to hospital and at the start of active monitoring of labour or the decision to perform Caesarean section. Two sub-populations were used in sensitivity analyses; the steroids-only population contained 1,213 participants, with 1,495 included in the population of women who received counselling. All three of the chosen exposures – antenatal steroids, tocolysis and Caesarean delivery – demonstrated important associations in unadjusted analyses with both outcomes: birth in a good condition and perinatal death.

The first set of analyses described in detail related to antenatal steroids. After adjustment, and including the effects of both counselling variables and an interaction seen with type of delivery, these were associated with improved odds ratio 1.84 (95% CI: 1.20 – 2.82) for birth in a good condition for those born vaginally after a partial course, but only 1.63 (95% CI: 1.08 – 2.47) after a full course. There was no effect seen in those born by Caesarean delivery. Similar results were found for perinatal death, with odds ratio of 0.34 (95% CI: 0.21 – 0.55) and 0.41 (95% CI: 0.26 – 0.64) for partial and full courses of steroids in those born vaginally after adjustment for all variables. The sensitivity analyses were conducted in the population of women who received

Table 9.8: Odds ratios of the effect of delivery by caesaerean section on the presence of a heart rate greater than 100 at minutes of age or to perinatal death in the population of babies born to women who received a partial or full course of antenatal steroids in England in 2006.

Model	N	OR	(95% CI)	P-value
Condition at birth				
Baseline	1213	2.31	(1.61 – 3.31)	<0.001
Baseline + GA [a]	1213	1.46	(0.99 – 2.15)	<0.001
Models with effect modification by labour type:				
No labour [b]	1212	4.08	(1.96 – 8.50)	0.003
Spont labour [b]		0.86	(0.52 – 1.43)	
Clinical: no labour [c]	1172	8.67	(3.47 – 21.70)	0.0001
Clinical: spont labour [c]		1.22	(0.71 – 2.09)	
Complete: no labour [d]	1172	5.55	(2.13 – 14.47)	0.016
Complete: spont labour [d]		1.03	(0.60 – 1.79)	
Perinatal death				
Baseline	1213	0.38	(0.24 – 0.61)	<0.001
Baseline + GA [a]	1213	0.84	(0.51 – 1.39)	<0.001
Models with effect modification by labour type:				
No labour [b]	1212	0.19	(0.08 – 0.46)	0.0002
Spont labour [b]		1.47	(0.76 – 2.82)	
Clinical: no labour [e]	1178	0.05	(0.02 – 0.17)	<0.001
Clinical: spont labour [e]		0.90	(0.45 – 1.81)	
Complete: no labour [f]	1210	0.10	(0.03 – 0.34)	0.033
Complete: spont labour [f]		1.18	(0.58 – 2.43)	

[a] GA: gestational age.

[b] Adjusted for gestational age.

[c] Adjusted for: GA, labour type,*in utero* transfer, presentation at delivery, NICU level at birth hospital, maternal pre-eclampsia and placental abruption.

[d] Adjusted for the same factors as the clinical model ('c'), plus: who the parents were counselled by, whether a decision was made not to perform Caesarean section in cases of fetal distress, and what choice was expressed for resuscitation.

[e] Adjusted for: GA, labour type, antenatal steroids, presentation at delivery, maternal pre-eclampsia and placental abruption.

[f] Adjusted for the same factors as the clinical model ('e'), plus: who the parents were counselled by, whether a decision was made not to perform Caesarean section in cases of fetal distress, what choice was expressed for resuscitation, and whether withholding care was discussed.

counselling and found similar results with respect to the confounding variables and effect modification by mode of delivery; however, the odds ratios were more consistent in this population with a dose-response effect from steroids. For condition at birth, the OR for a partial course of steroids was 1.88 (95% CI: 1.20 – 2.94) and for a full course, 2.27 (95% CI: 1.44 – 3.57); for perinatal death, the respective ORs were 0.44 (95% CI: 0.26 – 0.73) and 0.31 (95% CI: 0.18 – 0.52).

The second set of analyses related to use of tocolysis, with the sensitivity analyses carried out using the population of women who received antenatal steroids. Adjusted results between the two populations were relatively consistent, with the fully-adjusted models in the complete population finding odds ratios of 1.45 (95% CI: 1.05 – 2.00) for birth in a good condition and 0.48 (95% CI: 0.32 – 0.73) for perinatal death in those receiving tocolysis.

Mode of delivery – specifically, delivery by Caesarean section – was the focus of the last group of analyses. Two key interaction terms were identified in the main analysis: gestational age, and whether the woman was spontaneously labouring or not. In the absence of labour, there were extremely strong odds of being born in a good condition for those born at 25 (OR 4.94, 95% CI: 1.44 – 16.90) or $\leq$ 24 weeks gestational age (12.67, 95% CI: 2.79 – 57.60), but no difference at 26 weeks (1.56, 95% CI: 0.16 – 14.81). Similarly, the odds of perinatal death were 0.12 (95% CI: 0.01 – 1.63) at 26 weeks, 0.13 (95% CI: 0.03 – 0.55) at 25 weeks and 0.03 (95% CI: 0.01 – 0.21) at $\leq$ 24 weeks. The sensitivity analyses used the population who received steroids; spontaneous labour continued to modify the effect but in this population gestational age acted only as a confounder. This meant the there was an odds ratio of 5.5 (95% CI: 2.13 – 14.47) for birth in a good condition and 0.10 (95% CI: 0.03 – 0.34) for perinatal death in the absence of labour, but no difference if the mother was labouring.

Chapter 10

Discussion of the obstetric antecedents of prematurity investigation

This chapter presents a discussion of the findings from the investigation of the obstetric antecedents of extreme prematurity – specifically, the effects of administration of antenatal steroids, the use of tocolysis, and Caesarean delivery on the presence of a heart rate greater than 100 beats per minute at five minutes of age ("good condition"), and on "perinatal death", defined for this scenario as death during labour or in the delivery room. Key results are presented first, followed by discussion of the limitations of the study.

10.1 Key results

The full study cohort included 1,722 women with singleton pregnancies where the fetus was alive at admission to hospital and at the the beginning of labour monitoring or the point at which it was decided to perform Caesarean section; 1,213 women received antenatal steroids and 437 tocolysis, of whom 406 received both. Some form of antenatal counselling was provided to 1,495 women, and 356 babies were delivered surgically.

Even after accounting for potential decision-making influences, there was strong evidence that antenatal steroids were associated with improved survival at birth as well as being born in better condition, particularly in those babies born vaginally.

Fully adjusted odds ratios (ORs) for birth in a good condition were 1.84 (95% CI: 1.20 – 2.82) and 1.63 (95% CI: 1.08 – 2.47) for partial and full courses of steroids, respectively, in those born vaginally. For babies delivered by Caesarean section, no effects were seen (table 9.3). These findings were similar to those for perinatal death: for a partial course of steroids in those born vaginally, the OR was 0.34 (95% CI: 0.21 – 0.55) and, for babies born after a full course of steroids, 0.41 (95% CI: 0.26 – 0.64; full results in table 9.4). In the population of women who were recorded as receiving antenatal counselling, effects were more marked in those who received a complete course of steroids (table 9.5).

This study also showed an association between the use of tocolysis and improvements in both outcomes, shown in table 9.6. There was an increase in the adjusted odds ratio for birth in a good condition to approximately 1.4 that was highly statistically significant ($p< 0.0001$) in all of the final analyses. For perinatal death, the odds ratio was approximately 0.5 for those who received tocolysis – slightly lower in the restricted population of only women who received steroids. There was no evidence of effect modification by gestational age or any other factor.

In contrast, mode of delivery was subject to effect modification by gestational age *and* by the presence or not of spontaneous labour. Consequently, there was no evidence of an effect by Caesarean section at any gestation in women with spontaneous labour, but in women without spontaneous labour there were bigger effects at lower gestations. This finding was even more pronounced for the outcome of perinatal death than for condition at birth. Babies born by Caesarean section to women who were not in spontaneous labour at 24 weeks gestation or below had 12.67 (95% CI: 2.79 – 57.60) times the odds of being born in a good condition than those who were born vaginally; there was no difference at 26 weeks (OR 1.56, 95% CI: 0.16 – 14.81). With respect to perinatal death, the OR at $\leq$ 24 weeks was 0.03 (95% CI: 0.01 – 0.21) whereas at 26 weeks there was no evidence of an association, with an OR of 0.12 (95% CI: 0.01 – 1.63).

10.2 Strengths and limitations

There are a number of potential limitations to this study. As previously, these may be classified according to the type of error that is introduced: random (chance), non-

random (bias) and confounding.

10.2.1 Chance

The study benefits from the size and the completeness of the data that were available: it covers the entire population of England, and data collection was comprehensive.[55] It was therefore possible to obtain robust statistical evidence in support of results. This provides a high degree of certainty – evidenced by extremely low p-values in many of the analyses – about how the usage of antenatal steroids, tocolysis and operative delivery affected immediate outcomes for these extremely premature babies in 2006 in England. However, many of the regression analyses were complex, particularly when considering interaction. This sometimes lead to individual cells containing small numbers, thus resulting in an increased possibility of random error being introduced. This is reflected by wide confidence intervals in results; these are particularly evident in the mode of delivery analyses.

A related issue was the poor distribution of data within some variables. As there are very few Caesarean sections performed at the lowest gestational ages, those ≤ 24 weeks had to be grouped into one category. Similarly, there were low numbers of women receiving tocolysis, and an inadequate spread of data among those who did (most women received atosiban or nifedipine, as shown in table 9.1), hence it was not possible to examine the effects of individual medications. For the analyses that were performed, however, it seems unlikely that chance had an important impact on the results.

10.2.2 Bias

The study may be biased. Some data *were* missing, which may have lead to a non-reporting, selection bias. It is most likely that this would have reduced the effect of a variable, possibly even to nullify it. This occurred for any effect from gestational age at booking: no difference was seen from booking late in pregnancy; however, it is equally plausible that there was no effect anyway. Further reassurance is provided by the relative size of final adjusted models to the numbers at baseline: in all but one (the analysis comparing steroids to condition at birth) were there greater than 95% of the subjects included in the final model.

More importantly, the counselling variables were incompletely filled out. Analyses were therefore based on whether there was evidence that counselling was carried out or not. This may have led to the confounding impact of these variables being underestimated if there was under-recording of positive responses; however, it is more likely that positive responses will have been recorded, and that in those subjects where a response was missing the relevant conversation is less likely to have occurred. The incomplete counselling variables may also have affected the sensitivity analyses for the effect of antenatal steroids. In actual fact, the point estimates obtained from the sensitivity analyses are *more* biologically plausible than the main analyses as a dose-response relationship is seen in all models, including those with effect modification. Reassuringly, however, effect magnitudes and confidence intervals are similar across both sets of analyses (using the main and counselled populations).

A further selection bias that may have occurred relates to women who receive interventions but then did not deliver before 27 weeks gestational age. Some women who presented with threatened preterm labour at 20 to 26 weeks completed gestation may have received antenatal steroids, tocolytic therapy and/or some other treatment, but subsequently have been well enough to be discharged, hence continuing her pregnancy beyond the gestational age ranges included in the study. There is no way of knowing how many women this applies to, nor the effect that might have been seen if they were included. For tocolysis particularly, delaying delivery may be a positive outcome – if the fetus stays healthy. If this were the scenario, exclusion of these women may bias results towards no effect, yet an effect was seen in the investigation of tocolysis.

Another form of selection bias that may occur is also known as "confounding by indication",[220] and occurs when selection into one of the exposure groups is conditional upon a factor related to the outcome. For instance, antenatal steroids may be more likely to be prescribed if it is thought the fetus will have a good outcome. This study attempted to control for some of these factors through the inclusion of variables related to antenatal counselling. However, it is unlikely to have fully accounted for these influences. This was evident in the antenatal steroids analysis, where a partial course of steroids had a greater impact than a full course of steroids on condition at birth in those born vaginally. A potential solution to rectify this problem is through the use of methods such as propensity scores[221, 222], or instrument variable analysis if suitable instruments can be identified.

Information bias also may have influenced the results. Recall bias was minimised as data were collected contemporaneously, but there may have been differential reporting at the various study centres leading to differential misclassification. Clear instructions were provided with the initial data collection form in order to prevent this, and the data are internally consistent.[55] Thus, this is likely to have had little if any impact on the results. Similarly, loss to follow-up bias is unlikely to have played a role due to the involvement of principle investigators in ensuring data returns; for this study, outcome data were complete.

Within the statistical analyses, a conservative attitude was taken in interpreting ambiguous data; this meant that interpretation was restricted on occasion. For example, chorioamnionitis is a diverse condition that is poorly diagnosed on both clinical and histological grounds. [223] Because of the wording of the questions that had been used, the analysis was only able to look at the combined group of "clinical suspicion of chorioamnionitis at any time before birth" and those in whom "chorioamnionitis [was] noted at time of birth", rather than more precise patient categories. Any misclassification that resulted from this would have happened in a non-differential fashion, and therefore be likely to bias results towards the null.

10.2.3 Confounding

It might be argued that the choices of confounding variables in the analyses are wrong; particularly, in the steroids analyses, the inclusion of mode of delivery as a covariate may appear bizarre. This is because it appears to be on the causal pathway, as shown when considering figures 9.3 and 9.5 together. However, mode of delivery may instead be considered as a marker for *intention* to deliver – for, having made a decision to deliver operatively (e.g. in the presence of intrauterine growth restriction), one would first ensure that steroids had been administered; in this scenario, the model makes sense. It might also be argued that variables not included in the final adjusted models should have been. This, however, is not an issue as the study was designed to only retain those variables that had an important confounding impact on the relationship between exposure and outcome; variables not retained did *not* exhibit a confounding influence.

The results of the study may, however, be impacted by residual confounding. More detail could have been obtained about the current variables: for instance, there is very

little detail regarding tocolysis – how long women were administered drugs for, or at which doses. However, the most important variables are believed to have been captured: this is demonstrated through confounding by gestational age and the use of antenatal steroids, which had consistently large impacts throughout all analyses. Furthermore, a longer study period or a larger baseline population would be required to power more detailed investigations.

More specifically, residual confounding may have occurred in relation to the counselling information. The low completion rates for these variables in comparison to other questions, together with the patterns of overlap in the responses, suggests there was some confusion in answering these questions. This raises the possibility that the true confounding effects of perinatal decision-making at the borders of viability have not been accounted for. However, when analyses excluding counselling variables ("clinical" models in tables 9.3 – 9.8) are compared with those that include them (the "complete" models), there tended to be a *more* pronounced effect in the complete (i.e. including counselling factors) models.

10.3 Interpretation

It is reassuring that the effects seen in the results are in the direction and of a magnitude that might be expected for each of the exposures. The study further benefits from the fact that it is a complete birth cohort and not based solely on neonatal admissions data – which, by definition, would not be able to answer questions about delivery room outcomes.

One point to highlight is that perinatal death was redefined in the analyses to include deaths *only* during labour or occurring shortly thereafter in the delivery room. This is of interest to obstetricians and midwives as it is a more direct reflection on antenatal and intrapartum care. However, it may mean that results are less comparable with other studies – studies that may have used a more conventional definition such as including neonatal deaths occurring within the first seven days of life, as in the calculation of the traditional perinatal mortality rate.[217] Causality is another important issue: the evidence provided by this study relates only to *associations* that are detectable in pregnancies that deliver at these extremely low gestational ages. Are the associations *causative*? This cannot be stated from the present study alone, as

"consistency" between multiple studies is one of the nine Bradford Hill criteria that are required to be met.[224] However, a causal interpretation may certainly be suggested by the results.

10.3.1 Caesarean section

There were very marked improvements in survival and condition at birth in babies born via operative delivery at gestations below 26 weeks to women who were not in labour, but no difference at 26 weeks. In contrast, no benefit was seen from performing Caesarean sections at any gestation for women who were already in labour.

The primary interpretation of this is that there is an important influence from case selection: very few Caesarean sections were performed at the lowest gestations (1, 6 and 42 at 22, 23 and 24 weeks respectively) where the odds of a favourable outcome were the highest. These findings are compatible with a Cochrane Review investigating the effect of Caesarean delivery in premature singleton babies that found no evidence of a difference in mortality.[225] However, the review identified only six studies, three of which examined mortality, and only two of these including births as low as 26 weeks gestational age.[225] The EPICure results are also compatible with another observational study, using linked registry data in the United States.[226] This used a "trimming procedure" to ensure that only births with congruent birth weight and gestational age data were included. Neonatal mortality – primarily on the day of birth – was shown to be lower in babies born operatively at 25 weeks and below.[226] Differences became more pronounced with decreasing gestational age; although the size of effect seen was lower than in the present study, perhaps due to non-differential misclassification which would bias towards the null; this is an important similarity.

An alternative explanation for the findings is that vaginal delivery is more dangerous than Caesarean delivery at the very lowest gestations – at least, in terms of the baby's outcome. This is plausible: there is likely to be increased trauma to the baby during vaginal delivery, and circumstances may be less controlled. However, it is a less plausible explanation than the primary interpretation, above, despite the fact that it has been postulated previously.[226] It also lacks coherence with findings from the other analyses in this study, particularly steroid analyses, which demonstrated important effect modification by mode of delivery.

10.3.2 Tocolysis

Tocolysis, on the other hand, was associated with an improved outcome in all groups, even after accounting for antenatal steroids, gestational age, perinatal counselling and other factors. This is important knowledge, as it helps answer the question relating to women with threatened preterm labour at these gestations, "is tocolysis beneficial?" This question arises because of the potential negative impact of a "hostile" *in utero* environment.[216] The present study demonstrated an improved condition at birth and increased survival beyond the delivery room for babies born to women receiving tocolysis. These data therefore suggest that therapy was appropriately targeted within the population of women who presented during 2006 with threatened extremely preterm labour by English obstetrical practice.

How does this compare with other studies? There are a plethora of studies and systematic reviews relating to specific questions about tocolysis – for instance, a recent network meta-analysis identified 95 randomised controlled trials comparing tocolytics with either each other or placebo.[227] This found no tocolytics that were associated with improved neonatal mortality or decreased morbidity. However, there are no data presented on the effect of gestational age in any of the analyses [227] and, in fact, there are very few data available anyway about the impact of tocolysis on neonatal mortality at extremely low gestational ages.[228, 229] It is therefore unfortunate that in the present obstetric antecedents investigation there were insufficient data to look at the effects of individual tocolytics (most women received either atosiban or nifedipine, as shown in table 9.1), or to more specifically examine the circumstances in which they were given.

10.3.3 Antenatal steroids

Perhaps most notably, in independent analysis, antenatal steroids were consistently associated with improved outcomes; they were also among the strongest confounders for the other exposures. This is worth highlighting as, while the results are similar to other studies that have looked at longer term survival [34, 38, 230, 231] or shorter term outcomes in the population of babies admitted into neonatal intensive care,[231] there is little biological reason apparent for steroids to cause an improvement in condition at birth or perinatal survival as it was defined in this study. However, if these findings

are true, there *is* a biological effect from the administration of antenatal steroids that improves survival – an effect that is particularly marked in those born by vaginal delivery.

An alternative possibility is that steroids may simply be a marker of other reasons for improved survival. This warrants further investigation. Prior intention (by both parents and clinical teams) *was* adjusted for in these analyses through the use of data related to perinatal counselling. Despite the questions being crude, it appears that they captured an important aspect of the management as all but one variable (whether withholding care was discussed or not) made extremely important contributions towards the final models in all analyses. Thus, a future study could investigate the effects of antenatal steroids further by including more – and more detailed – questions specifically related to the intention to resuscitate and provision of further care, but difficulty remains in ensuring adequate response rates. Instead, it could be investigated through the use of propensity scores or similar methods, as discussed in section 10.2.2.

10.4 Generalisability

Can the results from these investigations be applied to other populations? Indeed, do findings from the English population of 2006 apply to the English population in the present time? These investigations accounted for numerous factors, including demographic factors such as maternal age, socioeconomic status and ethnicity, and clinical factors such as maternal past medical and obstetric history. The majority of these factors had no impact on the results, indicating that the effects seen are likely to truly reflect those of the exposures. They are also commensurate with findings in studies that investigated antenatal steroids [34, 38, 230, 231] and tocolysis [228, 229] in less premature babies or different geographic regions. Consequently, the results *are* likely to apply to other populations, especially those in developed nations exposed to a similar style of medical care. Furthermore, attitudes in England towards the care of women with threatened extremely preterm labour have converged since 2006, with the publication of a Scientific Impact Paper and accompanying guidelines by the RCOG a recent example of this.[228] This standardisation is likely to mean that more women receive antenatal steroids and tocolysis at the appropriate time, hence the interpretations will be applicable.

10.5 Conclusions

This chapter discussed the results obtained from the investigation into three specific obstetric antecedents of prematurity. The effect of antenatal steroids, tocolysis and mode of delivery on the chances of being born in a good condition and of perinatal death for babies delivered at 26 completed weeks gestational age or below in an English birth cohort were summarised first. The strengths and limitations were then discussed in section 10.2, focusing again on issues related to chance, bias and confounding. The size and the completeness of the data set were noted, and the fact that some bias may have occurred in relation to the counselling data which had lower rates of completion than the rest of the data. The chapter continued with possible interpretations being discussed for each of the three exposures. It was suggested that the results for Caesarean section were influenced by case selection, use of tocolysis in the current fashion is appropriate, and that steroids may exert a biological effect. The broader applicability of these findings were then discussed.

In conclusion, the results of this study demonstrate the importance of antenatal steroids in babies born vaginally, and provide contemporary estimates of the impact of tocolysis and Caesarean section that may be used when counselling patients and their families.

Part IV

Trends in unit organisation over time

Chapter 11

Methods for the assessment of trends over time in hospital organisation

The effect of changes in provision of neonatal services over time in England was investigated using data from the UK Neonatal Staffing Study (UKNSS) and the Unit Profile Studies (UPS) of 2006 and 2011. This chapter describes the methods used in the study. It commences with a brief recap of the data sets that are available, followed by a description of the variables. The chapter then moves on to detail the exposure and the outcomes that are considered; the latter part of this contains detail on how the outcome measures were derived. A discussion about the populations that the study covers is then presented, following which I describe the statistical methods and sensitivity analyses that were performed, before ending with a brief summary.

11.1 Data sources

Data sets from three studies were used in this analysis. The first was the UKNSS, conducted by the UKNSS Collaborative Group in 1997 (described in section 4.4); the second, the UPS conducted by the EPICure Research Group in 2006 (section 4.3); and the third, a repeat UPS carried out in 2011 by the NESCOP collaboration, for which I led the development and revision of the questionnaire, as well as the actual data collection (section 4.5). Among the three, there were a number of common questions

and available variables representing unit activity, staffing and infrastructure; these may be used to investigate changes that have occurred over the time. Specifically, the investigation sought to answer questions on whether nurse staffing and neonatal unit activity had changed with national changes in the configuration of neonatal care.

Hospitals for inclusion were defined as those hospitals that provided a maternity facility concurrent with neonatal services. This excluded both stand-alone, midwifery-led birthing centres as well as neonatal facilities located within an exclusively paediatric hospital.

11.2 Data analysis

Common activity data collected by the three surveys covering the most recent complete year for which data were available were: the total number of deliveries (defined as the number of women giving birth rather than the number of babies born) in the hospital; the total number of admissions (including re-admissions) to the neonatal unit and the number of admissions of babies with a birth weight of less than 1500g; the number of babies provided with any respiratory support (excluding babies who only received supplementary nasal cannula or head box oxygen but including those who received CPAP or high flow oxygen), the number of those who only received invasive ventilatory support, and the total number of days of support in each category (any support or invasive support only).

Nurse staffing data collected related to the establishment number of nurses currently employed in whole-time equivalent (WTE) for all nurses and for those with a specialist qualification. In relation to medical staffing, the total number of consultants contributing to the on-call rota for neonates and the number of consultants with 50% or more of their clinical time dedicated to neonatology, as well as the number of consultant-led ward rounds per week were collected.

Additionally, infrastructural data about the units were collected on the number of cots available, broken down by level of care (intensive, high dependency or special care), whether the unit provided ongoing neonatal intensive care or temporary care only and, for the latter surveys, the designated level of the unit (level 1, 2 or 3 in 2006, equating to special care unit (SCU), local neonatal unit (LNU) or network neonatal intensive care unit (NICU) in 2011). In addition, the local strategic health authority (defined as

the strategic health authority (SHA)o regions in existence in January 2006) for each unit was noted, as were the networks in existence in the later surveys.

11.2.1 Exposure

The time period at which each survey was carried out was considered the exposure for this study. Data were therefore available from before (1996-97) and after (2010-11) the introduction of managed clinical networks, as well as shortly after their introduction (2005-6) in what is likely to have still been a transition phase.

11.2.2 Outcomes

The focus of the study was on assessing adequacy of nurse staffing and changes in activity over time. This immediately raises two issues. First, as both of these outcomes may be assessed or measured in different ways, it was critical to define in advance which measures would actually be used. Secondly, some form of standardisation is required, to ensure that the chosen measures are compared in the same way at each time point.

Of the two outcomes, the most straight-forward was nurse staffing: this was available as the total number of nurses in each unit in WTE, as well as the WTE number of QIS nurses. These figures are also known as the "establishment number" of nurses. Due to the potential changes in size and configuration of individual neonatal units during the study period, it was decided to standardise these results by expressing both in relation to the number of cots.

The number of cots that was used was based on the British Association of Perinatal Medicine (BAPM) guidelines, which state that an individual nurse should not be responsible for more than one baby in an intensive care cot, two babies in high dependency cots or four babies in special care cots.[18, 19, 21] Such guidance has been re-iterated in publications by the Department of Health.[20, 65] I therefore used these ratios to create the standardised number of "BAPM" cots (N_{BAPM}), as per the following formula:

$$N_{\mathrm{BAPM}} = N_{\mathrm{IC}} + \frac{N_{\mathrm{HD}}}{2} + \frac{N_{\mathrm{SC}}}{4} \tag{11.1}$$

where, N_{IC} = number of intensive care cots; N_{HD} = number of high dependency cots; and, N_{SC} = numbers of special care cots. Using this, the number of nurses required per

unit to meet current guidelines for staffing were estimated. Current guidelines recommend that the total establishment number of nurses in order to provide continuous 24 hour nursing cover should be equivalent to the BAPM cot number plus 1 (representing an additional,'float' nurse), multiplied by 5.75 WTE to take into account annual, study and maternity and other leave.[67, 232] This was a slight increase from the previously recommended multiplier of 5.5 [26, 233], which was instead chosen for use in this study in order to preserve consistency with previous studies, and because it allowed more conservative estimates of change to be made, thereby potentially increasing the validity of results. The resulting equation used in this study for predicted number of all nurses in (WTE) (denoted as P_{all}) was:

$$P_{\text{all}} = 5.5 \times (1 + N_{\text{IC}} + \frac{N_{\text{HD}}}{2} + \frac{N_{\text{SC}}}{4}) \tag{11.2}$$

This facilitated calculation of the ratio of actual (denoted as A) to predicted number of nurses per unit at each time period, here called S_{nurse}, as follows:

$$S_{\text{nurse}} = \frac{A_{\text{nurse}}}{P_{\text{all}}} \tag{11.3}$$

However, as only the number of QIS nurses have previously been shown to be associated with mortality,[29] a further variable, S_{qis} was constructed,, using the same denominator of predicted number of (all) nurses (P_{all}) but with a numerator consisting of the number of actual QIS nurses (denoted as A_{qis}):

$$S_{\text{qis}} = \frac{A_{\text{qis}}}{P_{\text{all}}} \tag{11.4}$$

Unlike staffing, "activity" is not so clearly defined as a concept, and the term may represent different things to different people. One specific form of activity is "throughput", also known as "queuing time", defined using Little's law:

$$L = \lambda W \tag{11.5}$$

where, within a pre-specified system, L represents the number of units, and W the time per unit required to pass through the system, and λ is inversely proportional to the mean time between units – i.e. relates to the frequency (rate) at which units arrive.[179]

A different way of considering activity is through work *intensity*. This may be defined as how strenuous work activity is per unit measure – with units potentially representing time or, in the health research field, patients.[171]

In the present situation, these definitions give rise to a number of possible outcome measures. Either the number of admissions, the number of babies provided with respiratory support or the number of days of respiratory support provided – all measured per year – may be used and all are, in fact, measures of throughput. However, using the crude data returns from the units without any adjustment would potentially lead to serious confounding by unit size.

It was therefore decided to perform standardisation by using these measures as numerators to create ratios. For the denominator, there was a choice between the number of cots or the number of admissions per year – using the number of births made no sense as it reflects activity of the whole perinatal department rather than of the neonatal unit specifically. By using the number of cots, a time element is retained in the formula, hence retaining a measurement of throughput: the "unit" simply becomes admissions (or babies or days) per cot. If, instead, the number of babies admitted per year is chosen as the denominator, the time elements in the equations cancel out, thus intensity is measured.

Of these options, it was decided to use the number of babies admitted and the number of days of respiratory support provided per year to measure throughput (denoted here as T). These were each standardised using the adjusted number of cots shown in equation 11.1 to produce:

$$T_{\mathrm{a}} = \frac{N_{\text{admissions/year}}}{N_{\mathrm{BAPM}}} \tag{11.6}$$

for the number of admissions per cot per year ("admissions throughput" – T_{a}), and:

$$T_{\mathrm{r}} = \frac{N_{\text{days respiratory support provided/year}}}{N_{\mathrm{BAPM}}} \tag{11.7}$$

for the number of days respiratory support provided per cot per year (T_{r}).

For intensity, it clearly was not possible to use the number of admissions as a numerator, hence there were only two possible measures, both of which were used. The first was the number of babies provided with respiratory support per admission:

$$I_{\mathrm{b}} = \frac{N_{\text{babies receiving respiratory support/year}}}{N_{\text{admissions/year}}} \tag{11.8}$$

where I_{b} is the intensity ratio. By comparison, I_{d} denotes the number of days of respiratory support provided per admission, as follows:

$$I_{\mathrm{d}} = \frac{N_{\text{days respiratory support provided/year}}}{N_{\text{admissions/year}}} \tag{11.9}$$

As noted earlier, it is clear from the way that both equations 11.8 and 11.9 are written that the time factor cancels out.

11.3 Study populations

Two potential study populations were implicit in the data: the group of intensive care units themselves, or the underlying (human) population that is served by those services. Given that the surveys covered the whole of England, it seemed sensible to conduct analyses on the basis of the underlying population, as this makes it possible to attribute a denominator to the crude totals.

However, analysis in this fashion was limited by the fact that there were no data available for maternity units *without* neonatal intensive care facilities from 1997, whereas such data did exist for the other two time points. There were, therefore, two analyses possible. The first was to use only those hospitals at each of the three time points that were intensive care units, thus representing the availability of neonatal *intensive* care in England across the entire study period. The second option was to use all hospitals but to restrict comparison to the most recent two time periods; this would also permit investigation of the changes in staffing and activity across the whole of England, but for the broader category of *neonatal* care between 2006 and 2011.

Consequently, the primary set of analyses were performed with the group of hospitals that provided neonatal intensive care in all three time periods – i.e. all hospitals for whom data were available in 1997, level 2 and level 3 units in 2006, and LNUs and NICUs in 2011. All analyses were then repeated using all the available data from just 2006 and 2011; this secondary set of analyses assessed changes in the outcomes for the provision of all categories of neonatal care during the second time interval.

11.4 Statistical analysis

Available variables were examined for the medians and interquartile ranges, with outlying responses noted for further investigation. Those variables for which complete data from all hospitals were available were summated in order to estimate the overall provision for England at each time period. Crude differences over time were assessed using chi-squared test for trend.

For all subsequent investigations, potential confounders were identified *a priori* and their impact assessed on the relationship between exposure and outcome using linear regression. This was done after checking data from constructed outcome variables approximated the Gaussian distribution using visual inspection and assessment of kurtosis and skewness; where necessary, data were transformed prior to analysis. Multiple regression models were built using forward stepwise regression, including confounding variables based on the strength of their impact on the bivariate relationship between exposure and outcome as well as the degree of missingness. Strength of association for the individual relationships between exposure and outcome was tested using the Wald test, and models were compared using the F-test. Statistical significance for all results was set at $p<0.05$. Additionally, for all models a geographic variable was included to account for potential regional differences: in the primary analysis, Strategic Health Authority was used, in order to represent the potential pathways of neonatal commissioning; in the secondary analysis, neonatal network was used, so as to represent the organisational structures in place post introduction of managed care networks. The final models identified in this manner were then compared with models including all *a priori* identified variables that had a confounding effect, and the most parsimonious model chosen.

11.5 Sensitivity analyses

For the first staffing (S_{nurses}) and the admissions throughput (T_{a}) analyses, an additional sensitivity analysis was conducted using the group of hospitals that had data available at each of the three time points. However, as it was a biased population (due to selection bias), analyses using this population were not repeated for all outcomes.

11.6 Summary

In summary, this chapter described the methods used to investigate trends in neonatal staffing and unit activity using data from the UKNSS in 1997 and the two Unit Profile Studies of 2006 and 2011. Common variables identified from the surveys were grouped into staffing, activity and infrastructural variables. The exposure under consideration was "time", representing the changes in configuration to neonatal services that had occurred in England over the study period – notably, the introduction of managed clinical networks from 2004 onwards.[3]

The choice of outcome measures was described in detail. The study aimed to assess "staffing and activity", but only the former of these was clearly defined by any of the variables. There were further complications due to differing sizes among neonatal units, hence standardisation was performed using criteria from BAPM to categorise cot numbers; this is shown in equation 11.1, with the resulting staffing ratios developed in equations 11.2 – 11.4. Activity was broken down into "throughput" and "intensity" – throughput being a measure relating to time, whereas intensity is a measure of the workload per patient. Equations 11.6 – 11.9 were developed to represent these outcomes using the standardised number of cots as the denominator for throughput, and the number of admissions per year as the denominator for intensity.

The next point considered was which population the analyses related to. This is because there were fewer hospitals available from the UKNSS data set, meaning units providing temporary stabilisation of babies pending transfer elsewhere only had data available for 2006 and 2011, and only neonatal *intensive* care units had data at all three time points. Consequently, two sets of analyses are described: the primary analysis for each outcome used data representing the intensive care units from all three time points, with the secondary analysis comparing neonatal care provision between 2006 and 2011. All analyses were conducted using multiple linear regression.

Chapter 12

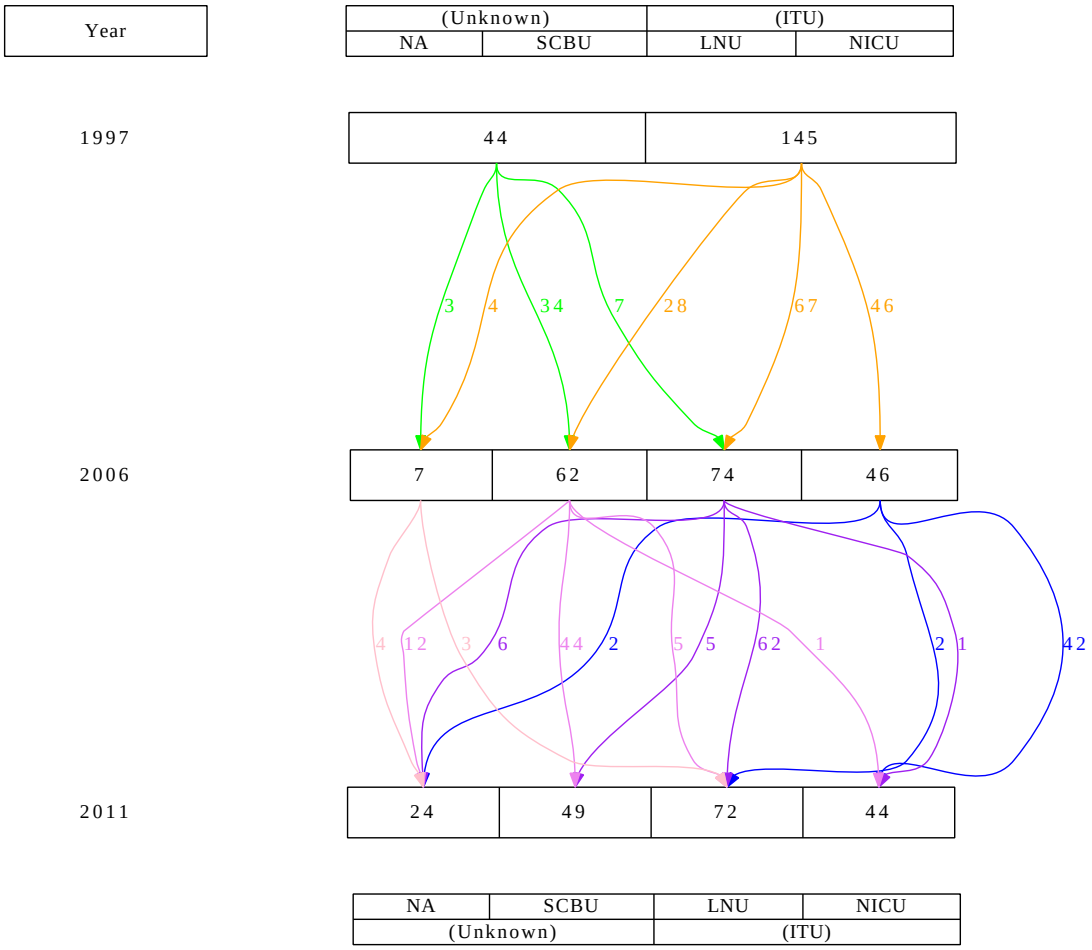
Results of the analysis of hospital trends over time

This chapter presents the results from the investigation into the organisation of neonatal units in England in 1997, 2006 and 2011 and the changes that have occurred during the intervening time periods. The first part of the chapter describes the overall changes in unit types and numbers over time. More detailed exploration of the data is then conducted, with an exploration of differences over time in summary totals and medians and IQRs. There are then three sections prior to the final summary: these deal with the main results arising from the regression analyses, used to explore in detail the changes in nurse staffing, throughput and intensity.

12.1 Unit types and numbers

One hundred and eighty-nine hospitals were identified that existed during at least one of the time points during the study period. There were 145 units providing neonatal intensive care in 1997 for which complete data were received from the original UKNSS investigators. In 2006, 182 neonatal units were identified, all of whom responded to the survey; 46 of these were Level 1 units and did not provide ongoing neonatal intensive care. In 2011, 159 of 165 units responded to the survey (96.4%); 49 Special care unit (equivalent to those designated as Level 1 in 2006) were identified, of whom six were the non-respondents to the survey (shown in figure 4.1 in chapter 4). One network neonatal intensive care unit (NICU) (Level 3 hospital) closed on 18th November 2011, hence

Figure 12.1: Types of neonatal units and numbers of each in England in 1997, 2006 and 2011.



activity data were available for the preceding year (2010) but staffing and infrastructural data were reported as null. The change in the designation of units over time is shown in figure 12.1.

12.2 Unadjusted changes over time

12.2.1 Differences by unit designation

At the time of the UKNSS, no distinction was made between different levels of provision of neonatal care, hence care provision was divided simply by whether or not intensive

care was provided on an ongoing basis. By the time of the Unit Profile Study (UPS) 2006, different levels of unit existed and, as can be seen in the graphs in figure 12.2, there were some clear differences between the types of unit. Level 3 units tended to be larger than other units with more women delivering in those hospitals (figure 12.2a), a higher number of total admissions (figure 12.2b) and of low birth weight admissions (figure 12.2c), more cots (figure 12.2f and figures 12.2g – 12.2i), greater provision of respiratory care (figures 12.2d and 12.2e, and 12.2j and 12.2k) and corresponding increases in staffing levels (with medical provision represented in figures 12.2l and 12.2m, and nurse staffing in figures 12.2n and 12.2o). Such differences were even more apparent in the 2011 data, as shown in the same set of graphs.

12.2.2 Country-wide changes

Data for perinatal centres providing neonatal intensive care at all three time points comprised the primary data set and were examined first. Response rates were high for basic activity data (number of women delivering and the number of admissions to neonatal intensive care units) and staffing and infrastructural data at all times, but poor for provision of respiratory support – particularly in the first two study epochs. Medians and IQRs for each of the variables by epoch are shown in table 12.1. The median size of units increased both in terms of the number of cots and of members of staff, and with respect to activity variables such as the number of babies admitted and the total number of days respiratory support provided.

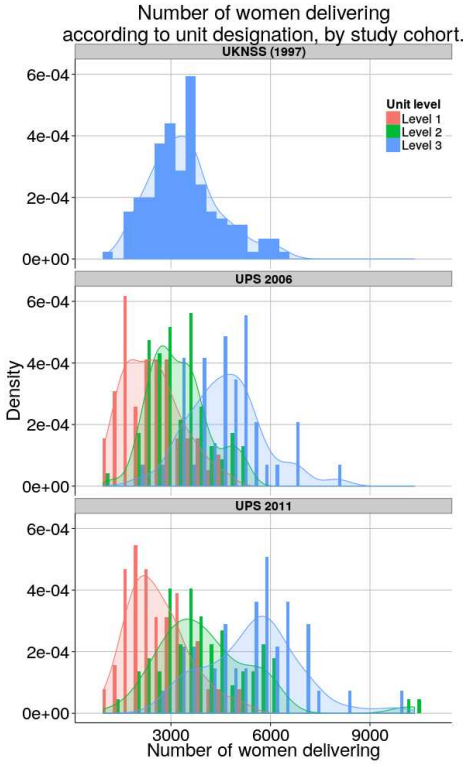
Country-wide totals were calculated for each of the available variables, to demonstrate crude national changes over time. These are shown in table 12.2. Despite increases in the number of women delivering, the number of admissions and the amount of respiratory support provided, the total number of cots available decreased from 2,725 in 145 units in 1997 to 2,683 in 116 units in 2011; both numbers of nurses and numbers of consultants increased. Similar results are shown in tables 12.3 and 12.4 for the secondary data set which covered 2006 and 2011 and additionally included level 1 neonatal units. Again, there was an increase in the median size of unit, with more cots, nurses and doctors, and increased admissions and other measures of activity except for the total days of respiratory support provided, which decreased from 137 to 126.5. There was, however, an increase in missing data in 2011 due to the six level 1

Table 12.1: Common variables from the 1997 UK Neonatal Staffing Study (UKNSS) and the Unit Profile Studies (UPS) of 2006 and 2011 for units providing neonatal intensive care in England: medians and interquartile ranges. In 1997, all hospitals for whom data were available were included; in 2006, level 2 and level 3 units were included; in 2006, these were known as local neonatal unit and network neonatal intensive care units, respectively. NA: not applicable. WTE: whole-time equivalents. PAs: clinical time.

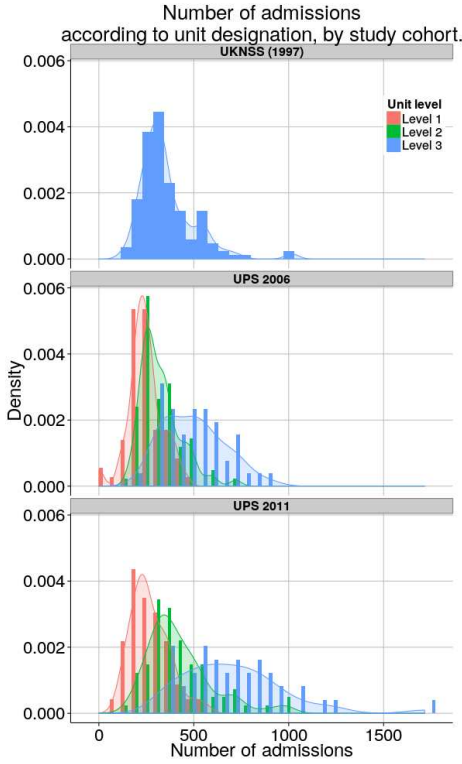
Variable	UKNSS (1997)			UPS (2006)			UPS (2011)		
	Units (n=145)	Median	(IQR)	Units (n=120)	Median	(IQR)	Units (n=116)	Median	(IQR)
Number of admissions <1500g	142	44	(32–71)	120	58	(41–98)	115	76	(54–118)
Number of admissions	145	323	(270–417)	118	350	(272–472)	114	466	(350–685)
Total babies provided with respiratory support	111	69	(42–110)	102	118.5	(71–179)	98	180	(108–281)
Total babies provided with invasive ventilatory support	27	54	(42–98)	38	64	(42–108)	112	69	(44–150)
Total days of respiratory support provided	95	455	(212–1028)	113	1027	(503–1786)	101	1304	(631–2568)
Total days of invasive ventilatory support provided	28	392	(234–817)	52	259	(137–724)	114	280	(112–913)
Number of women/deliveries	145	3250	(2608–3992)	120	3604	(2840–4603)	115	4453	(3432–5716)
Total number of ICU cots	145	4	(2–6)	120	4	(3–7)	116	4	(2–8)
Total number of HDU cots	145	4	(4–5)	120	3	(2–4)	116	4	(2–6)
Total number of SCBU cots	143	9	(7–12)	120	12	(10–14)	116	13	(10–16)
Total number of cots	145	18	(14–22)	120	20	(16–24)	116	22	(16–28)
Total number of nurses (WTE)	145	28	(20–41)	120	35	(26–52)	106	46	(32–73)
Number of neonatally-qualified nurses (WTE)	145	18	(12–26)	120	24	(16–33)	100	28	(19–45)
Number of consultants on the oncall rota for neonates	0	NA	NA	120	6	(5–7)	116	7	(6–8)
Number of consultants with >50% PAs for neonates	145	1	(0–2)	120	3	(1–4)	114	4	(1–6)
Number of consultant ward rounds	145	6	(4–7)	120	7	(5–12)	116	7	(7–12)

Figure 12.2: Density plots by unit designation and year for variables available from the UK Neonatal Staffing Study and the Unit Profile Studies of 2006 and 2011.

(a) Number of women delivering per year.



(b) Total number of admissions.



(c) Number of admissions <1500g.

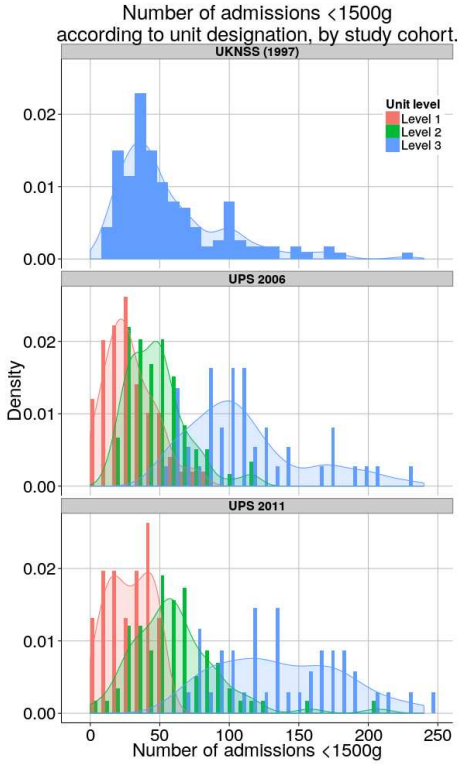
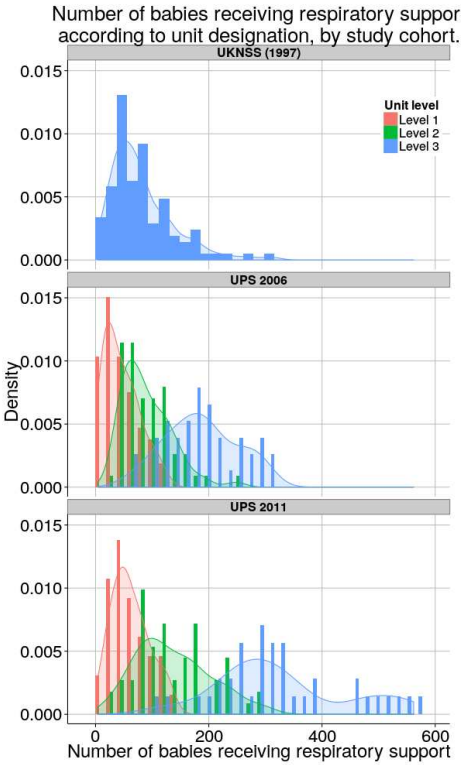
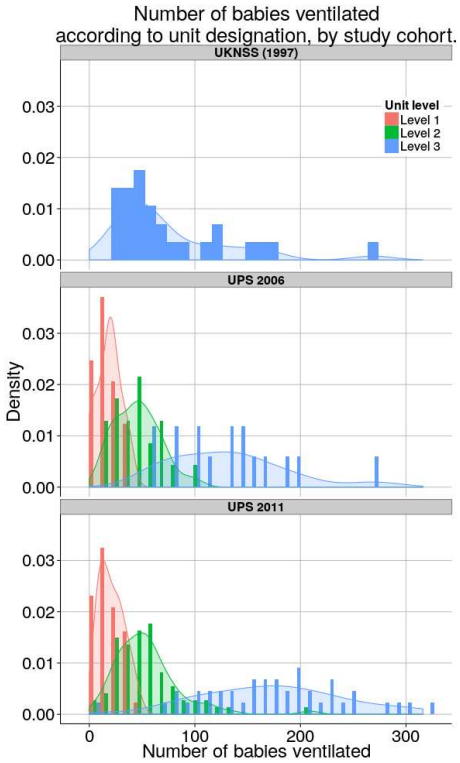


Figure 12.2: (Continued) Density plots by unit designation and year for variables available from the UK Neonatal Staffing Study and the Unit Profile Studies of 2006 and 2011.

(d) Total number of babies supported.



(e) Number of babies ventilated.



(f) Total number of cots.

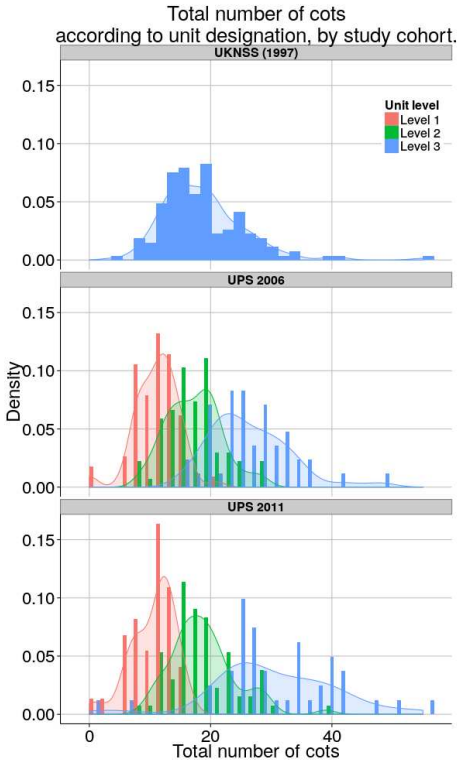
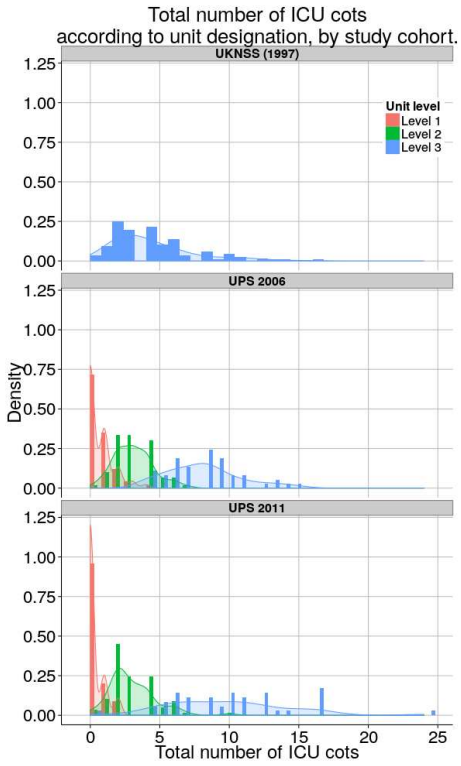
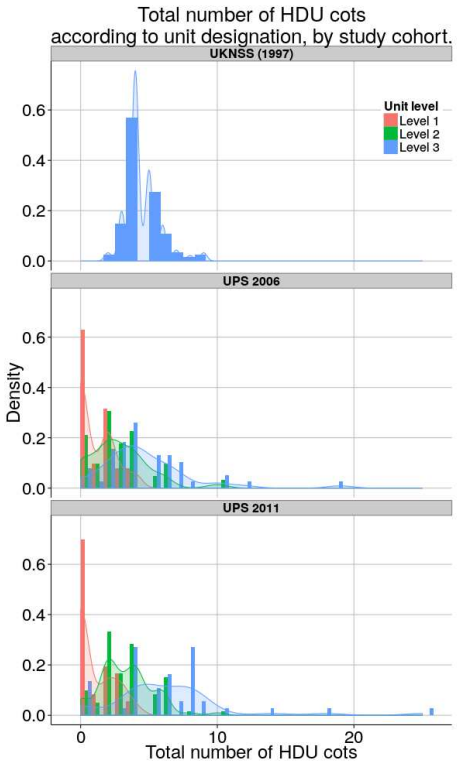


Figure 12.2: (Continued) Density plots by unit designation and year for variables available from the UK Neonatal Staffing Study and the Unit Profile Studies of 2006 and 2011.

(g) Number of ITU cots.



(h) Number of HDU cots.



(i) Number of SCBU cots.

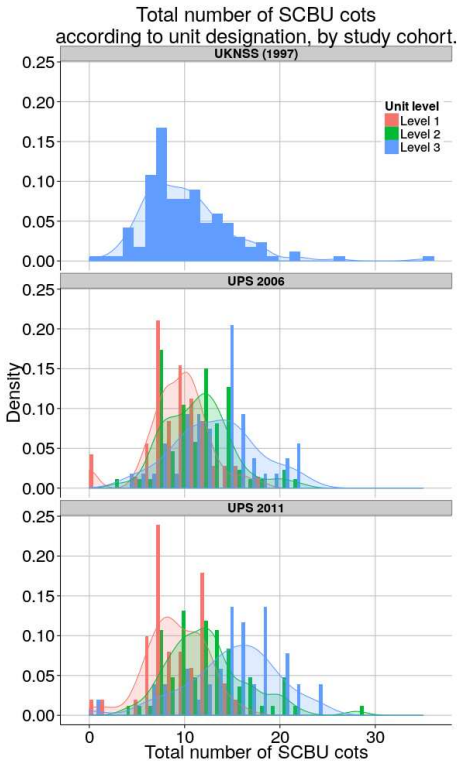
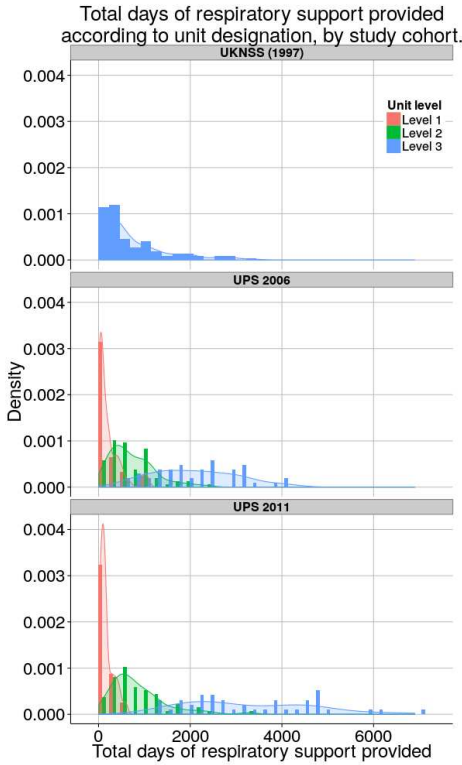
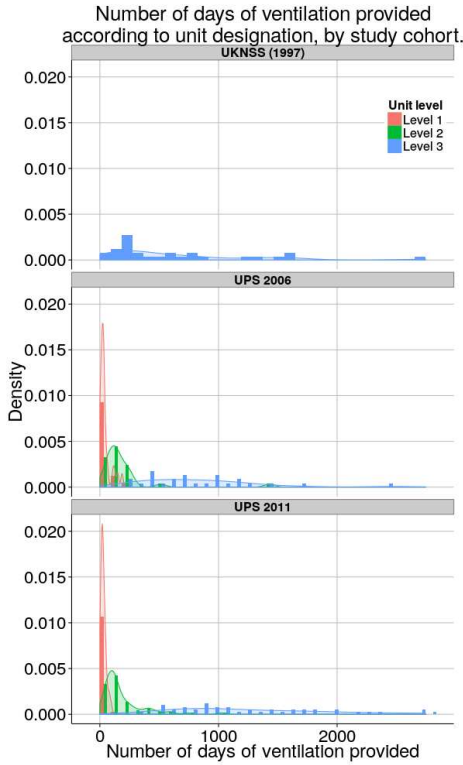


Figure 12.2: (Continued) Density plots by unit designation and year for variables available from the UK Neonatal Staffing Study and the Unit Profile Studies of 2006 and 2011.

(j) Total number of days respiratory support provided.



(k) Total number of days of invasive ventilatory support provided.



(l) Number of consultant ward rounds per week.

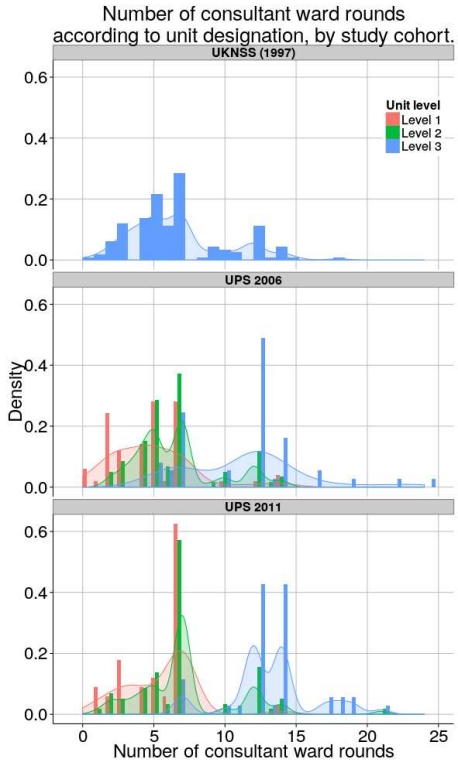
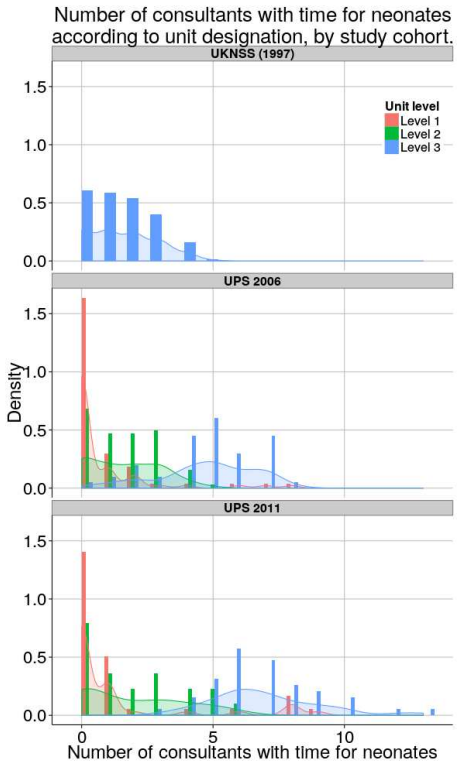


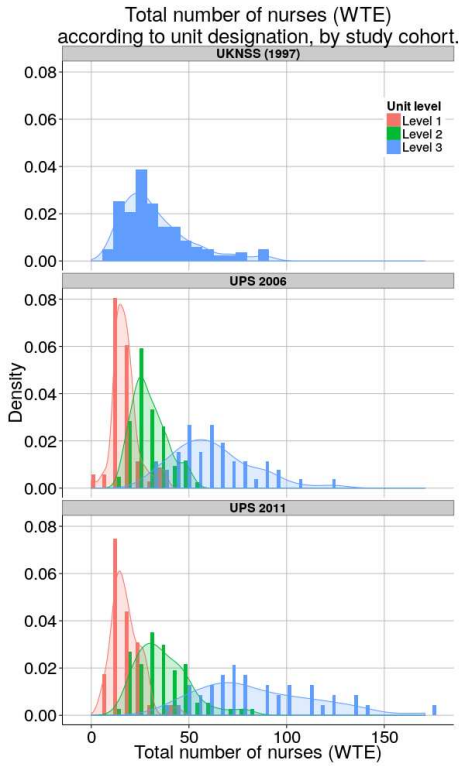
Figure 12.2: (Continued) Density plots by unit designation and year for variables available from the UK Neonatal Staffing Study and the Unit Profile Studies of 2006 and 2011.

(m) Number of consultants with >50% of PAs dedicated to neonatology.

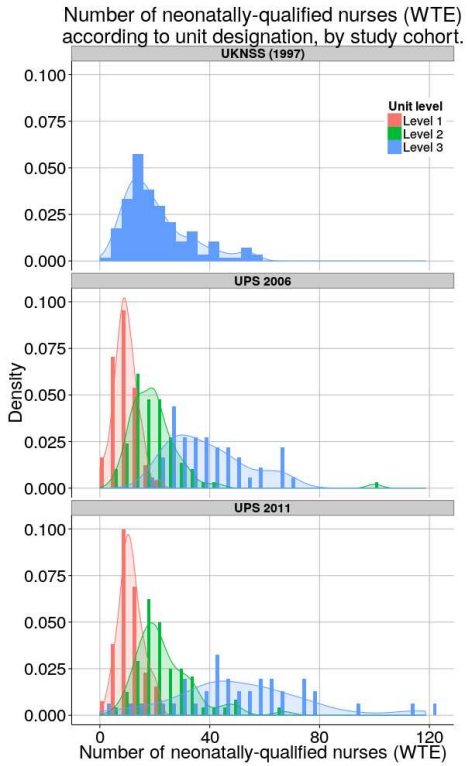


(n) Number of WTE nurses

.



(o) Number of ITU nurses who are qualified in specialty.



units that did not respond to the survey, meaning that there is less certainty regarding the total change that occurred country-wide (table 12.4). Nevertheless, it can be seen that overall activity appeared to increase, with the number of admissions (and low birth weight admissions) rising by 16.8% (9.4%) from 61,186 (10,379) to 71,477 (11,353). In contrast, the total number of cots decreased from 3,184 to 3,104, although the standardised derived variable, "BAPM cots", increased from 1,390 to 1,429.

Table 12.2: Common variables from the 1997 UK Neonatal Staffing Study (UKNSS) and the Unit Profile Studies (UPS) of 2006 and 2011 for units providing neonatal intensive care in England: total numbers. In 1997, all hospitals for whom data were available were included; in 2006, level 2 and level 3 units were included; in 2006, these were known as local neonatal unit and network neonatal intensive care units, respectively. NA: not applicable. WTE: whole-time equivalents. PAs: clinical time.

Variable	UKNSS (1997)		UPS (2006)		UPS (2011)	
	Units (n=145)	Total	Units (n=120)	Total	Units (n=116)	Total
Number of admissions <1500g	142	8102	120	8610	115	10278
Number of admissions	145	51864	118	46048	114	60690
Total babies provided with respiratory support	111	9071	102	13157	98	20524
Total babies provided with invasive ventilatory support	27	2079	38	3103	112	11042
Total days of respiratory support provided	95	69964	113	143777	101	188015
Total days of invasive ventilatory support provided	28	17891	52	26178	114	68954
Number of women/deliveries	145	493468	120	454103	115	532578
Total number of ICU cots	145	637	120	602	116	650
Total number of HDU cots	145	650	120	420	116	516
Total number of SCBU cots	143	1445	120	1455	116	1554
Total number of cots	145	2725	120	2493	116	2683
Total number of nurses (WTE)	145	4777	120	5051.94	106	5868.25
Number of neonatally-qualified nurses (WTE)	145	2978	120	3272.52	100	3449.06
Number of consultants on the oncall rota for neonates	0	0	120	753	116	854
Number of consultants with >50% PAs for neonates	145	225	120	337	114	444
Number of consultant ward rounds	145	954	120	1004	116	1106

Table 12.3: Medians and interquartile ranges for common variables from the Unit Profile Studies (UPS) of 2006 and 2011 for all units providing neonatal care in England. WTE: whole-time equivalents. PAs: clinical time

Variable	UPS (2006)			UPS (2011)		
	Units (n=182)	Median	(IQR)	Units (n=165)	Median	(IQR)
Number of admissions <1500g	182	47.5	(28–78.25)	153	57	(36–103)
Number of admissions	180	300	(235.75–408.5)	154	385.5	(297.25–575)
Total babies provided with respiratory support	159	78	(47.5–135.5)	133	128	(77–240)
Total babies provided with invasive ventilatory support	61	39	(20–75)	153	50	(26–109)
Total days of respiratory support provided	167	594	(201–1320.5)	136	772.5	(285–2195.5)
Total days of invasive ventilatory support provided	79	137	(36–460)	156	126.5	(43.5–614.75)
Number of women/deliveries	182	3173.5	(2405.75–4072)	156	3819	(2846.5–5416)
Total number of ICU cots	181	3	(1–5)	159	3	(1–6)
Total number of HDU cots	181	2	(0–4)	159	3	(1–5)
Total number of SCBU cots	181	11	(8–14)	159	12	(9–15)
Total number of cots	182	16	(12–21.75)	156	18	(12–25)
Total number of nurses (WTE)	181	27.17	(18.35–43.6)	146	35.68	(21.59–59.75)
Number of neonatally-qualified nurses (WTE)	182	16.95	(10.04–27.40)	133	20.8	(13.48–38.84)
Number of consultants on the oncall rota for neonates	182	6	(5–7)	158	7	(6–8)
Number of consultants with >50% PAs for neonates	182	1	(0–4)	155	3	(0–6)
Number of consultant ward rounds	182	7	(4.25–10)	158	7	(6–12)

Table 12.4: Total numbers for common variables from the Unit Profile Studies (UPS) of 2006 and 2011 for all units providing neonatal care in England. WTE: whole-time equivalents. PAs: clinical time

Variable)	UPS (2006)		UPS (2011)	
	Units (n=182)	Total	Units (n=165)	Total
Number of admissions <1500g	182	10379	153	11353
Number of admissions	180	61186	154	71477
Total babies provided with respiratory support	159	15757	133	22683
Total babies provided with invasive ventilatory support	61	3528	153	11902
Total days of respiratory support provided	167	154916	136	193827
Total days of invasive ventilatory support provided	79	27301	156	70176
Number of women/deliveries	182	604318	156	639689
Total number of ICU cots	181	641	159	663
Total number of HDU cots	181	485	159	559
Total number of SCBU cots	181	2027	159	1945
Total number of cots	182	3184	156	3104
Total number of nurses (WTE)	181	6114	146	6581.4
Number of neonatally-qualified nurses (WTE)	182	4732.86	133	3821.38
Number of consultants on the oncall rota for neonates	182	1061	158	1140.5
Number of consultants with >50% PAs for neonates	182	383	155	498
Number of consultant ward rounds	182	1299	158	1336

12.3 Nurse staffing

12.3.1 Actual to predicted numbers of all nurses.

Changes in nurse staffing levels were first examined using the number of whole-time equivalents (WTE) for all nurses working in neonatal intensive care in England at each of the three time points. In the unadjusted model, the ratio of actual to predicted number of nurses obtained from the UKNSS was 0.57 (95% CI: 0.55 – 0.59), with a crude increase between 1997 and 2006 of 0.12 (95% CI: 0.09 – 0.16, and an increase of 0.23 (95% CI: 0.19 – 0.26) for the overall time period from 1997 to 2011.

After assessing all potential confounding variables for inclusion in the final model, only the number of low birth weight admissions, the total number of deliveries in the hospital and the total number of days respiratory support provided had an impact on the relationship between the different time periods and the overall staffing ratio. This resulted in an increase in the staffing ratio from 1997 to 2006 of 0.10 (95% CI:

0.07 – 0.14), and from 1997 to 2011 of 0.17 (95% CI: 0.13 – 0.21); the addition of a geographical variable (strategic health authority) changed only the confidence intervals and not the estimates, as shown in table 12.5. Inclusion of units' designated network level in the model did not change the estimate or the confidence intervals from these results.

Change in staffing was then investigated for all neonatal units in England between 2006 and 2011. Initial analysis demonstrated a baseline ratio of 0.70 (95% CI: 0.67 – 0.72), with a subsequent unadjusted increase of 0.11 (95% CI: 0.07 – 0.15). Only two factors were identified that exhibited an important confounding effect: the total number of days of respiratory support provided and the number of consultants with 50% or more of their time dedicated to neonatology; the adjusted result showed an increase in the staffing ratio of 0.08 (95% CI: 0.04 – 0.12).

The final assessment of changes in nurse staffing were made using the tertiary data set that comprised neonatal intensive care units that had data available at all three time points. This selection of neonatal units demonstrated an almost identical picture (shown in table 12.5) in the unadjusted model – albeit with a higher baseline of 0.67 (95% CI: 0.57 – 0.78) – to that seen in the primary analysis. However, the fully adjusted model, accounting for total number of days of respiratory support, total number of consultants with $\geq$ 50% of their time dedicated to neonatology and, additionally, the network designated unit level, demonstrated smaller increases in the staffing ratio than either of the preceding analyses.

12.3.2 Actual to predicted numbers of nurses qualified-in-specialty.

The baseline ratio of the number of nurses qualified in specialty (QIS) actually working on neonatal units to the predicted number of all nurses was 0.35 (95% CI: 0.32 – 0.38) in 1997 using the primary data set. During the first time interval, up to 2006, the unadjusted model showed an increase of 0.10 (95% CI: 0.06 – 0.015). The factors identified that confounded the relationship between time and the QIS ratio in the primary data set were the number of admissions of low birth weight babies, the number of consultants with $\geq$ 50% of their time dedicated to neonatology, and both the designated level and SHA of the neonatal unit. Following adjustment for these, only weak evidence of a difference was identified for the first time period, with an estimated increase of 0.07, p=0.045 and the 95% confidence interval ranging from 0.00 – 0.13. The evidence

Table 12.5: Regression estimates for the absolute change in the ratio of actual to predicted numbers of nursing staff available for neonatal care in England between 1997 and 2011 using three different underlying populations [Primary: all intensive care units at all time points; Secondary: all neonatal units in 2006 and 2011; Tertiary: intensive care units with data at all three time points].

Model	Baseline – 2006		Baseline – 2011	
	Estimate	**(95%CI)**	**Estimate**	**(95%CI)**
Primary analysis (Baseline = 1997)				
Unadjusted model	0.12	(0.09–0.16)	0.23	(0.19–0.26)
Full adjusted model	0.10	(0.07–0.14)	0.17	(0.13–0.21)
Fully adjusted model, including geography	0.10	(0.04–0.15)	0.17	(0.10–0.24)
Secondary analysis (Baseline = 2006)				
Unadjusted model	—	—	0.11	(0.07–0.15)
Full adjusted model	—	—	0.08	(0.04–0.12)
Tertiary analysis (Baseline = 1997)				
Unadjusted model	0.13	(0.09–0.17)	0.22	(0.18–0.26)
Full adjusted model	0.06	(0.01–0.11)	0.12	(0.06–0.18)

became stronger when looking at change over the entire study period, from 1997 to 2011, with an estimated increase in the ratio of 0.10 (95% CI: 0.02 – 0.17) and p-value of 0.014.

When this analysis was repeated using the secondary data set, the total number of days of respiratory support provided along with the neonatal network in which the unit was located were the only two factors that had a confounding effect. Including these in the model reduced the estimated change between 2006 and 2011 from 0.07 (95% CI: 0.03 – 0.11) to 0.05 (95% CI: 0.01 – 0.10). Complete results for the secondary analysis are shown in table 12.6.

12.4 Throughput

Throughput was assessed by looking at the number of babies admitted into neonatal units in relation to the number of cots available (T_{a}, as shown in equation 11.6) as well as by using the number of days respiratory support provided per cot (T_{r}, shown in equation 11.7).

Table 12.6: Regression estimates for changes in the ratio of actual number of QIS nursing staff compared to the predicted number of total nurses available for neonatal care in England between 1997 and 2011 using two different underlying populations. Primary: all intensive care units at all time points; Secondary: all neonatal units in 2006 and 2011.

Model	Baseline – 2006		Baseline – 2011	
	Estimate	**(95%CI)**	**Estimate**	**(95%CI)**
Primary analysis (Baseline = 1997)				
Unadjusted model	0.10	(0.06–0.15)	0.15	(0.11–0.20)
Full adjusted model	0.07	(0.00–0.13)	0.10	(0.02–0.17)
Secondary analysis (Baseline = 2006)				
Unadjusted model	—	—	0.07	(0.03–0.11)
Full adjusted model	—	—	0.05	(0.01–0.10)

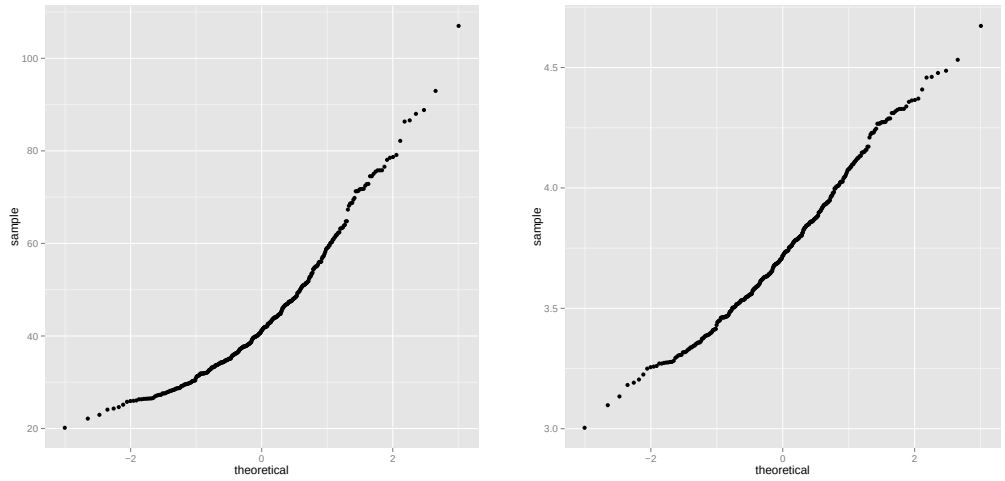
12.4.1 Admissions per cot per year

Admissions throughput was examined first. For this constructed variable, T_{a}, the distribution obtained was not normal, exhibiting skewness of 1.089 and kurtosis of 1.095, shown as a QQ plot in figure 12.3a. Consequently, this was log-transformed to provide a more acceptable measure (shown in figure 12.3b) with skewness and kurtosis of 0.3676 and -0.3818, respectively.

Initial, unadjusted analysis showed that there was no difference between 1997 and 2006 (39.54 (95% CI: 37.69 – 41.49) v.s 40.31 (95% CI: 38.23 – 42.50) admissions per cot), but that between 1997 and 2011 there was an increase of 22% (95% CI: 14% – 31%) to 48.31 (95% CI: 45.75 – 51.00). Consultant and nursing numbers, and the designated unit level within the network, as well as both the available geographical variables, SHA and network, were identified as potential confounding variables. Of these, only the network designated unit level and the numbers of nursing staff had an important impact on the overall model, causing an increase in the baseline number of admissions per cot to 60.56 (95% CI: 46.54 – 78.80); however, no major difference was shown for the effect of time compared with the unadjusted model (table 12.7). Consequently, it was decided to use 'neonatal network' as a geographical variable, as it was more relevant to the latter time period of the study (from 2006 to 2011), and because it permitted greater variation due to the increased number of categories. The number of deliveries at each hospital was also included, as this is plausibly related to

Figure 12.3: QQ plots of actual versus expected distributions for admissions throughput values in data available to examine the trends over time.

(a) Admissions ratio (untransformed). **(b)** Log-transformed admissions ratio.



both the number of admissions and the size of the neonatal unit, as well as having changed over time. Thus, a second adjusted model was created; again, this showed little difference to the crude model, with no change during the first period of the study (41.59 (95% CI: 31.46 – 54.98) admissions per cot 1997 compared with 41.61 (95% CI: 33.08 – 52.34) in 2006) but evidence of a 23% (95% CI: 11% – 36%) increase overall, to 51.05 (95% CI: 40.63 – 64.13) in 2011.

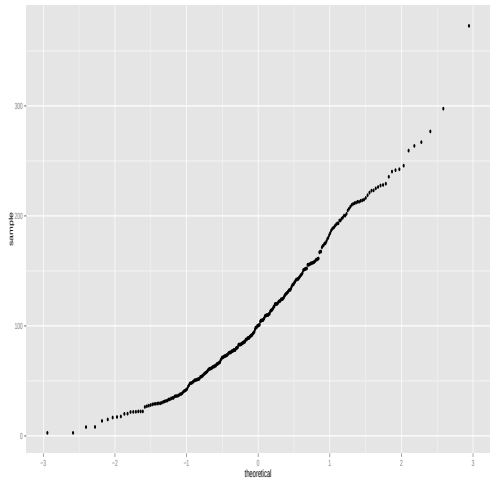
12.4.2 Days of respiratory support per cot per year

As for admissions, the throughput variable constructed with respect to the number of respiratory days support provided per cot was not normally distributed (figure 12.4). However, the overall fit was not improved with either a log or square transformation (figures 12.4b and 12.4c). The use of a square-root transformation would have required overly complex formulae to obtain results for individual changes and, as the majority of the data were linear with only a few values at the extremes distorting the data, it was decided to use the untransformed variable.

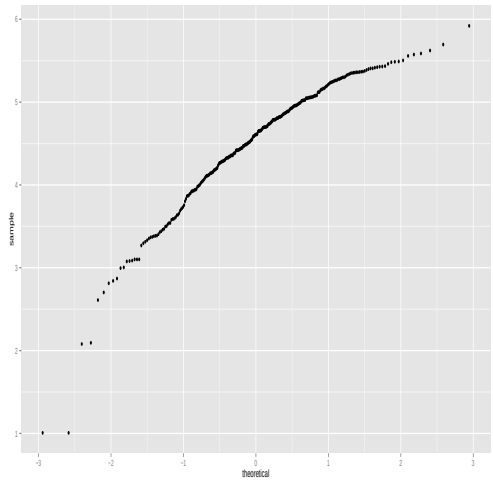
In the unadjusted model, there were 70.01 days (95% CI: 58.29 – 81.73 days) of respiratory care provided per standardised cot per year in 1997, which increased by 46.18 (95% CI: 30.28 – 62.07) between 1997 and 2006, and by 71.95 (95% CI: 55.59 – 88.31

Figure 12.4: QQ plots of expected versus actual distributions for respiratory throughput (T_r) values in data available from the UK Neonatal Staffing Study (1997) and two Unit Profile Studies (2006 and 2011).

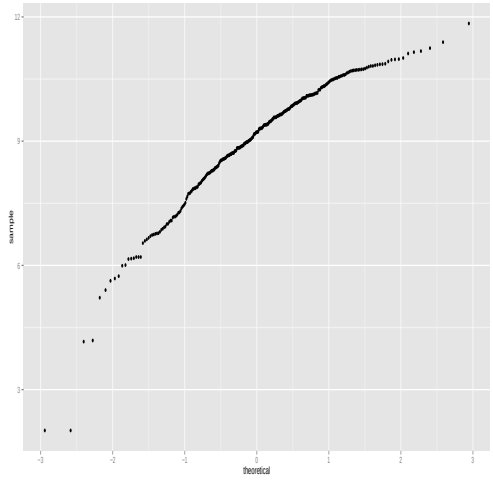
(a) T_r (untransformed).



(b) $\log T_\mathrm{r}$.



(c) T_r^2.



(d) $\sqrt{T_\mathrm{r}}$.

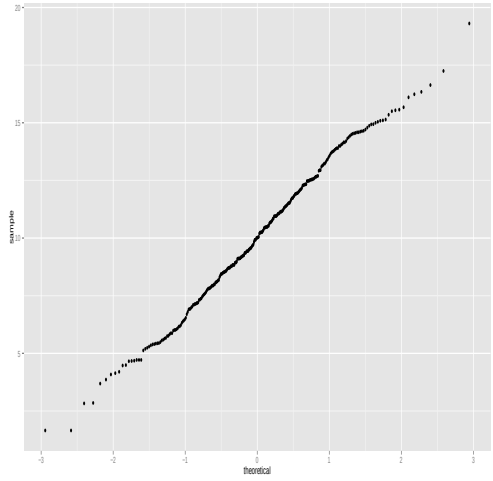


Table 12.7: Regression estimates for changes in the ratio of admissions to neonatal care per standardised cot in England between 1997 and 2011 using two different baseline populations. Primary: all intensive care units at all time points; Secondary: all neonatal units in 2006 and 2011.

Model	Baseline – 2006		Baseline – 2011	
	Estimate	(95%CI)	Estimate	(95%CI)
Primary analysis (Baseline = 1997)				
Unadjusted model	1.02	(0.95–1.09)	1.22	(1.14–1.31)
Fully adjusted model	0.99	(0.90–1.08)	1.23	(1.10–1.38)
Fully adj. inc geog	1.00	(0.92–1.09)	1.23	(1.11–1.36)
Secondary analysis (Baseline = 2006)				
Unadjusted model	—	—	1.18	(1.08–1.30)
Fully adjusted model	—	—	1.23	(1.15–1.32)

days) over the entire study period, from 1997 to 2011 – thereby effectively doubling. Six factors were identified as confounding this relationship, however: the number of consultants with greater than 50% of their time dedicated to neonatal medicine, the total number of nurses, the number of admissions and the total number of babies provided with any form of respiratory support, as well as the level of the unit designated by the local neonatal network and the strategic health authority. When these factors were accounted for, the only evidence for change was during the first period of the study, from 1997 to 2006: during this interval, days of respiratory support provided per cot per year increased consistently across the whole of England by 36.35 (95% CI: 17.49 – 55.20) days. There was no evidence of a change beyond this time point – change in 2011 compared to 1997 was extremely similar to that seen in 2006 (table 12.8), and no change was seen when comparing 2011 to 2006 (−3.07 days per cot per year of respiratory support, 95% CI: −16.66 – 10.51).

These findings were supported by results using the secondary data set. As with the primary data, there was evidence of a change between 2006 and 2011 (table 12.8) before adjustment, but after including just a single one of the identified potential confounders, the total number of deliveries, any relationship between epoch and the ratio of days respiratory support provided per cot was lost.

Table 12.8: Regression estimates for changes in the ratio of days of respiratory support provided per standardised cot available for neonatal care in England between 1997 and 2011 using two different baseline populations. Primary: all intensive care units at all time points; Secondary: all neonatal units in 2006 and 2011.

Model	Baseline – 2006		Baseline – 2011	
	Estimate	**(95%CI)**	**Estimate**	**(95%CI)**
Primary analysis (Baseline = 1997)				
Unadjusted model	46.18	(30.28–62.07)	71.95	(55.59–88.31)
Full adjusted model	36.35	(17.49–55.20)	33.28	(10.04–56.51)
Secondary analysis (Baseline = 2006)				
Unadjusted model	—	—	23.38	(7.96–38.80)
Full adjusted model	—	—	0.90	(-11.72–13.51)

12.5 Intensity

The third set of analyses related to intensity.

12.5.1 Babies provided with respiratory support per admission

The number of babies receiving respiratory support per admission (I_{b}) in 1997 was 0.23 (95% CI: 0.21 – 0.25) in the initial, unadjusted analysis; this was accompanied by an increase in the ratio of 0.10 (95% CI: 0.06 – 0.13) between 1997 and 2006, and an increase of 0.16 (95% CI: 0.12 – 0.19) between 1997 and 2011. Potential confounders of the relationship between time and I_{b} identified from the available variables were the infrastructural variables, including the designated unit level, number of cots and the area of the country – as this may act as a proxy for regional population differences that may influence the incidence of the underlying diseases (e.g. rates of prematurity or genetic problems), and activity variables – specifically, the number of deliveries and the number of low birth weight admissions. Staffing variables were not considered as the number of staff present were assumed to affect the duration of stay (i.e. whether the baby could be cared for on-site, or whether *ex utero* transfer elsewhere was required) rather than whether a baby became "sick" (i.e. required ventilation) in the first place.

Of these, the designated unit level and the number of low birth weight admissions exhibited the greatest confounding effects, hence were included first in the model. Fol-

lowing this, the standardised ("BAPM") cots variable did not improve the overall model at all, but inclusion of the numbers of intensive care, high dependency and special care cots as separate variables *did* have important effects. The total number of deliveries in each hospital then did not change the model further, but inclusion of SHA showed that there were differences around the country. However, there was no evidence of an interaction between any of the variables, meaning that the overall change between 1997 and 2006 was 0.11 (95% CI: 0.07 – 0.15) and, between 1997 and 2011, 0.17 (95% CI: 0.12 – 0.21).

Results from the secondary data set were similar except that the standardised number of cots was used instead of individual cot numbers, and network was used instead of SHA as a geographic variable. In the unadjusted analysis, there was a baseline in 2006 of 0.27 babies provided with respiratory support per admission (95% CI: 0.25 – 0.30 babies per admission). After adjusting for the above factors, there was a country-wide increase in the intensity ratio of 0.05 (95% CI: 0.03 – 0.08).

Table 12.9: Regression estimates for changes in the ratio of number of babies provided with respiratory support per admission into neonatal care in England between 1997 and 2011 using two different baseline populations [Primary: all intensive care units at all time points; Secondary: all neonatal units in 2006 and 2011].

Model	Baseline – 2006		Baseline – 2011	
	Estimate	(95%CI)	Estimate	(95%CI)
Primary analysis (Baseline = 1997)				
Unadjusted model	0.10	(0.06–0.13)	0.16	(0.12–0.19)
Full adjusted model	0.11	(0.07–0.15)	0.17	(0.12–0.21)
Secondary analysis (Baseline = 2006)				
Unadjusted model	—	—	0.07	(0.04–0.11)
Full adjusted model	—	—	0.05	(0.03–0.08)

12.5.2 Days of respiratory support per admission

With respect to I_{d} – the number of days respiratory support provided per admission (equation 11.9) – the baseline ratio in 1997 was 1.98 (95% CI: 1.61 – 2.34) and the unadjusted estimates were of an increase by 1.07 days (95% CI: 0.57 – 1.57) in the first interval, to 2006, and an increase of 1.23 (95% CI: 0.71 – 1.75) days overall. All

infrastructural, staffing and activity variables were considered as potential confounders, with only one variable, the number of consultants on the on call rota, not showing a statistical association with I_d (p=0.203). The optimal model obtained using the forward stepwise approach to regression model building resulted in an adjusted model containing six confounding factors: unit designation, number of low birth weight admissions, the standardised number of cots, total number of deliveries, and of babies provided with respiratory support, and SHA. This model demonstrated an increase from 1997 to 2006 of 1.14 (95% CI: 0.63 – 1.65) days in the intensity ratio, with an overall change of 0.73 days (95% CI: 0.10 – 1.37) between 1997 and 2011. Results from this and from a regression model that included all of the potential confounding variables initially identified are shown in table 12.10. The latter analysis demonstrated reduced effects for both time periods, with no evidence for a change between 1997 and 2011.

Table 12.10: Regression estimates for changes in the ratio of days of respiratory support provided per admission into neonatal care in England between 1997 and 2011 using two different baseline populations [Primary: all intensive care units at all time points; Secondary: all neonatal units in 2006 and 2011].

Model	Baseline – 2006		Baseline – 2011	
	Estimate	**(95%CI)**	**Estimate**	**(95%CI)**
Primary analysis (Baseline = 1997)				
Unadjusted model	1.07	(0.57–1.57)	1.23	(0.71–1.75)
"Optimally"-adjusted model [a]	1.14	(0.63–1.65)	0.73	(0.10–1.37)
Model with all variables included [b]	0.85	(0.29–1.42)	0.49	(-0.23–1.21)
Secondary analysis (Baseline = 2006)				
Unadjusted model	—	—	0.27	(-0.18–0.72)
Full adjusted model [c]	—	—	0.02	(-0.28–0.33)

[a] Adjusted for: unit designation, number of very low birth weight (VLBW: < 1500g) admissions, adjusted number of cots, total deliveries, total number of babies provided with respiratory support, strategic health authority.

[b] Adjusted for all the above factors, plus total number of all nurses in whole-time equivalents, number of consultants with time dedicated to neonatalogy, and number of consultant ward rounds per week.

[c] Adjusted for unit designation only.

Using the secondary data set to look at the same outcome showed no effect in unadjusted analysis: there was a point estimate of 0.27 days increase in the number of days respiratory support provided per admission, but the 95% confidence interval crossed zero, ranging from -0.18 to 0.72, indicating there was no difference. Of the

potential confounding variables, only the total number of nurses in WTEs had an important impact on the effect over time, suggesting a decrease in the number of days respiratory support provided of 0.41 (95% CI: 0.08 – 0.74) days. This was cancelled out by the inclusion of other variables; in the model including all potentially confounding variables, the estimated effect was -0.13 with a 95% confidence interval of -0.44 to 0.18. The fully adjusted model shown in table 12.10 included only the first additional factor that was added into the model – unit designation, which exerted the strongest confounding influence.

12.6 Summary of results

In summary, this chapter has presented the results from the assessment of trends over time of hospital organisation, specifically focusing on activity and staffing. First, unit types and numbers for each epoch were shown: 189 units were identified that existed during at least one of the time periods studied. Within these were two groupings of hospitals: those providing intensive care at each of the time points (used for the primary analyses) or, for the data from 2006 and 2011, *all* neonatal units (secondary analyses). Unadjusted trends over time were then examined. Density values for each of the variables was plotted (figure 12.2) and tables constructed of medians and IQRs (tables 12.1 and 12.3 and country-wide totals (tables 12.2 and 12.4).

The first outcome examined was nurse staffing, expressed in terms of the ratio of actual number of nurses compared to the recommended number of nurses according to BAPM guidelines. This demonstrated an increase in the total number of nurses over both time periods (1997 – 2006 and 2006 – 2011) in both the primary and secondary analyses; however, staffing in the initial period was noted to be low (unadjusted ratio of 0.57, 95% CI: 0.55 – 0.59) and did not reach the recommended level at any point for which data were available. A tertiary analysis using only the 123 units with data at all three of the time points additionally identified similar results. A further analysis (section 12.3.2) examining only those nurses who were qualified in specialty also demonstrated evidence of an increase over time, albeit lower than for all nurses.

"Activity" was measured as both throughput and intensity. Throughput reflected both the number of admissions per cot per year (section 12.4.1) and also the number of days of respiratory support provided per cot per year (section 12.4.2). The admissions

throughput ratio was log-transformed prior to analysis. In the primary analysis, no change was seen between 1997 and 2006 but a 23% increase was found during the second time period; this was confirmed in the secondary analysis using data from all neonatal units who supplied data for 2006 and 2011. Conversely, the throughput variable related to the number of days respiratory support provided demonstrated an increase only during the first time interval in adjusted analysis. No evidence was seen of any change between the two Unit Profile Studies in either the primary or secondary analyses.

The final set of analyses related to intensity. This was measured per admission, first as the number of babies provided with respiratory support and then as the number of days of respiratory support provided. Both sets of analyses demonstrated increases in intensity from 1997 to 2006. From 2006 to 2011, the increase in the number of babies requiring respiratory support per admission was smaller, although the time period was also shorter by a similar proportion (approximately half). The days of respiratory support provided did not change after models were fully adjusted with all available variables; however, in an "optimally"-adjusted model, a small increase was seen. No difference was seen between 2006 and 2011 in the secondary analysis either before or after adjustment.

Chapter 13

Discussion of the trends analysis

The investigation into changes that occurred between the UK Neonatal Staffing Study (UKNSS) in 1997 and the second Unit Profile Study (UPS) in 2011 focused on three areas of workload assessment: neonatal staffing and unit activity, examined in terms of throughput and intensity. This chapter discusses the findings from the study, commencing with a brief summary of the key results and, similar to the previous discussion chapters for the Admissions Validation Study and obstetric antecedents investigation, continues by examining the limitations of the exercise. The results are then interpreted and their generalisability discussed before the chapter is summarised.

13.1 Key results

There were 189 neonatal units existing at one of the three time points for study; data were only available for units providing intensive care in the first time period (data from the UKNSS) for 145 units, whereas at each of the other two time points (data from the UPS 2006 and UPS 2011) there were data from units providing both intensive care and temporary care only. This meant that investigations were conducted with references to two baseline populations: the population of babies receiving neonatal *intensive* care (from 1997 to 2011) and the population of babies receiving *any* form of neonatal care (in 2006 and 2011). Data from all units in 1997, and level 2 and 3 units (corresponding to local neonatal units (LNUs) and network neonatal intensive care units (NICUs) in 2011) in the Unit Profile Studies were used in "primary" and "secondary" analyses, respectively.

13. DISCUSSION OF THE TRENDS ANALYSIS

Table 13.1 provides an overview of the analyses that were performed. For each of the areas of workload assessment, two sets of questions were asked, each in both of the available (primary and secondary) data sets. With respect to nurse staffing, there were increases in the total number of nurses as well as those who were qualified in specialty at both time points. However, the ratio of actual to predicted nurses in intensive care units in 1997 in the unadjusted primary analysis was 0.57 (95% CI: 0.55 – 0.59) and, by 2011, this had only increased by 0.23 (95% CI: 0.19 – 0.26) in unadjusted analysis, and by 0.17 (95% CI: 0.13 – 0.21) following adjustment – thus still remaining well below the level at which the actual number of nurses equalled the minimum recommendation. These findings were mirrored for the model using QIS nurses as the numerator, albeit the increases seen were smaller, barely attaining statistical significance for the adjusted analysis looking at the change between 2006 and 2011 in all neonatal units (ratio increase: 0.05, 95% CI: 0.01 – 0.10).

Table 13.1: Direction of effects seen in adjusted analyses examining nurse staffing, throughput and intensity in English neonatal units between 1997 and 2011 using data from the UK Neonatal Staffing Study, Unit Profile Study (UPS) 2006 and UPS 2011. '↔' indicates no change seen.

	Analysis	Symbol	Time interval	
			1997 - 2006	2006 – 2011
Nurse staffing	All nurses per cot	S_{nurse}	↑	↑
	QIS nurses per cot	S_{qis}	↑	↑
Throughput	Admissions per cot per year	T_{a}	⟷	↑
	Days of respiratory support per cot per year	T_{r}	↑	⟷
Intensity	Babies provided with respiratory support per admission	I_{b}	↑	↑
	Days of respiratory support per admission	I_{d}	↑	⟷

In the throughput analyses, the number of *admissions* per cot per year was not associated with any change between 1997 and 2006, whereas the number of *days* of respiratory support per cot per year was. This finding was reversed between 2006 and 2011: the number of days of respiratory support per admission per year remained the same, whereas the number of admissions per cot per year increased. For the intensity analyses, both measures resulted in an increase after adjustment for confounding factors between the UKNSS and the UPS 2006. The number of days of respiratory support again remained the same in the following period, from 2006 to 2011; the number of babies supported per admission, however, increased.

13.2 Limitations

This study had a number of limitations, as well as some strengths. Issues were particularly related to information bias, but also relate to selection bias, chance and potential confounding.

13.2.1 Selection bias

The study's main strength was that it was so comprehensive. It included 100% of intensive care units at all three time points, as well as all level 1 units in 2006 and 43 of 49 special care units (SCUs) in 2011. This means that the primary data set used in the investigations – including *intensive* care units – cannot be affected by selection bias, as data were complete. However, there is a possibility that the results of the secondary analyses were affected by selection bias because there were six units that did not respond to the 2011 survey. This is unlikely to have been a major source of error as the missing data belonged to SCUs that do not regularly care for babies requiring intensive care and cater primarily to low-risk populations.[20, 22] A bigger concern relates to why there was a complete absence of data from units providing temporary care only in 1997 when it is known those data *were* collected originally; however, that does not affect the results from this study, merely who the results may be applied to, as discussed in section 11.3 on the study populations.

A related worry is with the data missingness for some variables – particularly, those related to the "activity" measures of respiratory support (number of babies and number of days of different modalities). For the UKNSS, this was reported as being due

in part to differing (and inadequate) data collection practises at individual hospitals, although response rates for these questions varied from 18% to 76%.[28] The low reporting numbers can, however, also be seen in the other time periods (see table 12.1) – and, certainly in 2011, these data should have become much easier to gather.

13.2.2 Information bias

Many of the limitations of this study relate to potential information biases within the data that were collected and available for analysis.

Firstly, there is the unknown provenance of the UKNSS data. The data collection is well described [29, 234] and a copy of the questionnaire is available (appendix C). However, the data weren't checked at the time they were received from the UKNSS investigators in 2005 and subsequently turned out to not only have a substantial proportion of missing data, but also to contain data that were not collected in the initial survey! This is a crucial point to note, as the additional variable represented the number of high dependency cots at units, hence was a vital part of the "BAPM cot number" (equation 11.1) that was used to standardise results between units of different size.

The presence of this variable also highlights the fact that we don't *really* know what happened with data collection for the other variables. For example, were the supplied data only obtained via the UKNSS survey, or are they supplemented with information obtained from elsewhere? If they did only come from the survey, what data collection methods did individual hospitals base their reporting on? It is unlikely that many hospitals had computerised data collection systems for recording information such as that collected by the UKNSS; indeed, even now, not all of the data asked for in the surveys are routinely collected in all hospitals, nor gathered at regional or national levels in a systematic fashion. If they were, it wouldn't have been necessary to carry out the Unit Profile Studies.

Indeed, the concern relating to potential differences in data collection by units during the UKNSS also pertains to differences in data collection by (the same or different) units *between* the three surveys. It is unlikely that units are still recording all of this data in the same fashion that they were in the mid to late 1990s; methods of measurement may also have changed – for example, from retrospectively counting the number

of individual babies admitted from those recorded in a log book, to prospectively reporting via systems like BadgerNet.[114] Most systems will have improved over time, hence data may be more accurate for the later time points.

Conversely, knowledge about what the data refer to may have decreased over time, thus affecting their collection. For the 2011 Unit Profile Study, data were extensively cleaned and checked, as described in section 4.5.4. Confusion was noted in responses to staffing questions, particularly the difference between requests for total numbers or whole-time equivalents. This cannot be explained by a lack of detail in the questionnaire: these questions were very similar to those of the two previous surveys, and had been piloted. Therefore, it must reflect on the respondents: either they did not read the questions properly, or they did not understand them. Other data also lead to the same thought: a number of hospitals reported the same value for the total number of admissions (including re-admissions) and the total number of individual babies admitted (questions 4 and 5 of the survey, appendix E). Were these data true records of what actually happened, or was there confusion related to interpretation of the questions?

Some questions were, with hindsight, poorly thought out. The failure to obtain usable data relating to junior medical staff has already been highlighted in section 4.5.1. These questions have also been examined by a more specific "Medical Workforce Census" carried out by the RCPCH on 30th September, 2011.[235] This examined paediatric and neonatal staffing, including some hospital level measures, across the whole of the UK. However, the breakdown of staffing into neonatal or paediatric was not provided at a country level, and description of neonatal units did not distinguish between those that were stand-alone or that were accompanied by obstetric services.[235]

The data consistency checks and other cleaning that were performed for the UPS 2011 do not appear to have happened contemporaneously for the 2006 survey. There were, though, established contacts at each hospital for the EPICure study at that time who may therefore have been more motivated to complete the questionnaires. And, as already noted, data provenance for the UKNSS is uncertain.

What about coding of the variables that happened during the investigative stage? Specifically, were the units coded appropriately in 1997? As can be seen from figure 12.2, the units were categorised as equivalent to level 3 units (network NICUs); they might also have been coded as level 2 (LNUs), or given an entirely separate categorisation. This was accounted for by repeating the analyses using just the 2006 and 2011

data sets – which additionally included SCUs. In all analyses, the results from the secondary data sets mirrored those obtained in the primary data. Furthermore, the primary function of the 'unit designation' variable was to distinguish between units *within* that time period; as there was only one kind of unit (intensive care units) in 1997, there weren't sufficient groups to distinguish between, and hence the choice of category was irrelevant. If data for the other units *had* been present, then it might have been useful to include the variable reflecting units that gave care on a temporary basis only.

Were the unit designations correct subsequently? There are differences in reported numbers from other publications.[85, 86, 232, 235] In the 2011 RCPCH workforce report, 170 units are noted in England, of which 48 are reported as being neonatal intensive care units, 76 local neonatal unit and 46 special care unit; the methods state only that levels are based on BAPM designation [21] and the 2009 Department of Health (DoH) Toolkit.[20] In this study, the methods have been clearly stated, including exactly how a "perinatal unit" was defined; the methodologies employed by others are less clear: for instance, surgical centres located within children's hospitals were probably also included.[232, 235]

Another concern previously noted in section 4.5.4 relates to the reporting of cot numbers by individual units. Were the numbers of open/available cots reported (i.e. those for which staff were available?) or did responses reflect the total numbers of cots that *would be* available if there were sufficient staff? Plausibly, this would apply equally to all three of the surveys, meaning misclassification is distributed non-differentially over time and, thus, most probably caused a dilution of any effects seen.

13.2.3 Chance

Chance could have been a factor. However, important effects were seen in almost all analyses with confidence intervals that did not overlap the null value, and were consistent between the data sets used in terms of direction and magnitude.

Additionally, attempts were made to guard against the possibility of chance findings by being cautious in the creation of the ratios used as outcomes. Calculation of staffing was based on a conservative estimate of the number of (all) nurses required (as shown in equation 11.3), based on previous studies,[26, 233] rather than the estimate that is currently used.[67, 232] The use of a lower multiplier to establish the predicted number

of nurses means the resultant staffing ratios had a lower denominator, hence would have been higher overall. Therefore, fewer units might have been expected to have staffing ratios below a certain threshold. In actual fact, nearly all units were understaffed, with a ratio below one, and the average ratio in 1997 was found to be 0.57 (95% CI: 0.55 – 0.59) – i.e. on average, units were operating with only 57% of the recommended number of staff.

13.2.4 Confounding

The standardisation of outcomes by creating ratios was performed to minimise the effects of confounding. By basing ratios on cot numbers for the staffing and throughput analyses, potential confounding by unit size was controlled for. For each of these analyses, too, all other variables were considered as potential confounders. The intensity analyses, in comparison, sought to examine how much work was required for each baby – both in terms of how many babies required invasive respiratory support, and how many days of support were provided. In the first (but not the second) of these two analyses, staffing data were not considered as potential confounders, as they were considered to be related to length of respiratory support provided (low staffing might necessitate transfer elsewhere for continued care) rather than how sick a baby might be. However, all other factors were considered as potential confounders.

Consequently, there is unlikely to be significant confounding by the factors measured and accounted for in this study. Instead, the possibility of residual confounding needs to be considered. Within the data collected for the study, this possibility was minimised by using continuous data wherever possible; the maximum information available was therefore extracted from the data. Thus, concerns about residual confounding affecting the current investigations relate primarily to data not collected by all three surveys, and therefore not available for comparison. This may be an issue, although the surveys included in this study garnered a wide variety of data reflecting multiple domains of neonatal care, and much of which was compatible enough to facilitate comparison. Furthermore, the data collected represented measures that were proximal to each other; other potential confounding variables such as policy could be considered to be more remote from neonatal care – and, in fact, be represented by the exposure used in this study, “time”.

13.3 Interpretation

Given these many concerns, one interpretation might be that the results from this study are completely invalid and should be disregarded entirely. That would be unfair: the data are comprehensive in terms of the units covered, as well as the breadth of questions asked, with the majority of data complete. Chance is unlikely to have played a role in the findings and the most important potential confounders were accounted for. Information biases, although many, tended to be non-differential, leading towards the likelihood of any effects being reduced. However, marked effects *were* found.

Further comfort is provided by the consistency between analyses. Most importantly, the primary and secondary groupings of units that were used produced virtually identical results throughout. With respect to staffing, increases were seen across the board – in all nurses and in only those who were QIS – in both time intervals. This implies the increases are true. However, the fact that staffing levels remain well below recommended standards even in the most recent survey data from 2011 is deeply worrying, as the recommendation has remained virtually unchanged throughout the entire study period.[18, 19, 67, 232] However, it should also be remembered that these guidelines are based on consensus opinion rather than a strong evidence base.[18, 19]

There is less certainty regarding the results from the activity analyses due to the increased missingness affecting the activity variables, and interpretation is therefore less easy. By looking at table 13.1, it would appear that between 1997 and 2006, the admissions rate (per cot) stayed the same, but that babies who were admitted were sicker than previously (more were provided with respiratory support per admission) and required more respiratory support overall (reflected by increases in the number of days respiratory support provided both per cot per year *and* per baby). This is congruent with the knowledge that increasing numbers of extremely premature babies were being admitted into neonatal intensive care units in England [55] – exactly the babies who might be expected to be sicker and require more respiratory support.

Between 2006 and 2011, the data are even more confusing. There was an increase in the admissions rate (per cot) and the babies were sicker (more requiring respiratory support per admission). However, the number of days respiratory support provided did not increase – either per cot or per admission – suggesting that, in 2011, babies were sicker but being provided support for shorter periods of time than previously. It is hard

to know what to make of this: it could be erroneous, or it could be true. The data were not complete, and the outcome measures in both analyses that found no effect were based on the number of days of respiratory support provided. But this variable was more complete than the variable representing the number of babies provided with respiratory support. Furthermore, if these results are due to bias, non-differential misclassification would mean that they would most likely be biased toward no effect being seen. Hence, it is possible that there was an increase (or even a decrease) in admissions, more babies were provided with respiratory support *and* the number of days of support provided increased.

Alternatively, if the findings of the study are thought to be accurate, this might indicate changes in practise, for example, towards less invasive ventilation strategies for extremely preterm babies. This change in practice is supported by the increasingly widespread use of high-flow oxygen therapy over the interval period, necessitating the inclusion of this modality in the UPS 2011 questionnaire (appendix E).

13.4 Generalisability

Finally, to the utility of these results. The data described by this study identified changes in care in England between 1997 and 2011. They are, therefore, of crucial importance to those who are concerned with the NHS and provision of neonatal services in this country. It is also possible that these findings can be applied outside of England. The NHS operates in the other countries in the United Kingdom, and influences over time have been similar – albeit, not the same – in each of them. This is particularly true for clinical influences, but less true for political influences which may have affected neonatal care organisation at regional or national levels differently to the organisation of care in England and, consequently, affecting the individual hospitals providing perinatal care differently as well.

Applicability of the results beyond the countries of the UK, however, is more questionable. Any impact from managed clinical networks has occurred within a nationalised health system where care is provided free at the point of delivery. Neither these kind of networks nor this philosophy of care exist overseas. In contrast, the study methods would be of use in evaluating care in other scenarios, and these results from English data may therefore be of value as a benchmark.

13.5 Conclusion

The investigation into trends over time in factors relating to the organisation of perinatal units in England was discussed in this chapter. Key findings from two populations of unit were described: a primary data set comprising neonatal *intensive* care units from all three surveys (UKNSS, UPS 2006 and UPS 2011) and a secondary data set containing information from all units who responded to either of the Unit Profile Studies. Summary results for the six outcome measures were presented in table 13.1.

There were many potential limitations, and some strengths, to these investigations. The data were comprehensive and complete in terms of the units that were covered, as well as with regards to the majority of individual factors assessed. However, there was a large potential for the information to be biased, as discussed in section 13.2.2. The provenance of the UKNSS data was not entirely clear, and recording discrepancies for different variables were noted between the data at different time points in this investigation, as well as between this and other studies. Further limitations that were discussed were the potential effects of chance and confounding; particularly, residual confounding may have been an issue, but it was thought that a majority of the most important, proximal confounders were accounted for.

The main interpretations that arise from these findings were described in section 13.3. Staffing levels have increased through out the study period for all nurses and for those QIS; staffing remains below recommended levels, however. Important changes were noted in activity levels during both time periods, although the characteristics were different in each and could be due to a variety of reasons and not just the reorganisation of care into managed clinical networks. These results are of particular importance to those concerned with neonatal care in England, and are also of value to other nations within the UK. This study may also be of relevance to those conducting similar investigations overseas.

Part V

Integrating the findings

Chapter 14

Discussion

This chapter discusses the findings of the different investigations carried out in the thesis and how they may be integrated with each other, bringing together the various results in relation to the original aims. Section 14.1 provides an overview of the key results that were obtained for each of the study objectives. The limitations of the principal findings are discussed next, particularly those attributable to bias, but also in relation to possible confounding factors and to chance. The results are interpreted in the context of the overall thesis, and alternative interpretations considered. The chapter continues with a discussion about the generalisability of the results. This covers the principle findings of the thesis: a possible increase in the numbers of extremely premature babies; issues with routine data collection; organisational relevance; and clinical importance. The implications for further research are discussed in section 14.5, commencing with the outstanding questions. Population characteristics and data requirements are also discussed. In the final section, the overall conclusions of the thesis are stated.

14.1 Key results

Before presenting the key results, it is first useful to review the study objectives in light of the investigations that were carried out. This thesis sought to explain the demographic, clinical and organisational factors that resulted in a 44% increase in admissions to neonatal intensive care of babies born between 22 and 25 weeks and six days gestation between 1995 and 2006 in England [55] and, where possible, to extend

that knowledge to the present time. Three specific aims relating to the overall project were explored:

1. Validation of the EPICure data sets was attempted using HES data to find out whether the perceived increase in admissions to neonatal intensive care units was true. It was not possible to fully investigate the reasons behind this increase.

2. The impact of antenatal steroids, tocolysis and delivery by Caesarean section on perinatal outcomes was evaluated using data from the EPICure 2 cohort.

3. Organisational changes that have occurred in neonatal care in England since the 1990s were examined.

Results are presented according to each of these objectives.

14.1.1 Investigation of increased admissions to NICU

The first objective of the Admissions Validation Study was to perform data linkage to garner additional information with which to investigate the increased number of admissions seen between 1995 and 2006 in the English EPICure data.[55] The second objective – which, in fact, was a necessary prerequisite to the first – was to investigate both the completeness of the EPICure data *and* of the routine data set, Hospital Episode Statistics. However, it turned out that linkage was extremely poor, with only around 60% of EPICure records successfully matched. Furthermore, many variables within the HES data sets were incomplete or altogether missing (appendix F); the data that were present displayed marked inconsistencies, for example, between gestational age and birth weight (tables 6.4 and 6.5). These and similar problems have also been noted by others,[110, 126, 132, 133, 236] and are not evident in the EPICure data sets, both of which were subject to close supervision during periods of data collection.[54, 55] The conclusion drawn from this was that HES data are of poor quality when considering births at the borders of viability and that the EPICure data represent the most accurate picture available of such births that occurred during 1995 and 2006.

Despite the data linkage from the AVS not being as successful as initially hoped, it was still possible to query the HES data. The crude numbers of births identified in 1995 by HES were 621 for live births and 867 for all births; in 2006, it was 887 and 1,188 respectively (table 6.16). This equated to a 42.8% increase in the number of live births –

similar to the increase in the number of admissions seen in EPICure. The total number of births at all gestational ages recorded in HES increased between 1995 and 2006 by 9.7%, from 575,509 to 631,401; a 0.7% increase was seen in corresponding registry data, from 613,257 to 635,748.[194] Combined, these results support the hypothesis that the increase seen in the EPICure admissions was due – at least in part – to increased numbers of births, rather than just to improved survival. They do not distinguish, however, between whether the increased births are due to increased fertility rates (more women giving birth, or more children being born per woman) or an increased population size with current fertility rates maintained. Unfortunately, the HES data are not able to provide us with a way to answer this question.

14.1.2 Obstetric antecedents of prematurity

Indeed, increases in the number of admissions to neonatal intensive care are determined by prior events: specifically, by how many babies are born, and by whether those babies that are delivered survive labour and the first few minutes of life while in the delivery room. While the AVS was designed to explore the former possibility, the investigation of obstetric antecedents aimed to explore the latter. The factors leading to improved perinatal survival may relate to changes in obstetric or neonatal management, including both clinical and organisational factors, or to changes in decision-making.

Part III of the thesis used data from the EPICure 2 cohort to examine the effects of three obstetric antecedents, the use of antenatal steroids, tocolysis and Caesarean delivery, on delivery room outcomes. These outcomes were condition at birth, defined as the presence of a heart rate above 100 beats per minute at five minutes of age, and "perinatal death", which included deaths during labour and those occurring in the delivery room. This study additionally accounted for perinatal decision-making, including variables such as who counselled women and which topics were discussed. Although there were missing data among the counselling variables, results in the main and sensitivity analyses were consistent.

Caesarean section was not demonstrated to be associated with a survival advantage or with the baby being in a better condition at birth, although there was marked evidence of improved outcomes in those born at extremely low gestations where the mother was not in spontaneous labour (OR 12.67, 95% CI: 2.79 – 57.60 for an improved condition at birth; 0.03, 95% CI: 0.01 – 0.21 for perinatal death). This finding

was thought to represent unmeasured effects of obstetric decision-making. Tocolysis and antenatal steroids were both independently associated with improved outcomes at birth in multiply-adjusted regression analyses. Comparison of the EPICure cohorts demonstrated important increases in the use of antenatal steroids (9%, 95% CI: 4% – 13%) and tocolysis (7%, 95% CI: 2% – 11%) between 1995 and 2006 in those who were admitted into neonatal intensive care,[55] hence it is likely that survival was higher in 2006 as a result of this.

14.1.3 Trends in unit organisation over time

Organisational factors relating to neonatal units were assessed in November 2011 through a second Unit Profile Study carried out as part of this thesis; methods were presented in chapter 4. The study was successful in achieving a 96.4% response rate (159 of 165 units, shown in figure 4.1), failing only to get responses from six special care units (SCUs). Forty-four network neonatal intensive care units (NICUs) and 72 local neonatal units (LNUs) responded (figure 12.1). For responding units, data were virtually complete (tables 12.1 – 12.4), and these demonstrated important differences by unit type (table 4.2 and figure 12.1).

The data from the 2011 UPS were used alongside data from the UKNSS in 1997 and the first UPS in 2006 to explore the final objective, the investigation of changes over time in organisational aspects of neonatal care in England. Common data from the three surveys included information about staffing (nursing and medical), activity (relating to admissions and provision of respiratory support) and unit infrastructure (for example, the number of cots). UKNSS data were available from all intensive care units providing care in England in 1997 – but not from units that provided temporary care only. UPS data from 2006 were available for all neonatal units in England. Therefore, primary analyses were conducted looking at changes between 1997 and 2011 in units providing neonatal *intensive* care, and secondary analyses performed using data from all neonatal units providing data in 2006 or 2011.

First, staffing levels were examined in relation to a standardised measure of unit size, created using an algorithm based on the number of cots and nurse staffing recommendations from BAPM.[18, 19, 21] There were increases seen in all analyses: in fully adjusted analyses, there was an increase in the staffing ratio of 0.17 (95% CI: 0.13 – 0.21) between 1997 and 2011 using the primary data, and an increase of 0.08 (95%

CI: 0.04 – 0.12) seen in the secondary analysis. The starting point in 1997 was only 0.57 (95% CI: 0.55 – 0.59), however, indicating there were only 57% of the total recommended number of nurses; this meant the ratio failed to approach the point where actual staffing equalled recommended levels at any point during the study period.

The activity analyses were divided into "throughput" – based on Little's law [179] – and "intensity". Throughput was measured per cot per year in terms of admissions and days of respiratory support provided. Intensity was measured as either the number of babies provided with or the number of days provided of respiratory support per admission. Between 1997 and 2006, there were important increases in all measures of activity other than the number of admissions per cot per year. This indicates that the number of admissions remained in proportion to the number of cots but a higher proportion of babies required and were provided with respiratory support in 2006 than in 1997. Consequently, the number of days of respiratory support provided per cot per year also increased. These findings are consistent with those of the EPICure study and also from the comparison of the HES data sets. They suggest that there was an increase in the numbers of extremely premature babies that was probably disproportionate to the number of overall babies, hence contributing a group of babies who were both sicker and required ventilation for longer than the number of babies that had previously been cared for.

From 2006 until 2011, in contrast, admissions per cot per year and the number of babies provided with respiratory support per admission increased, but there was no change in the number of days of respiratory support provided either per cot per year or per admission. As discussed in section 13.3, this indicates that more babies needed respiratory support in 2011 than 2006 but did not required as much support each. It is not clear whether this finding is due to changes in practice during the intervening time interval, or whether it is erroneous.

14.2 Limitations

Many limitations impacted the results. The extent of the problems with the HES data was one of the most significant hurdles, and had not been fully anticipated. This meant it was not possible to combine results from the AVS and obstetric antecedents investigation in the manner originally envisioned. Data from the AVS were meant to be

used to identify how much of the increase in admission numbers seen in EPICure data was attributable to an increased number of births. However, some of the increase was also thought possibly to be due to increased survival, and it may have been possible using the data from the obstetric antecedents investigation along with information gleaned from HES to quantify the fraction attributable to this. Therefore, it is only possible to state that there is likely to have been "some" increase in the number of births without being able to quantify the amount. Similarly, it is possible to say that "some" of the increase seen in admissions is likely to have been due to changes in management – that is, due to the effect of steroids and tocolysis in babies who would not previously have been treated – but it is again not possible to quantify this.

Furthermore, the original expectation was that the HES data would be of sufficient quality to then use the equivalent data for 2011 to assess outcomes in combination with the repeat UPS. Although this was not possible, objectives were able to be satisfied using outcome measures derived directly from the organisational data collected – albeit, with reference to all babies cared for on neonatal units, rather than just those born extremely prematurely.

These facts, however, do not detract from the findings of the investigations that *were* carried out, nor from examination of how they reflect on each other and a discussion of potential limitations. Answers *were* obtained for each of the original objectives but, as for the individual investigations, how these are interpreted depends upon the impact of any confounding factors as well as chance, bias and any other methodological considerations.

14.2.1 Confounding

The primary goal of this thesis was to examine whether the increase over time seen in the EPICure data of neonatal admissions was true; the exposure therefore is "time", with the outcome being the number of admissions contained in the data set. Although the linkage did not fulfil its original purpose, the HES data showed commensurate increases in live and all births to EPICure. Importantly, this may have been confounded by the data completeness of HES, which had 575,509 (93.68%) records compared with 613,257 registered births in 1995, and 631,401 (99.32%) of 635,748 in 2006;[194] accuracy of coding may also have improved over time.

In both the investigations of obstetric antecedents and the analysis of trends over time in unit organisational characteristics, a common strength was that statistical models were built using forward stepwise regression techniques. This took place in both cases after *a priori* consideration of potential confounding factors from amongst the variables available for analysis. However, in both cases, the possibility of residual confounding remains, as discussed in sections 10.2.3 and 13.2.4, respectively. This is particularly noteworthy regarding the influence of perinatal decision-making on the increased numbers admitted into intensive care seen in EPICure. Despite taking many counselling factors into account, there was still evidence of probable decision-making influences on outcome following Caesarean delivery. The importance of this residual confounding is likely to be more significant when the study is interpreted more broadly – for instance, if interpreting the results as supportive of the notion of an improvement in survival between 1995 and 2006 due to changes in management or decision-making.

In contrast, the concern about residual confounding affecting interpretation of the results diminishes when the results from all three investigations are combined. This is because the finding that admission rates per cot remained constant between the first study epoch (1997) and the second (2006) but babies were sicker and needed more respiratory support is congruent with the findings in the HES and EPICure data sets (between 1995 and 2006) that there were increased numbers of extremely premature babies being born, potentially forming a greater proportion of those admitted. There is thus a consistent story being told for this time interval across multiple data sets, each collected in a different manner. Residual confounding factors will vary across these three studies, therefore potentially cancelling each other out – although some may still be shared in common.

14.2.2 Chance

There is no single, overall statistical test which can consolidate the findings from the different pieces of work undertaken. A qualitative rather than quantitative answer is all that can be offered in response to the question, "has there been an increase in extremely premature babies?" Multiple pieces of evidence support the notion that there *was* a true increase in admissions to neonatal intensive care, but none provide sufficient information with which to confirm it, or to attribute possible causes or estimate the magnitude of their effects.

14.2.3 Bias

The remaining area of worry with regard to the overall interpretation of the thesis is that the findings are biased. The data quality concerns expressed about HES with respect to missingness indicate the possibility of selection bias affecting the conclusion that extremely premature births have increased between 1995 and 2006. It is not known what data were missing, nor why there was a difference with respect to the proportions of all births supplied by the NHS HSCIC. A further feature of HES was data inconsistencies, which also must be factored in. These considerations are, however, both countered by the fact that results from the HES data *support* the EPICure finding: there were consistent increases in populations of extremely premature babies identified within the HES data in different fashions – using gestational age as coded in HES; using those defined as "true" matches from the combined linkage analyses; and using the population of babies who were both reported by HES *and* identified in the linkage analyses. The low quality of the HES data also point to the conclusion that the EPICure data should be considered the "gold standard", hence reinforcing the conclusion that there has been a probable increase in (live) births at extremely low gestations.

Strong evidence about obstetric interventions, particularly antenatal steroids and tocolysis, was provided in the obstetric antecedents of prematurity investigation – for those women who *delivered* below 27 completed weeks of gestation. As discussed in section 10.2.2, we cannot know what happened to *all* women who may have received treatment. Does this affect interpretation of the thesis? The obvious answer is "unknown" – for it is not possible to know what effect such potential bias might have. But this applies only to magnitude (and direction) of effect; the true answer needs to acknowledge the population under consideration and, therefore, yes: interpretation *is* affected. Results refer only to those women who are both admitted with threatened extremely preterm labour *and* that go on to deliver; we cannot know with certainty in advance *who* will deliver, hence any guidance (e.g. for counselling and decision-making, or regarding clinical interventions) drawn from the conclusions will be biased.

For the investigation of trends over time in unit organisational characteristics, the possibility of information bias was discussed in section 13.2.2, with selection bias unlikely due to the high response rates. Due to concerns regarding provenance of the UKNSS data and variations in reporting, it is useful to re-examine the findings in the

context of the other investigations. Again, the congruence of messages between studies during the first time interval, from 1995-7 to 2006, is reassuring, whereas for the 2006 – 2011 time period, it is harder to comment. One discrepancy highlighted previously that is worth revisiting is determination of unit numbers, particularly for the 2011 survey. Figure 4.1 shows that 12 units merged to become six from the initial list of units provided by NNAP. An important question to ask here is how a "neonatal unit" is defined: in the 2011 survey, several respondents stated that they were including data for two sites – for example, a LNU and accompanying SCU in a nearby hospital that was covered by the same group of consultants operating a single rota. Thus, is a "unit" the buildings, the organisation (i.e. NHS trust) or the doctors (and other staff) who work together? It is important again to ask how this might affect interpretation of the thesis. Unfortunately, no other data were available covering this time period, nor have other studies looked at changes in neonatal care at that time. Any bias in the 2011 data arising from potential unit classification is likely to be small, however, as only a small number (fewer than 10% of units) merged, and likely to be non-differentially distributed among the different unit levels.

14.2.4 Methodological considerations

The final way the results of the thesis may have been affected is through the choice of methods used. I chose to use free software – 'free as in "free speech," not as in "free beer"'.[237] This meant that R [219] rather than a commercial package like Stata, [182] which co-investigators used in other studies (for example, Marlow et al [70]), was used for statistical analysis. Analyses might have been quicker if I had been able to access help from colleagues. There are, on the other hand, benefits to using open source software such as greater versatility. With respect to other aspects of the investigations, the approaches I used were conventional: rationales for choice of potential confounding variables in the analyses are presented where appropriate, and are consistent with those used by other investigators. The work was entirely dependent upon data that had already been collected, hence the choice of variables for use in analyses was inherently guided by the original study investigators, be they the EPICure research team, UKNSS Collaborative Group, or the many parties involved in the development and collection of HES data over the years. Furthermore, participation in the Neonatal Economic, Staffing and Clinical Outcomes Project ensured that all collaborators were able to explore ideas

with contemporaries. This greatly benefited the UPS 2011 and, consequently, the subsequent investigation into workload assessment over time described in Part IV.

14.3 Interpretation

Having considered the potential limitations of this thesis, interpretation can be structured into two components: first, interpretation of the overall aims and objectives of the thesis; and second, alternative interpretations – ones that may be less likely but should not be discarded entirely.

14.3.1 Overall interpretations

This thesis sought first to determine if the increased number of admissions seen in England for babies aged 22 to 25 completed weeks of gestation between the two EPICure studies was true. Changes seen within the HES data supported this, despite their potential limitations; the changes in unit activity that were seen between 1997 and 2006 are also consistent. Additionally, if there was increased usage between 1995 and 2006 of antenatal steroids, which seems likely, and more appropriate use of tocolysis, which is certainly possible, then it is plausible to think that these together have lead to improvements in the survival rate. Thus, the conclusion is that there is likely to have been an increase in admissions, although it is not possible to quantify exactly through which mechanisms it occurred.

The thesis also sought to explore how unit and network level effects on outcome following extremely premature delivery had changed over time. It was not possible to do this, as outcome data did not exist for extremely premature babies at the third time point, in 2011. Instead, unit level measures for staffing, throughput and intensity were used as outcomes. In order to interpret the findings, it is helpful to divide the investigation into the two time intervals, from 1997 to 2006, and then from 2006 to 2011. The changes that occurred during the first study interval were consistent with both an increasing number and increasing proportion (among all babies admitted) of extremely premature babies: more babies required respiratory support overall, and more days of respiratory support were provided per admission and per cot. The data therefore support (and are supported by) the conclusion to the first aim, leading to

the interpretation that they are an accurate representation of changes that occurred at that time.

There is less certainty regarding the activity data from the second period, 2006 to 2011: it appears that admissions increased in relation to the number of cots available in England and that, of the babies admitted, more babies required respiratory support. Confusingly, days of respiratory support required in relation to either cots or admissions remained unchanged. While the interpretation that results from the first interval are correct tends to support the validity of the data during the second time period – as all methods were the same (both for data collection and analysis) – other data will be required to contextualise the findings and interpret them appropriately.

14.3.2 Alternative possibilities

Other possible interpretations exist. One interpretation might be that there were no increases in extremely premature births, survival of extremely premature babies at birth, or admissions into neonatal intensive care. This seems very unlikely, but cannot be entirely excluded due to the limitations with the data that have already been described. However, even if the interpretation is different, the implications are likely to remain similar. This is because the data are historical, with a primary objective of assessing mortality and morbidity in extremely premature babies.[54, 55] To understand the current situation, more data – with a greater degree of accuracy than previously available from HES – are required.

Indeed, any of the interpretations presented in the thesis thus far may be false. For instance, results obtained in the obstetric antecedents analysis in the population of women who delivered may not be the same as those that would be obtained for the population of women who actually receive such treatments. This is important information to know for counselling purposes. Thus, interpretation becomes that a further study is required with a more inclusive population base of women who are at risk of delivering at these extremely low gestations.

Similarly, perhaps there were no real changes in the workload measures examined in Part IV. There were serious issues with the data that make any interpretation hazardous. Indeed, can the trends over time study even be called an assessment of workload? It is certainly incomplete: “workload” was not actually assessed, merely proxy measures used. Of the three surveys that contributed data, only the UKNSS properly

attempted to measure workload,[29] and that was in a second, separate phase from which data were not available for this thesis. As discussed in section 2.10, "workload assessment" is a nebulous concept, with many intertwining factors contributing towards it being a 'wicked' problem.[178] An alternative interpretation is to say that the information gained is not useful. This view, however, fails to take account of the fact that *something* was shown to have changed by the study – even if it wasn't "workload". Only by trying to understand this further will it be possible to untangle exactly what, and to get a better grip of the problem. This is a key characteristic of 'wicked' problems,[178] which do not have a unique solution and for which definitive descriptions of the actual problem remain elusive.

14.4 Generalisability

The inability to use routine data for investigations of select populations like those who are born extremely premature has implications spanning arenas of clinical medicine, research, ethics, public policy and finances. Moreover, the finding that the number of extremely premature babies has probably increased, seemingly out of proportion to increases seen in the total number born means there will be increased numbers of individuals and families affected by extremely premature birth. There is also broader applicability: the individual investigations have relevance in other developed countries as described previously, and there may also be applicability to health care systems in less well developed countries.

14.4.1 Increasing extreme prematurity

The most important finding of this work is that there was a probable increase in extremely premature babies – particularly, those below 26 weeks gestation – between 1995 and 2006 that appears to have been disproportionate to increases in the total number of babies born and admitted into neonatal intensive care. Although this increase occurred between 1995 and 2006, there have been continued changes since then in the English population with respect to both ethnicity and socioeconomic status, and these changes are ongoing.[238, 239] Ethnicity and socioeconomic status are, additionally, the two factors that were initially thought to be the most important determinants of any change, but were unfortunately not able to be investigated due to data inadequacies.

If this increase in extreme prematurity is true, then it is of great consequence: it leads to questions about *why* this increase has occurred or is occurring; about whether any effects can be mitigated; and, even, if the trend can be reversed.

The impact of extremely premature birth on individuals can be enormous. Survival without impairment improves with gestational age and with birth weight adjusted for gestational age, and is better for girls compared to boys.[240] Survival without morbidity is also higher after antenatal steroids [34, 38, 240]. However, at 11 years old, 45% of survivors who were assessed from the first EPICure study had some kind of serious – particularly cognitive – disability.[58] Less severe morbidity is even more frequent: both respiratory [76] and cardiovascular [77] findings were impaired in the EPICure cohort members compared to controls who were born at term. There also appears to be an increased risk of behavioural problems, especially autistic spectrum disorder, that affects survivors of extremely preterm birth.[241] The effect on other family members is less clear, with one study reporting a negative impact on family life from extremely preterm birth,[242] while another found no effect on marriage break-up or maternal psychological factors.[243]

For society, extreme prematurity is costly, not just financially but also in terms of the time and resources allocated to caring for the survivors of extremely preterm birth.[244] Examples are the increased need for educational support [75] and the transportation requirements for additional medical appointments.[245] Financial costs to society result from the initial course of hospitalisation, and continue to accrue over the childhood (and lifetime) of those born prematurely.[78–80, 246, 247]

Moreover, changing demographics in England [238, 239] affect attitudes and opinions such as "sanctity" or "quality" of life [11] (as discussed in section 2.6.3) or about resource allocation.[11] Some decision-making was considered in the obstetric antecedents investigations, and the effects of resources (particularly, cots and staffing) were considered in the trends over time analysis. However, attitudes and opinions will have changed since the data were collected, hence further evaluation is required.

There are also broader social concerns relating to birth at periviable gestations. Abortion is legal in England below 24 weeks gestational age. This leaves a group of women – specifically, those who may deliver at 23 weeks gestation – for whom it is legal to terminate the pregnancy but whose baby may receive full intensive care if a decision has been made to resuscitate. Such an issue catches the media eye – for example, in

a documentary shown on BBC television [248] or, more recently, in some of the right wing press.[249, 250] Abortion is even more of a debate in other parts of the UK: in Northern Ireland, it is still not legal.[251] This means women are forced to travel to other countries and fund it themselves.[251] Around the world, abortion laws impact on the health of women and their babies.[252]

14.4.2 Routine data issues

The poor quality of the HES data for babies born at the edges of viability is a concern for English society. Data accuracy is important not only for record linkage [104] but also to ensure investigation results are unaffected by information bias. Completeness of coverage is required to prevent selection bias. HES data currently suffer from both of these problems with respect to babies born extremely prematurely. Given that they are primarily administrative data,[108, 110, 124–126] ethical concerns might be raised about their use for allocating resources. This arises because spending on neonatal care also has consequences on other parts of the health service. Budgets that are based on flawed data may be incorrectly estimated. With the current splintering of the NHS,[253] such issues can be difficult to identify. This challenges us as a society to improve the data and, consequently, our understanding.

14.4.3 Organisational relevance

Individual hospitals in England are increasingly being pushed into competitive practices [106, 254] – but effective implementation of managed clinical networks to cope with increased demand for neonatal care requires cooperation.[255] There is a strong history of successful centralisation (or, regionalisation) for important conditions within the NHS. This includes the development of extracorporeal membrane oxygenation (ECMO) services in the 1990s following the success of the UK ECMO trial in demonstrating improved survival for babies cared for in one of five supraregional centres that were established.[256] More recently, there have been implementations of networked care for adult stroke and myocardial infarction.[257–259] However, changes in the number of extremely premature babies being cared for may affect service provision for individual hospitals differently, depending both on the population that chooses to deliver there and on the service that is provided. Difficulties may be compounded by the fact MCNs are focused on babies rather than women.[3]

There are further related considerations that arise from the investigation of hospital trends over time. Increases seen in staffing levels are promising but hospitals remain significantly understaffed compared to the levels currently recommended in England.[18, 19, 21] Changes seen in activity between the different epochs need further exploration with other data sources if we are to understand them fully. However, they suggest there has been improved utilisation of cots over time. Importantly, effects on patients are not known for any of the findings, nor were recommendations for staffing based on strong evidence. Increases in resource utilisation cannot be sustained indefinitely. Combining findings from the different parts of the thesis showed that nurse staffing numbers had increased in comparison to the number of cots but so had the number of babies admitted, implying not only were there more staff but they were also working harder. More investigation is required to determine optimal staffing levels for best patient care.

The findings of this investigation are also generalisable beyond England. There is an increasing literature on the effects of hospital level organisational characteristics on outcomes (for example, [61, 62, 70, 87, 255, 260]). However, there is a lack of standardisation of measures between studies, making comparison difficult.[62, 87] This, along with a lack of data, has also been identified as a key issue in the Every Newborn Action Plan published by *The Lancet*.[261] There may, then, be global usage for such measures of activity and staffing.

14.4.4 Clinical importance

Direct patient relevance is seen in the results from the obstetric antecedents analysis. Like other studies,[90, 225, 226] there do not appear to be any advantages to Caesarean delivery for babies of 26 weeks gestational age or below, save in very select circumstances. Babies whose mothers received steroids antenatally experience a survival benefit over those who don't; this is strongly supported by evidence from other studies [34–36, 38, 230, 231, 262–265] and, consequently, administration should be considered routine for mothers expected to deliver between 22 and 26 completed weeks of gestation. Similarly, tocolysis appears to be associated with improved outcome at birth. However, supporting evidence is mixed, hence this finding should probably not influence practice unless it is replicated in other investigations. These findings are generalisable to populations of extremely premature babies less than 27 weeks gestational

age in other countries as well as England: many important potentially confounding factors including demographics and clinical history were accounted for in the analyses, and the EPICure 2 was very comprehensive.[55]

The caveat remains, however, that it is impossible to distinguish at presentation whether a woman will deliver prematurely or not. This may change, with broader use of diagnostic tests such as fetal fibronectin [266, 267] or measurement of cervical length.[268] Diagnosis also impacts on whether or not women require antenatal transfer to a tertiary referral centre. There is little current knowledge about this, and information is hard to collect.[269]

14.5 Implications for future research

Several questions remain unanswered and others arise from these investigations. More importantly, there are lessons to be learned that may have an impact on future studies of extreme prematurity in England.

14.5.1 Outstanding questions

- Despite the current investigation, it is not confirmed that the increases in extremely premature babies seen between 1995 and 2006 were true, nor whether there have been continued increases in births or admissions since. Thus, it is essential to again examine what the trends have been for extremely preterm births and admissions into neonatal intensive care.

- Furthermore, the impact of obstetric management in the population of women who present with threatened preterm labour is unknown. This requires investigation.

- A specific component of perinatal management about which there is currently very little data is antenatal transfer. It is currently unknown how successful policies are in achieving transfer for women to network NICUs.

- Linked to this, the performance of markers of imminent delivery such as fetal fibronectin or cervical length measurement require further assessment.[270]

- It is unclear if there are changes in attitude in relation to perinatal decision-making. A repeat study similar to the EPICure 2 cohort but with improved questions may help to disentangle this component from the effects of perinatal management.

- There is little information available about the effects of current management strategies on longer term outcomes such as neurological or other system morbidities.

- A further gap in knowledge is the relationship between staffing and activity – using measures of intensity and throughput, as presented in this thesis – on individual patient outcomes. Repeating the Unit Profile Study at the same time as an investigation of extremely premature births would enable this to be examined.

14.5.2 Population characteristics

Two additional factors relating to the collection of population data for the investigation of these questions are proposed. The first is that the gestational age range should be extended to included all women presenting at less than 28 weeks completed gestational age. This would mean all those who are currently recommended [65, 271] to be managed in network NICUs are included. Secondly, future investigations should take place over a longer time period to increase power for the lowest gestational age groups – those born at 23 and 24 weeks; a minimum of two years is suggested. This is of particular relevance for longer term follow-up, where the number of survivors may be few.

14.5.3 Data requirements

The current routine data sources that are collected in England relating to extremely premature births are:

- HES for all women and babies admitted to an NHS hospital.

- MBRRACE-UK for maternal and perinatal deaths.

- NNRD for all babies admitted into neonatal intensive care.

None of these comprehensively collect data on all births occurring at extremely low gestations; there are some overlaps, and some notable omissions. To conduct further investigations, it is therefore necessary to collect supplementary data – although the majority of data collection could be performed in collaboration with MBRRACE-UK and the NNRD.

Specifically, it is proposed that additional data should be collected on all women who are admitted into hospital, and of all births occurring, at less than 28 weeks completed gestational age. Data should be recorded about the women during their entire admission through to the point at which they deliver or are discharged home from hospital. Either of these events may occur after 28 weeks gestation. Women who are transferred with their fetus *in utero* to another hospital for ongoing care would be subject to ongoing monitoring.

For the cohort of mothers admitted to hospital (women with threatened preterm labour or requiring delivery for maternal reasons) at less than 28 weeks gestation, data could be collected via a secure, web-based portal, similar to those described for the Stroke Improvement National Audit Project (SINAP)[1] [258, 272] and the Myocardial Ischaemia National Audit Project (MINAP).[257] The software should be designed to function together with the existing MBRRACE-UK collection system for maternal and perinatal deaths and stillbirths. Identifiable information would need to be collected to facilitate data linkage with registry and/or NNRD (if the baby was admitted) data for those women who deliver in an admission subsequent to one commencing at less than 28 weeks gestation.

It is proposed that data about three groups of babies should be collected. The first group would comprise those babies born at less than 28 weeks gestation. Detailed information about these babies could be obtained from the NNRD if they are admitted onto a neonatal unit, else from MBRRACE-UK. The second group of babies would be those born at 28 weeks gestational age or beyond to mothers who continue to be hospital inpatients following their primary admission at less than 28 weeks gestation. These babies may be admitted to neonatal units, discharged directly home, or not survive; complete data should be obtained, which may be from the NNRD for those admitted, or from registry data. However, a novel data collection system in tandem

[1]This is now known as SSNAP – the Sentinal Stroke National Audit Project: `https://www.rcplondon.ac.uk/projects/ssnap-acute-organisational-audit`

with the maternal record may also be required. The final group of babies would be those born to women who were discharged home after their primary admission at less than 28 weeks gestation without having yet delivered. These babies may again be admitted to the neonatal unit, discharged directly home after birth or not survive; however, they are likely to be a more difficult group of babies to trace for data collection. Linkage would be performed using both NNRD and registry data to examine morbidity and mortality outcomes.

14.6 Overall conclusions

This thesis supported the hypothesis that the 44% increase seen in the numbers of premature babies born below 26 weeks completed gestation in England who were admitted to neonatal intensive care between 1995 and 2006 was real. It was not possible to confirm this increase with routine data, nor to determine the reasons behind it. Increases were seen in HES data of the total numbers of babies born and of those born alive. These were disproportionate to the increase seen in the total number of births at all gestational ages, and of a commensurate magnitude and direction to those seen in EPICure. Additionally, improvements in management are likely to have improved condition at birth and survival beyond the delivery room.

Specific conclusions from the individual investigations were that:

- Hospital Episode Statistics are a poor source of information about those born extremely prematurely.
- There have been no improvements over time in the coding of HES.
- Research governance mechanisms in England are complicated and may be cumbersome.
- Antenatal steroids are strongly associated with improved outcome at birth.
- Tocolysis appears to be of benefit as it was used in 2006, but this needs investigation in other populations.
- Caesarean section is not associated with improved outcome for babies born below 27 completed weeks of gestational age.

- Staffing numbers have increased more than cot numbers between 1997 and 2011, but remain significantly below recommended levels.
- There have been important changes in activity since 1997. Data up to 2006 are consistent with other studies, suggesting a greater proportion of extremely premature babies were being admitted to neonatal intensive care. Further changes were noted between 2006 and 2011 although their significance is unclear.

The overall interpretation of these findings was that there were inadequate data available from routine data and insufficient information from historical sources to guide national policy, hospital organisation or individual patient care. However, the changes noted in population trends, particularly with respect to the proportion of extremely premature babies that are born and then subsequently admitted into neonatal intensive care, were likely to have important – and probably ongoing – consequences. Suggestions were made in section 14.5 of improvements that could be made to future investigations by collecting supplementary data to those which are currently collected by MBRRACE-UK and for the NNRD.

References

[1] Goldenberg R, Culhane J, Iams J, Romero R (2008) Epidemiology and causes of preterm birth. Lancet 371: 75–84.

[2] Zeitlin J, Ancel PY (2011) Interpreting data on the health outcomes of extremely preterm babies. Archives of Disease in Childhood Fetal and Neonatal Edition 96: F314-6.

[3] Marlow N, Bryan Gill A (2007) Establishing neonatal networks: the reality. Archives of Disease in Childhood Fetal and Neonatal Edition 92: F137-42.

[4] Dunn PM (2007) The birth of perinatal medicine in the United Kingdom. Seminars in Fetal & Neonatal Medicine 12: 227–238.

[5] Kassebaum N, Bertozzi-Villa A, Coggeshall M, Shackelford K, Steiner C, et al. (2014) Global, regional, and national levels and causes of maternal mortality during 1990–2013: a systematic analysis for the global burden of disease study 2013. The Lancet May 2nd.

[6] Macfarlane A, Mugford M (2012) Epidemiology. In: Rennie J, editor, Rennie and Roberton's Textbook of Neonatology, 5^{th} *Edition*, London, UK: Churchill Livingstone Elsevier. pp. 3–51.

[7] Howson C, Kinney M, Lawn J, World Health Organization (2012) Born Too Soon: The global action report on preterm birth. Technical report, March of Dimes, Partnership for Maternal, Newborn & Child Health, Save the Children and World Health Organization., Geneva. URL `http://www.who.int/pmnch/media/news/2012/preterm_birth_report/en/index1.html`.

[8] Saigal S, Doyle L (2008) An overview of mortality and sequelae of preterm birth from infancy to adulthood. Lancet 371: 261–269.

[9] Smith L, McKay K, van Asperen P, Selvadurai H, Fitzgerald D (2010) Normal development of the lung and premature birth. Paediatric Respiratory Reviews 11: 135–142.

[10] Nkadi P, Merritt T, Pillers DA (2009) An overview of pulmonary surfactant in the neonate: genetics, metabolism, and the role of surfactant in health and disease. Molecular Genetics and Metabolism 97: 95–101.

[11] Brazier M, Archard D, Campbell A, Franck L, Green B, et al. (2006) Critical care decisions in fetal and neonatal medicine: ethical issues. London: Nuffield Council on Bioethics.

[12] Royal College of Obstetricians and Gynaecologists (2001). Further issues relating to late abortion, fetal viability and registration of births and deaths. URL http://www.rcog.org.uk/womens-health/clinical-guidance/further-issues-relating-late-abortion-fetal-viability-and-registrati.

[13] World Health Organization (2010). International Classification of Diseases and related health problems, 10th revision. URL http://apps.who.int/classifications/icd10/browse/2010/.

[14] Larroque B, Breart G, Kaminski M, Dehan M, Andre M, et al. (2004) Survival of very preterm infants: EPIPAGE, a population based cohort study. Archives of Disease in Childhood Fetal and Neonatal Edition 89: F139-44.

[15] Subramanian KS, Barton AM, Montazami S (2011) Extremely low birth weight infant. Medscape Reference December 15.

[16] Battaglia F, Lubchenco L (1967) A practical classification of newborn infants by weight and gestational age. The Journal of Pediatrics 71: 159–163.

[17] British Association of Perinatal Medicine, Neonatal Nurses Association (1992) Report of working group of the British Association of Perinatal Medicine and Neonatal Nurses Association on categories of babies requiring neonatal care. Archives of Disease in Childhood 67: 868–869.

[18] Anonymous (1996) Standards for hospitals providing neonatal intensive care. Technical report, British Association of Perinatal Medicine.

[19] British Association of Perinatal Medicine (2001) Standards for hospitals providing neonatal intensive care and categories of babies requiring neonatal care. Technical report, British Association of Perinatal Medicine. URL http://www.bapm.org/publications/documents/guidelines/hosp_standards.pdf.

[20] Department of Health (2009) Toolkit for high quality neonatal services. Technical report, Department of Health. URL http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_107845.

[21] British Association of Perinatal Medicine (2010) Service standards for hospitals providing neonatal care. Technical report, British Association of Perinatal Medicine. URL http://www.bapm.org/publications/documents/guidelines/BAPM_Standards_Final_Aug2010.pdf.

[22] British Association of Perinatal Medicine (2011) Categories of care. Technical report, British Association of Perinatal Medicine, online. URL http://www.bapm.org/publications/documents/guidelines/CatsofcarereportAug11.pdf.

[23] Cornette L, Miall L (2006) Development of regional guidelines: the way forward for neonatal networks? Archives of Disease in Childhood Fetal and Neonatal Edition 91: F318-9.

[24] Bohin S, Draper E, Field D (1996) Impact of extremely immature infants on neonatal services. Archives of Disease in Childhood Fetal and Neonatal Edition 74: F110-3.

[25] Stark A (2004) Levels of neonatal care. Pediatrics 114: 1341–1347.

[26] Milligan D (1997) Neonatal intensive care provision in the United Kingdom 1992-3. British Association of Perinatal Medicine. Archives of Disease in Childhood Fetal and Neonatal Edition 76: F197-200.

[27] The ECSURF (Economic Evaluation of Surfactant) Collaborative Study Group (1998) Limited comparability of classifications of levels of neonatal care in UK units. The ECSURF (Economic Evaluation of Surfactant) Collaborative Study Group. Archives of Disease in Childhood Fetal and Neonatal Edition 78: F179-84.

[28] Tucker J, Tarnow-Mordi W, Gould C, Parry G, Marlow N (1999) UK neonatal intensive care services in 1996. on behalf of the UK Neonatal Staffing Study Collaborative Group. Archives of Disease in Childhood Fetal and Neonatal Edition 80: F233-4.

[29] UK Neonatal Staffing Study Group (2000) A prospective evaluation of risk-adjusted outcomes of neonatal intensive care in relation to volume, staffing and workload in UK neonatal intensive care units. Technical report, British Association of Perinatal Medicine. URL http://www.bapm.org/publications/documents/general/UKNNS4bapm.pdf.

[30] Hamilton K, Redshaw M, Tarnow-Mordi W (2007) Nurse staffing in relation to risk-adjusted mortality in neonatal care. Archives of Disease in Childhood Fetal and Neonatal Edition 92: F99-F103.

[31] Parry G, Tucker J, Tarnow-Mordi W (2003) CRIB II: an update of the clinical risk index for babies score. Lancet 361: 1789–1791.

[32] Parmanum J, Field D, Rennie J, Steer P (2000) National census of availability of neonatal intensive care. British Association for Perinatal Medicine. BMJ (Clinical Research ed) 321: 727–729.

[33] Hicks J, Allen G (1999) A century of change: Trends in UK statistics since 1900. Technical report, House of Commons Library. URL http://www.parliament.uk/commons/lib/research/rp99/rp99-111.pdf.

[34] Crowley P, Chalmers I, Keirse M (1990) The effects of corticosteroid administration before preterm delivery: an overview of the evidence from controlled trials. British Journal of Obstetrics and Gynaecology 97: 11–25.

[35] Gilstrap LC, Christensen R, Clewell WH, D'Alton ME, Jr ECD, et al. (1995) Effect of corticosteroids for fetal maturation on perinatal outcomes. NIH Consensus Development Panel on the effect of corticosteroids for fetal maturation on perinatal outcomes. JAMA: the Journal of the American Medical Association 273: 413–418.

[36] Royal College of Obstetricians and Gynaecologists (2010). Antenatal corticosteroids to reduce neonatal morbidity (Green-top 7).

[37] Crowley P (2006) Withdrawn: Prophylactic corticosteroids for preterm birth. The Cochrane Database of Systematic Reviews Issue 3: Cd000065.

[38] Roberts D, Dalziel S (2006) Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. Cochrane Database of Systematic Reviews (Online) 3: Cd004454.

[39] Halliday H (2008) Surfactants: past, present and future. Journal of Perinatology: Official Journal of the California Perinatal Association 28 Suppl 1: S47-56.

[40] Jobe A (1993) Pulmonary surfactant therapy. The New England Journal of Medicine 328: 861–868.

[41] Fujiwara T, Maeta H, Chida S, Morita T, Watabe Y, et al. (1980) Artificial surfactant therapy in hyaline-membrane disease. Lancet 1: 55–59.

[42] Tovey D, The Cochrane Library (2014). About the Cochrane Library. URL `http://www.thecochranelibrary.com/view/0/AboutTheCochraneLibrary.html`.

[43] Halliday H (2005) History of surfactant from 1980. Biology of the Neonate 87: 317–322.

[44] Logan J, Moya F (2009) Animal-derived surfactants for the treatment and prevention of neonatal respiratory distress syndrome: summary of clinical trials. Therapeutics and Clinical Risk Management 5: 251–260.

[45] Speer C, Gefeller O, Groneck P, Laufkotter E, Roll C, et al. (1995) Randomised clinical trial of two treatment regimens of natural surfactant preparations in neonatal respiratory distress syndrome. Archives of Disease in Childhood Fetal and Neonatal Edition 72: F8-13.

[46] Anonymous (1996) A multicenter, randomized trial comparing synthetic surfactant with modified bovine surfactant extract in the treatment of neonatal respiratory distress syndrome. Vermont-Oxford Neonatal Network. Pediatrics 97: 1–6.

[47] Bloom B, Kattwinkel J, Hall R, Delmore P, Egan E, et al. (1997) Comparison of Infasurf (calf lung surfactant extract) to Survanta (Beractant) in the treatment and prevention of respiratory distress syndrome. Pediatrics 100: 31–38.

[48] Hudak M, Martin D, Egan E, Matteson E, Cummings N, et al. (1997) A multicenter randomized masked comparison trial of synthetic surfactant versus calf lung surfactant extract in the prevention of neonatal respiratory distress syndrome. Pediatrics 100: 39–50.

[49] Capilla D (2012) FDA reprimands competing drug companies for false and misleading statements. Wolters Kluwer Law & Business [website] November 13.

[50] Morley C (1997) Systematic review of prophylactic vs rescue surfactant. Archives of Disease in Childhood Fetal and Neonatal Edition 77: F70-74.

[51] Soll R, Özek E (1997) Prophylactic animal derived surfactant extract for preventing morbidity and mortality in preterm infants. Cochrane Database of Systematic Reviews Issue 4: CD000511.

[52] Cartlidge P, Stewart J (1995) Effect of changing the stillbirth definition on evaluation of perinatal mortality rates. Lancet 346: 486–488.

[53] Tin W, Wariyar U, Hey E (1997) Changing prognosis for babies of less than 28 weeks' gestation in the north of England between 1983 and 1994. Northern Neonatal Network. BMJ (Clinical Research ed) 314: 107–111.

[54] Costeloe K, Hennessy E, Gibson A, Marlow N, Wilkinson A (2000) The EPICure study: outcomes to discharge from hospital for infants born at the threshold of viability. Pediatrics 106: 659–671.

[55] Costeloe K, Hennessy E, Haider S, Stacey F, Marlow N, et al. (2012) Short term outcomes after extreme preterm birth in England: comparison of two birth cohorts in 1995 and 2006 (the EPICure studies). BMJ (Clinical Research ed) 345: e7976.

[56] Wood N, Marlow N, Costeloe K, Gibson A, Wilkinson A (2000) Neurologic and developmental disability after extremely preterm birth. EPICure study group. The New England Journal of Medicine 343: 378–384.

[57] Marlow N, Wolke D, Bracewell M, Samara M (2005) Neurologic and developmental disability at six years of age after extremely preterm birth. The New England Journal of Medicine 352: 9–19.

[58] Johnson S, Fawke J, Hennessy E, Rowell V, Thomas S, et al. (2009) Neurodevelopmental disability through 11 years of age in children born before 26 weeks of gestation. Pediatrics 124: e249-57.

[59] Larroque B (2000) EPIPAGE: epidemiologic study of very premature infants. Protocol of the survey. Archives de Pediatrie: Organe Officiel de la Societe Francaise de Pediatrie 7 Suppl 2: 339s-342s.

[60] Bodeau-Livinec F, Marlow N, Ancel PY, Kurinczuk J, Costeloe K, et al. (2008) Impact of intensive care practices on short-term and long-term outcomes for extremely preterm infants: comparison between the British Isles and France. Pediatrics 122: e1014-21.

[61] Zeitlin J, Papiernik E, Breart G (2004) Regionalization of perinatal care in Europe. Seminars in Neonatology: SN 9: 99–110.

[62] Lasswell S, Barfield W, Rochat R, Blackmon L (2010) Perinatal regionalization for very low-birth-weight and very preterm infants: a meta-analysis. JAMA: the Journal of the American Medical Association 304: 992–1000.

[63] International Neonatal Network, Scottish Neonatal Constultants, Nurses Collaborative Study Group (2000) Risk adjusted and population based studies of the outcome for high risk infants in Scotland and Australia. International Neonatal Network, Scottish Neonatal Consultants, Nurses Collaborative Study Group. Archives of Disease in Childhood Fetal and Neonatal Edition 82: F118-23.

[64] Field D, Petersen S, Clarke M, Draper E (2002) Extreme prematurity in the UK and Denmark: population differences in viability. Archives of Disease in Childhood Fetal and Neonatal Edition 87: F172-5.

[65] Department of Health (2003) Neonatal intensive care services: report of the Department of Health Expert Working Group – consultation. Technical report, Department of Health. URL http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4002704.

[66] Hall S (2006) Premature babies transferred to hospitals up to 280 miles away. The Guardian, Wednesday, 12 July 2006, 11.20 BST.

[67] Kyte F, McDonald T, Ormiston-Smith N, Milligan C, ONeill M, et al. (2007) Caring for Vulnerable Babies: the reorganisation of neonatal services in England. Technical report, National Audit Office, London. URL http://www.nao.org.uk/report/caring-for-vulnerable-babies-the-reorganisation-of-neonatal-services-in-england/.

[68] Costeloe [PI] K (2005). EPICure 2: A study of extreme prematurity in 2006 [protocol]. URL http://www.epicure.ac.uk/professionals.

[69] Moody J, editor (2003) Antenatal care: routine care for the healthy pregnant woman. London: RCOG Press.

[70] Marlow N, Bennett C, Draper E, Hennessy E, Morgan A, et al. (2014) Perinatal outcomes for extremely preterm babies in relation to place of birth in England: the EPICure 2 study. Archives of Disease in Childhood Fetal and Neonatal Edition 99: F181-8.

[71] British Association of Perinatal Medicine (2010) Position statement on therapeutic cooling for neonatal encephalopathy. Technical report, British Association of Perinatal Medicine. URL http://www.bapm.org/publications/documents/guidelines/Position_Statement_Therapeutic_Cooling_Neonatal_Encephalopathy_July%202010.pdf.

[72] Azzopardi DV, Strohm B, Edwards AD, Dyet L, Halliday HL, et al. (2009) Moderate hypothermia to treat perinatal asphyxial encephalopathy. The New England Journal of Medicine 361: 1349–1358.

[73] Wood N, Costeloe K, Gibson A, Hennessy E, Marlow N, et al. (2003) The EPICure study: growth and associated problems in children born at 25 weeks of gestational age or less. Archives of Disease in Childhood – Fetal and Neonatal Edition 88: f492-f500.

[74] Bracewell MA, Hennessy EM, Wolke D, Marlow N (2008) The EPICure study: growth and blood pressure at 6 years of age following extremely preterm birth. Archives of Disease in Childhood Fetal and Neonatal Edition 93: F108-14.

[75] Johnson S, Hennessy E, Smith R, Trikic R, Wolke D, et al. (2009) Academic attainment and special educational needs in extremely preterm children at 11 years of age: the EPICure study. Archives of Disease in Childhood Fetal and Neonatal Edition 94: F283-9.

[76] Fawke J, Lum S, Kirkby J, Hennessy E, Marlow N, et al. (2010) Lung function and respiratory symptoms at 11 years in children born extremely preterm: the EPICure study. American Journal of Respiratory and Critical Care Medicine 182: 237–245.

[77] McEniery C, Bolton C, Fawke J, Hennessy E, Stocks J, et al. (2011) Cardiovascular consequences of extreme prematurity: the EPICure study. Journal of Hypertension : 1.

[78] Petrou S, Henderson J, Bracewell M, Hockley C, Wolke D, et al. (2006) Pushing the boundaries of viability: the economic impact of extreme preterm birth. Early Human Development 82: 77–84.

[79] Mangham L, Petrou S, Doyle L, Draper E, Marlow N (2009) The cost of preterm birth throughout childhood in England and Wales. Pediatrics 123: e312-27.

[80] Petrou S, Abangma G, Johnson S, Wolke D, Marlow N (2009) Costs and health utilities associated with extremely preterm birth: evidence from the EPICure study. Value in Health: the Journal of the International Society for Pharmacoeconomics and Outcomes Research 12: 1124–1134.

[81] Arora S, Charlesworth A, Kelly E, Stoye G (2013) Public payment and private provision: the changing landscape of health care in the 2000s. Technical report, Nuffield Trust; Institute for Fiscal Studies, online. URL http://www.nuffieldtrust.org.uk/publications/public-payment-private-provision-2000s.

[82] Office for National Statistics (2008) Review of the National Statistician on births and patterns of family building in England and Wales, 2008. Technical report, Office for National Statistics. URL http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-226270.

[83] Appleby J (2013) The hospital bed: on its way out? BMJ (Clinical Research ed) 346: f1563.

[84] Phibbs CS (2012) Managed clinical networks in neonatal care. BMJ (Clinical Research ed) 344: e2423.

[85] Gale C, Santhakumaran S, Nagarajan S, Statnikov Y, Modi N (2012) Impact of managed clinical networks on NHS specialist neonatal services in England: population based study. BMJ (Clinical Research ed) 344: e2105.

[86] Watson SI, Arulampalam W, Petrou S, Marlow N, Morgan AS, et al. (2014) 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 4: e004856.

[87] Lorch SA, Myers S, Carr B (2010) The regionalization of pediatric health care. Pediatrics 126: 1182–1190.

[88] Greenland S (2000) An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology 29: 722–729.

[89] Hallsworth M, Farrands A, Oortwijn WJ, Hatziandreu E (2007) The provision of neonatal services: data for international comparisons. Technical report, RAND Corporation. URL http://www.nao.org.uk/wp-content/uploads/2007/12/0708101_International_Comparisons.pdf.

[90] Fellman V, Hellstrom-Westas L, Norman M, Westgren M, Kallen K, et al. (2009) One-year survival of extremely preterm infants after active perinatal care in Sweden. JAMA: the Journal of the American Medical Association 301: 2225–2233.

[91] The EXPRESS Group (2010) Incidence of and risk factors for neonatal morbidity after active perinatal care: extremely preterm infants study in Sweden (EXPRESS). Acta Paediatrica (Oslo, Norway: 1992) 99: 978–992.

[92] Serenius F, Kallen K, Blennow M, Ewald U, Fellman V, et al. (2013) Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden. JAMA: the Journal of the American Medical Association 309: 1810–1820.

[93] Markestad T, Kaaresen PI, Ronnestad A, Reigstad H, Lossius K, et al. (2005) Early death, morbidity, and need of treatment among extremely premature infants. Pediatrics 115: 1289–1298.

[94] Leversen KT, Sommerfelt K, Ronnestad A, Kaaresen PI, Farstad T, et al. (2011) Prediction of neurodevelopmental and sensory outcome at 5 years in Norwegian children born extremely preterm. Pediatrics 127: e630-8.

[95] Leversen KT, Sommerfelt K, Elgen IB, Eide GE, Irgens LM, et al. (2012) Prediction of outcome at 5 years from assessments at 2 years among extremely preterm children: a Norwegian national cohort study. Acta Paediatrica (Oslo, Norway: 1992) 101: 264–270.

[96] Keller M, Felderhoff-Mueser U, Lagercrantz H, Dammann O, Marlow N, et al. (2010) Policy benchmarking report on neonatal health and social policies in 13 European countries. Acta Paediatrica (Oslo, Norway: 1992) 99: 1624–1629.

[97] Gallagher K, Martin J, Keller M, Marlow N (2014) European variation in decision-making and parental involvement during preterm birth. Archives of Disease in Childhood Fetal and Neonatal Edition .

[98] Chan K, Kean L, Marlow N (2006) Staff views on the management of the extremely preterm infant. European Journal of Obstetrics, Gynecology, and Reproductive Biology 128: 142–147.

[99] Draper E, Manktelow B, Field D, James D (1999) Prediction of survival for preterm births by weight and gestational age: retrospective population based study. BMJ (Clinical Research ed) 319: 1093–1097.

[100] Salathe M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, et al. (2012) Digital epidemiology. PLoS Computational Biology 8: e1002616.

[101] Khoury MJ, Lam TK, Ioannidis JPA, Hartge P, Spitz MR, et al. (2013) Transforming epidemiology for 21st century medicine and public health. Cancer Epidemiology, Biomarkers & Prevention: a Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology 22: 508–516.

[102] Doctorow C (2008). Big data: Welcome to the petacentre. doi:10.1038/455016a.

[103] Anonymous (2014) The NSA files. Online .

[104] Hagger-Johnson G, Harron K, Goldstein H, Parslow R, Dattani N, et al. (2014) Making a hash of data: what risks to privacy does the NHS's care.data scheme pose? BMJ (Clinical Research ed) 348: g2264.

[105] Moser K, Macfarlane A, Chow Y, Hilder L, Dattani N (2007) Introducing new data on gestation-specific infant mortality among babies born in 2005 in England and Wales. Health Statistics Quarterly / Office for National Statistics 35: 13–27.

[106] Pollock AM, Macfarlane A, Godden S (2012) Dismantling the signposts to public health? NHS data under the Health and Social Care Act 2012. BMJ (Clinical Research ed) 344: e2364.

[107] Health and Social Care Information Centre (2014). Who we are and what we do. URL http://www.hscic.gov.uk/whoweare.

[108] United Kingdom of Great Britain and Northern Ireland Government (2011). HESOnline: About the NHS Information Centre. URL http://webarchive.nationalarchives.gov.uk/tna_la_nhs/20110926161752/http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=981.

[109] Appleby J (2014) Care.data – your bits in their hands. BMJ Blogs 13 February.

[110] Spencer SA (2011) Hospital Episode Statistics (HES): Improving the quality and value of hospital data (a discussion document). Technical report, The Health and Social Care Information Centre. URL http://www.aomrc.org.uk/doc_view/9379-hospital-episode-statistics-improving-the-quality-and-value-of-hospital-data-discussion-document.

[111] Hilder L, Moser K, Dattani N, Macfarlane A (2007) Pilot linkage of NHS Numbers for Babies data with birth registrations. Health Statistics Quarterly / Office for National Statistics 33: 25–33.

[112] Watson M, Davis K, Oleka G, Gray D, Fitz-Simon N (2012) Annual report, National Neonatal Audit Programme 2012. Extended on line version: August 2013. Technical report, Royal College of Paediatrics and Child Health. URL http://www.rcpch.ac.uk/system/files/protected/page/NNAP%20Annual%20Report%202012%20-%20Extended%20Version%20Final%20Version_1.pdf.

[113] Neonatal Data Analysis Unit, Imperial College London (2013). About the National Neonatal Research Database. URL http://www1.imperial.ac.uk/departmentofmedicine/divisions/infectiousdiseases/paediatrics/neonatalmedicine/ndau/data/nnrd/.

[114] CloudStore, Crown copyright (2013). The BadgerNet Platform. URL http://govstore.service.gov.uk/cloudstore/the-badgernet-platform4-g4-1071-001.

[115] Clevermed Ltd (2013). BadgerNet pricing template. URL http://assets-production.govstore.service.gov.uk/G4/Clevermed_Ltd-1071/523dc07b354067a4241d1d6b/QD4/Pricing%20Template.pdf.

[116] Weindling A (2003) The confidential enquiry into maternal and child health (CEMACH). Archives of Disease in Childhood 88: 1034–1037.

[117] Anonymous (2012) MBRRACE-UK appointed to conduct the National Maternal, Newborn and Infant Review Programme. Online 13 June.

[118] HM Government (1953). Births and Deaths Registration Act. URL http://www.legislation.gov.uk/ukpga/Eliz2/1-2/20.

[119] Alberman E, Blatchley N, Botting B, Schuman J, Dunn A (1997) Medical causes on stillbirth certificates in England and Wales: distribution and results of hierarchical classifications tested by the Office for National Statistics. BJOG: An International Journal of Obstetrics and Gynaecology 104: 1043–1049.

[120] Maher J, Macfarlane A (2004) Trends in live births and birthweight by social class, marital status and mother's age, 1976-2000. Health Statistics Quarterly / Office for National Statistics 23: 34–42.

[121] Collingwood Bakeo A (2004) Trends in live births by mother's country of birth and other factors affecting low birthweight in England and Wales, 1983-2001. Health Statistics Quarterly / Office for National Statistics 23: 25–33.

[122] The NHS Information Centre for health and social care (2014). What HES data are available? URL http://www.hscic.gov.uk/hes.

[123] The NHS Information Centre for health and social care (2010) Hospital Episode Statistics: Inpatient Data Dictionary. Health & Social Care Information Centre.

[124] Abrahams C, Davy K (2002) Linking HES maternity records with ONS birth records. Health Statistics Quarterly 13: 22–30.

[125] Spencer SA, Davies MP (2012) Hospital Episode Statistics: improving the quality and value of hospital data: a national internet e-survey of hospital consultants. BMJ Open 2.

[126] Anonymous (2011) Hospital Episode Statistics: How good is HES ethnic coding and where do the problems lie? Technical report, The Health and Social Care Information Centre, Online.

[127] Knight H, Gurol-Urganci I, Mahmood T, Templeton A, Richmond D, et al. (2013) Evaluating maternity care using national administrative health datasets: how are statistics affected by the quality of data on method of delivery? BMC Health Services Research 13: 200.

[128] Statnikov Y, Santhakumaran S, Nagarajan S, Modi N (2011) NDAU 2010 report. Technical report, Neonatal Data Analysis Unit. URL http://www1.imperial.ac.

uk/resources/4882CFDF-2054-42AF-86D9-57EC957664E9/complete2010reportver2.pdf.

[129] Neonatal Data Analysis Unit (2012) London Neonatal Networks 2011 Report. Technical report, Neonatal Data Analysis Unit. URL http://www1.imperial.ac.uk/resources/D67972D4-5ADE-4DE4-AE14-4C0FBF69553C/londonreport2011final.pdf.

[130] Horbar JD (1999) The Vermont Oxford Network: evidence-based quality improvement for neonatology. Pediatrics 103: 350–359.

[131] Oakley L, Maconochie N, Doyle P, Dattani N, Moser K (2009) Multivariate analysis of infant death in England and Wales in 2005-06, with focus on socio-economic status and deprivation. Health Statistics Quarterly / Office for National Statistics 42: 22–39.

[132] Dattani N, Datta-Nemdharry P, Macfarlane A (2011) Linking maternity data for England, 2005-06: methods and data quality. Health Statistics Quarterly / Office for National Statistics 49: 53–79.

[133] Dattani N, Datta-Nemdharry P, Macfarlane A (2012) Linking maternity data for England 2007: methods and data quality. Health Statistics Quarterly / Office for National Statistics 53: 4–21.

[134] Office for National Statistics (2014). Index page: Health Statistics Quarterly, no 49, spring 2011. URL http://www.ons.gov.uk/ons/rel/hsq/health-statistics-quarterly/spring-2011/index.html.

[135] Bohensky MA, Jolley D, Sundararajan V, Evans S, Pilcher DV, et al. (2010) Data linkage: a powerful research tool with potential problems. BMC Health Services Research 10: 346.

[136] Clark DE (2004) Practical introduction to record linkage for injury research. Injury Prevention: Journal of the International Society for Child and Adolescent Injury Prevention 10: 186–191.

[137] Sariyer M, Borg A (2010) The recordlinkage package: Detecting errors in data. The R Journal 2: 61.

[138] Harron K, MRC Centre for Epidemiology of Child Health, UCL Institute of Child Health (2013) Introduction to Data Linkage: course notes. London: UCL.

[139] Ordnance Survey, HM Government (2014). Code-Point® Open. URL http://data.gov.uk/dataset/code-point-open.

[140] Snae C (2007) A comparison and analysis of name matching algorithms. International Journal of Applied Science Engineering and Technology 4: 252–257.

[141] Fellegi IP, Sunter AB (1969) A theory for record linkage. Journal of the American Statistical Association 64: 1183.

[142] Contiero P, Tittarelli A, Maghini A, Fabiano S, Frassoldi E, et al. (2008) Comparison with manual registration reveals satisfactory completeness and efficiency of a computerized cancer registration system. Journal of Biomedical Informatics 41: 24–32.

[143] Borg A, Sariyar M (2012) Package RecordLinkage. URL `https://r-forge.r-project.org/projects/recordlinkage/`.

[144] Winkler WE (2000) Using the EM algorithm for weight computation in the Fellegi-Sunter model of record linkage. In: Association AS, editor, Proceedings of the Section on Survey Research Methods. American Statistical Association, pp. 667–671. doi:10.1.1.17.175.

[145] Silveira DPd, Artmann E (2009) Acurácia em métodos de relacionamento probabilístico de bases de dados em saúde: revisão sistemática. Revista de Saúde Pública 43: 875–882.

[146] Bohensky MA, Jolley D, Sundararajan V, Evans S, Ibrahim J, et al. (2011) Development and validation of reporting guidelines for studies involving data linkage. Australian and New Zealand Journal of Public Health 35: 486–489.

[147] Kendall JM (2003) Designing a research project: randomised controlled trials and their principles. Emergency Medicine Journal: EMJ 20: 164–168.

[148] Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, et al. (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Medicine 4: e297.

[149] Schulz KF, Altman DG, Moher D (2010) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. British Medical Journal (Clinical Research ed) 340: c332.

[150] Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Medicine 6: e1000097.

[151] Hernan M, Hernandez-Diaz S, Werler M, Mitchell A (2002) Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. American Journal of Epidemiology 155: 176–184.

[152] Greenland S, Pearl J, Robins JM (1999) Causal diagrams for epidemiologic research. Epidemiology (Cambridge, Mass) 10: 37–48.

[153] Altman DG, Gore SM, Gardner MJ, Pocock SJ (1983) Statistical guidelines for contributors to medical journals. British Medical Journal (Clinical Research ed) 286: 1489–1493.

[154] Sterne JA, Davey Smith G (2001) Sifting the evidence-what's wrong with significance tests? BMJ (Clinical Research ed) 322: 226–231.

[155] Gardner MJ, Altman DG (1986) Confidence intervals rather than p values: estimation rather than hypothesis testing. British Medical Journal (Clinical Research ed) 292: 746–750.

[156] Goodman S (2003) Commentary: The p-value, devalued. International Journal of Epidemiology 32: 699–702.

[157] Kirkwood BR, Sterne JAC, Blackwell Science Ltd (2003) Essential Medical Statistics (2nd ed.). Blackwell Publishing.

[158] Sedgwick P (2014) Pitfalls of statistical hypothesis testing: type I and type II errors. BMJ (Clinical Research ed) 349: g4287.

[159] Perneger TV (1998) What's wrong with Bonferroni adjustments. BMJ (Clinical Research Ed) 316: 1236–1238.

[160] Bennett CM, Baird AA, Miller MB, Wolford GL (2009) Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction. In: the 15th Annual Meeting of the Organization for Human Brain Mapping. URL `http://prefrontal.org/files/posters/Bennett-Salmon-2009.pdf`.

[161] Aickin M (1999) Other method for adjustment of multiple testing exists. BMJ (Clinical Research Ed) 318: 127.

[162] Bender R, Lange S (1999) Multiple test procedures other than bonferroni's deserve wider use. British Medical Journal (Clinical Research ed) 318: 600.

[163] Gluud LL (2006) Bias in clinical intervention research. American Journal of Epidemiology 163: 493–501.

[164] Sanderson S, Tatt ID, Higgins JPT (2007) Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. International Journal of Epidemiology 36: 666–676.

[165] Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, et al. (2004) Systems for grading the quality of evidence and the strength of recommendations i: critical appraisal of existing approaches the GRADE Working Group. BMC Health Services Research 4: 38.

[166] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, et al. (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ (Clinical Research ed) 336: 924–926.

[167] Dubben HH, Beck-Bornholdt HP (2005) Systematic review of publication bias in studies on publication bias. BMJ (Clinical Research ed) 331: 433–434.

[168] Goldstein H, Spiegelhalter DJ (1996) League tables and their limitations: Statistical issues in comparisons of institutional performance. Journal of the Royal Statistical Society Series A (Statistics in Society) 159: 385.

[169] Horner RD, Szaflarski JP, Jacobson CJ, Elder N, Bolon S, et al. (2011) Clinical work intensity among physician specialties: how might we assess it? what do we find? Medical Care 49: 108–113.

[170] Horner RD, Szaflarski JP, Ying J, Meganathan K, Matthews G, et al. (2011) Physician work intensity among medical specialties: emerging evidence on its magnitude and composition. Medical Care 49: 1007–1011.

[171] Horner R, Matthews G, Yi M (2012) A conceptual model of physician work intensity: guidance for evaluating policies and practices to improve health care delivery. Medical Care 50: 654–661.

[172] Milligan D, Carruthers P, Mackley B, Ward Platt M, Collingwood Y, et al. (2008) Nursing workload in UK tertiary neonatal units. Archives of Disease in Childhood 93: 1059–1064.

[173] Pillay T, Nightingale P, Owen S, Kirby D, Spencer A (2012) Neonatal nurse staffing and delivery of clinical care in the SSBC Newborn Network. Archives of Disease in Childhood Fetal and Neonatal Edition 97: F174-8.

[174] Sherenian M, Profit J, Schmidt B, Suh S, Xiao R, et al. (2013) Nurse-to-patient ratios and neonatal outcomes: a brief systematic review. Neonatology 104: 179–183.

[175] Wilson S, Bremner A, Hauck Y, Finn J (2011) The effect of nurse staffing on clinical outcomes of children in hospital: a systematic review. International Journal of Evidence-Based Healthcare 9: 97–121.

[176] Watkin SL (2005) Managing safe staffing. Seminars in Fetal & Neonatal Medicine 10: 91–98.

[177] Subirana M, Long A, Greenhalgh J, Firth J (2014) A realist logic model of the links between nurse staffing and the outcomes of nursing. Journal of Research in Nursing 19: 8–23.

[178] Rittel HWJ, Webber MM (1973) Dilemmas in a general theory of planning. Policy Sciences 4: 155–169.

[179] Little J (1961) A proof for the queuing formula:$L = \lambda W$. Operations Research 9: 383–387.

[180] EPICure Research Group (2011). EPICure@16 documents. URL `http://epicure.ac.uk/publications/epicure-16-documents/`.

[181] EPICure Research Group (2014). EPICure@19. URL `http://epicure.ac.uk/epicure-1995/epicure19/`.

[182] StataCorp LP (1996). Stata. URL http://www.stata.com.

[183] Costeloe K, editor (2006) EPICure 2 study documents, EPICure Research Group, chapter EPIcure 2 Labour Ward Flowchart. URL http://epicure.ac.uk/publications/study-documents-for-professionals/. Exported from refbase (https://refbase.redetoile.net/show.php?record=1901).

[184] ObjectPlanet AS Inc (2011). Opinio. URL http://www.objectplanet.com/opinio/changelog.html.

[185] Oracle Corporation. Mysql. URL http://dev.mysql.com/downloads/mysql/5.5.html.

[186] R Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

[187] NIGB Ethics and Confidentiality Committee, National Information Governance Board for Health and Social Care (2013). Section 251. URL http://webarchive.nationalarchives.gov.uk/20130513181011/http://www.nigb.nhs.uk/s251.

[188] NIGB Ethics and Confidentiality Committee, National Information Governance Board for Health and Social Care (2012). Proportionate review process. URL http://webarchive.nationalarchives.gov.uk/20130513181011/http://www.nigb.nhs.uk/s251/howtoapply/Proportionate-Review-Process-July2012.pdf.

[189] Gutmann P (1996) Secure deletion of data from magnetic and solid-state memory. In: Proceedings of Sixth USENIX Security Symposium. volume 14 of *Proceedings of Sixth USENIX Security Symposium*, p. 7790. URL https://www.usenix.org/legacy/publications/library/proceedings/sec96/full_papers/gutmann/.

[190] National Information Governance Board for Health and Social Care (2012). NIGB website homepage. URL http://webarchive.nationalarchives.gov.uk/20130513181011/http://www.nigb.nhs.uk/.

[191] Service NRE, Health Research Authority (2009). Defining research. URL http://www.nres.nhs.uk/EasySiteWeb/GatewayLink.aspx?alId=355.

[192] Saout J. Cryptsetup. URL https://code.google.com/p/cryptsetup/.

[193] Software in the Public Interest, Inc and others (1997). Debian. URL http://www.debian.org.

[194] Office for National Statistics, Office for National Statistics (2014). Vital statistics: Population and health reference tables (annual data). URL http://www.ons.gov.uk/ons/rel/vsob1/vital-statistics--population-and-health-reference-tables/spring-2014-update/annual-table.xls.

[195] Brocklehurst P, Hardy P, Hollowell J, Linsell L, Macfarlane A, et al. (2011) Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study. BMJ (Clinical Research ed) 343: d7400.

[196] Ripley BD, Hornik K (2001) Date-time classes. R News 1/2: 8–11.

[197] Eddelbuettel D (2014). CRAN Task View: High-performance and parallel computing with R. URL `http://cran.r-project.org/web/views/HighPerformanceComputing.html`.

[198] Adler D, Gläser C, Nenadic O, Oehlschlägel J, Zucchini W (2013) Package 'ff'. URL `http://cran.r-project.org/web/packages/ff/index.html`.

[199] Contiero P, Tittarelli A, Tagliabue G, Maghini A, Fabiano S, et al. (2005) The EpiLink record linkage software: presentation and results of linkage test on cancer registry files. Methods of Information in Medicine 44: 66–71.

[200] van Hauser vh@thcorg (1999-2004) `wipe`. In: `secure-delete`, http://packages.qa.debian.org/s/secure-delete.html, Last updated: `Fri, 05 Jul 2013`: Debian. URL `http://packages.qa.debian.org/s/secure-delete.html`.

[201] Macfarlane A (1998) At last – maternity statistics for England. BMJ (Clinical Research ed) 316: 566–567.

[202] Murray J, Saxena S, Modi N, Majeed A, Aylin P, et al. (2013) Quality of routine hospital birth records and the feasibility of their use for creating birth cohorts. Journal of Public Health (Oxford, England) 35: 298–307.

[203] Brennan L, Watson M, Klaber R, Charles T (2012) The importance of knowing context of hospital episode statistics when reconfiguring the NHS. BMJ (Clinical Research ed) 344: e2432.

[204] Barer D, Cassidy T (2014) Effects of diagnostic uncertainty and misclassification on hospital performance indicators for acute stroke care. Clinical Medicine (London, England) 14: 597–600.

[205] Hagger-Johnson G, Harron K, Gonzalez-Izquierdo A, Cortina-Borja M, Dattani N, et al. (2014) Identifying possible false matches in anonymized hospital administrative data without patient identifiers. Health Services Research .

[206] Quigley MA, Hockley C, Davidson LL (2007) Agreement between hospital records and maternal recall of mode of delivery: evidence from 12 391 deliveries in the UK Millennium Cohort Study. BJOG : an International Journal of Obstetrics and Gynaecology 114: 195–200.

[207] Bentley JP, Ford JB, Taylor LK, Irvine KA, Roberts CL (2012) Investigating linkage rates among probabilistically linked birth and hospitalization records. BMC Medical Research Methodology 12: 149.

[208] Ford JB, Roberts CL, Taylor LK (2006) Characteristics of unmatched maternal and baby records in linked birth records and hospital discharge data. Paediatric and Perinatal Epidemiology 20: 329–337.

[209] Office for National Statistics (2013) Births in England and Wales by characteristics of birth 2, 2012. Technical report, Office for National Statistics. URL http://www.ons.gov.uk/ons/dcp171778_338077.pdf.

[210] Newcombe RG (1998) Interval estimation for the difference between independent proportions: comparison of eleven methods. Statistics in Medicine 17: 873–890.

[211] Moser K, Macfarlane A, Dattani N (2008) Survival rates in very preterm babies in England and Wales. Lancet 371: 897–898.

[212] Dattani N, Datta-Nemdharry P, Macfarlane A, Thomas G (2012) Linking maternity data for Wales, 2005-07: methods and data quality. Health Statistics Quarterly / Office for National Statistics 54: 1–24.

[213] Snooks H, Hutchings H, Seagrove A, Stewart-Brown S, Williams J, et al. (2012) Bureaucracy stifles medical research in Britain: a tale of three trials. BMC Medical Research Methodology 12: 122.

[214] van Teijlingen ER, Douglas F, Torrance N (2008) Clinical governance and research ethics as barriers to uk low-risk population-based health research? BMC Public Health 8: 396.

[215] Walshe K (2010) Reorganisation of the NHS in England. BMJ (Clinical Research ed) 341: c3843.

[216] Alfirevic Z (2012) Tocolytics: do they actually work? BMJ (Clinical Research ed) 345: e6531.

[217] World Health Organisation (2014). Perinatal mortality rate specification. URL http://www.who.int/ceh/indicators/perinatalmortality.pdf.

[218] Laing I (2014) Where should extreme preterm babies be delivered? Crucial data from EPICure. Archives of Disease in Childhood Fetal and Neonatal Edition 99: F177-8.

[219] R Core Team, R Foundation for Statistical Computing (2013) R: A Language and Environment for Statistical Computing. URL http://www.R-project.org/.

[220] Hernan MA, Hernandez-Diaz S, Robins JM (2004) A structural approach to selection bias. Epidemiology (Cambridge, Mass) 15: 615–625.

[221] Austin PC (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behavioral Research 46: 399–424.

[222] Rubin DB (2007) The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. Statistics in Medicine 26: 20–36.

[223] Tita A, Andrews W (2010) Diagnosis and management of clinical chorioamnionitis. Clinics in Perinatology 37: 339–354.

[224] Hill AB (1965) The environment and disease: Association or causation? Proceedings of the Royal Society of Medicine 58: 295–300.

[225] Alfirevic Z, Milan SJ, Livio S (2013) Caesarean section versus vaginal delivery for preterm birth in singletons. The Cochrane Database of Systematic Reviews 9: Cd000078.

[226] Malloy MH (2008) Impact of cesarean section on neonatal mortality rates among very preterm infants in the United States, 2000-2003. Pediatrics 122: 285–292.

[227] Haas DM, Caldwell DM, Kirkpatrick P, McIntosh JJ, Welton NJ (2012) Tocolytic therapy for preterm delivery: systematic review and network meta-analysis. BMJ (Clinical Research ed) 345: e6226.

[228] Royal College of Obstetricians and Gynaecologists (2014) Perinatal management of pregnant women at the threshold of infant viability (the obstetric perspective) (SIP 41). Technical report, Royal College of Obstetricians and Gynaecologists. URL `http://www.rcog.org.uk/files/rcog-corp/SIPPretermbirth.pdf`.

[229] Royal College of Obstetricians and Gynaecologists (2011) Preterm labour, tocolytic drugs (green-top 1b). [internet] .

[230] Foix-L'Helias L, Marret S, Ancel PY, Marchand L, Arnaud C, et al. (2008) Impact of the use of antenatal corticosteroids on mortality, cerebral lesions and 5-year neurodevelopmental outcomes of very preterm infants: the EPIPAGE cohort study. BJOG: an International Journal of Obstetrics and Gynaecology 115: 275–282.

[231] Carlo W, McDonald S, Fanaroff A, Vohr B, Stoll B, et al. (2011) Association of antenatal corticosteroids with mortality and neurodevelopmental outcomes among infants born at 22 to 25 weeks' gestation. JAMA: the Journal of the American Medical Association 306: 2348–2358.

[232] Redshaw M, Hamilton K (2006) Neonatal networks, admissions and transfers: the perspectives of networks, neonatal units and parents. Technical report, National Perinatal Epidemiology Unit. URL `https://www.npeu.ox.ac.uk/newneonatalsurvey`.

[233] Williams S, Whelan A, Weindling A, Cooke R (1993) Nursing staff requirements for neonatal intensive care. Archives of Disease in Childhood 68: 534–538.

[234] Tarnow-Mordi WO, Tucker JS, McCabe CJ, Nicolson P, Parry GJ (1997) The UK neonatal staffing study: A prospective evaluation of neonatal intensive care in the UK. Seminars in Neonatology 2: 171–179.

[235] Royal College of Paediatrics and Child Health (2013) Rcpch medical workforce census 2011. Technical report, Royal College of Paediatrics and Child Health. URL http://www.rcpch.ac.uk/system/files/protected/page/RCPCH%20census%20FINAL_0.pdf.

[236] Datta-Nemdharry P, Dattani N, Macfarlane A (2012) Birth outcomes for African and Caribbean babies in England and Wales: retrospective analysis of routinely collected data. BMJ Open 2.

[237] Anonymous, Free Software Foundation, Inc (1996). What is free software? The Free Software Definition, Version 1.135. URL https://www.gnu.org/philosophy/free-sw.html.

[238] Coleman D (2010) Projections of the ethnic minority populations of the United Kingdom 2006-2056. Population and Development Review 36: 441–486.

[239] Imison C (2012) Future trends: an overview. Technical report, The Kings Fund. URL http://www.kingsfund.org.uk/time-to-think-differently/publications/future-trends-overview.

[240] Tyson J, Parikh N, Langer J, Green C, Higgins R (2008) Intensive care for extreme prematurity–moving beyond gestational age. The New England Journal of Medicine 358: 1672–1681.

[241] Johnson S, Marlow N (2011) Preterm birth and childhood psychiatric disorders. Pediatric Research 69: 11r-8r.

[242] Drotar D, Hack M, Taylor G, Schluchter M, Andreias L, et al. (2006) The impact of extremely low birth weight on the families of school-aged children. Pediatrics 117: 2006–2013.

[243] Saigal S, Pinelli J, Streiner DL, Boyle M, Stoskopf B (2010) Impact of extreme prematurity on family functioning and maternal health 20 years later. Pediatrics 126: e81-8.

[244] Petrou S, Eddama O, Mangham L (2011) A structured review of the recent literature on the economic consequences of preterm birth. Archives of Disease in Childhood Fetal and Neonatal Edition 96: F225-32.

[245] McCormick MC, Bernbaum JC, Eisenberg JM, Kustra SL, Finnegan E (1991) Costs incurred by parents of very low birth weight infants after the initial neonatal hospitalization. Pediatrics 88: 533–541.

[246] Petrou S (2003) Economic consequences of preterm birth and low birthweight. BJOG: an International Journal of Obstetrics and Gynaecology 110 Suppl 20: 17–23.

[247] Petrou S (2005) The economic consequences of preterm birth during the first 10 years of life. BJOG: an International Journal of Obstetrics and Gynaecology 112 Suppl 1: 10–15.

[248] Wishart A (Wed, 2nd March 2011) 23 Week Babies: The Price of Life. BBC2. URL http://www.bbc.co.uk/programmes/p00fdm8b.

[249] Templeton SK (31 August 2014) Abortion reform call as record number of babies survive birth at 23 weeks. The Sunday Times. URL http://www.thesundaytimes.co.uk/sto/news/uk_news/Health/article1453445.ece.

[250] Spencer B (31 August 2014) Abortion reform call as more babies born at 23 weeks survive... such as little Lily who defied doctors' predictions. The Daily Mail. URL http://www.dailymail.co.uk/news/article-2739278/Abortion-reform-call-babies-born-23-weeks-survive-little-Lily-defied-doctors-predictions.html.

[251] Furedi A (2014) The UK abortion anomaly that can no longer be ignored. BMJ (Clinical Research ed) 348: g3610.

[252] Winikoff B, Sheldon WR (2012) Abortion: what is the problem? The Lancet 379: 594–596.

[253] Gerada C (2014) Something is profoundly wrong with the NHS today. 16 June. BMJ Careers. URL http://careers.bmj.com/careers/advice/view-article.html?id=20018022.

[254] Reynolds L, McKee M (2012) Opening the oyster: the 2010-11 NHS reforms in England. Clinical Medicine (London, England) 12: 128–132.

[255] Poets C (2014) Perinatal regionalisation in the UK: an international perspective. Archives of Disease in Childhood Fetal and Neonatal Edition 99: F176.

[256] Bennett CC, Johnson A, Field DJ, Elbourne D (2001) UK collaborative randomised trial of neonatal extracorporeal membrane oxygenation: follow-up to age 4 years. Lancet 357: 1094–1096.

[257] Herrett E, Smeeth L, Walker L, Weston C (2010) The Myocardial Ischaemia National Audit Project (MINAP). Heart (British Cardiac Society) 96: 1264–1267.

[258] Bray BD, Ayis S, Campbell J, Hoffman A, Roughton M, et al. (2013) Associations between the organisation of stroke services, process of care, and mortality in England: prospective cohort study. BMJ (Clinical Research ed) 346: f2827.

[259] Intercollegiate Stroke Working Party (2012) National clinical guideline for stroke, 4th edition. London: Royal College of Physicians. URL http://www.emahsn.ac.uk/emahsn/documents/stroke-idocumentroyalcollegeofphysiciansnationalclinicalguidelineforstroke.pdf.

[260] Rochefort C, Lamothe L (2011) Forcing the system: a configuration analysis of a regionalized neonatal-perinatal health network. Health Care Management Review 36: 241–251.

[261] Mason E, McDougall L, Lawn JE, Gupta A, Claeson M, et al. (2014) From evidence to action to deliver a healthy start for the next generation. Lancet 384: 455–467.

[262] Been J, Degraeuwe P, Kramer B, Zimmermann L (2011) Antenatal steroids and neonatal outcome after chorioamnionitis: a meta-analysis. BJOG: an International Journal of Obstetrics and Gynaecology 118: 113–122.

[263] Kollee L, Cuttini M, Delmas D, Papiernik E, den Ouden A, et al. (2009) Obstetric interventions for babies born before 28 weeks of gestation in Europe: results of the MOSAIC study. BJOG: an International Journal of Obstetrics and Gynaecology 116: 1481–1491.

[264] Sotiriadis A, Makrydimas G, Papatheodorou S, Ioannidis J (2009) Corticosteroids for preventing neonatal respiratory morbidity after elective caesarean section at term. Cochrane Database of Systematic Reviews (Online) Issue 4: Cd006614.

[265] Brownfoot F, Crowther C, Middleton P (2008) Different corticosteroids and regimens for accelerating fetal lung maturation for women at risk of preterm birth. Cochrane Database of Systematic Reviews (Online) Issue 4: Cd006764.

[266] Berghella V, Hayes E, Visintine J, Baxter JK (2008) Fetal fibronectin testing for reducing the risk of preterm birth. The Cochrane Database of Systematic Reviews : Cd006843.

[267] Honest H, Bachmann LM, Gupta JK, Kleijnen J, Khan KS (2002) Accuracy of cervicovaginal fetal fibronectin test in predicting risk of spontaneous preterm birth: systematic review. BMJ (Clinical Research ed) 325: 301.

[268] Crane JMG, Hutchens D (2008) Transvaginal sonographic measurement of cervical length to predict preterm birth in asymptomatic women at increased risk: a systematic review. Ultrasound in Obstetrics & Gynecology : the Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology 31: 579–587.

[269] Gale C, Hay A, Philipp C, Khan R, Santhakumaran S, et al. (2012) In-utero transfer is too difficult: results from a prospective study. Early Human Development 88: 147–150.

[270] van Baaren GJ, Vis JY, Wilms FF, Oudijk MA, Kwee A, et al. (2014) Predictive value of cervical length measurement and fibronectin testing in threatened preterm labor. Obstetrics and Gynecology 123: 1185–1192.

[271] Neonatal Clinical Review Group, NHS Commissioning Board (2013) 2013/14 NHS standard contract for neonatal critical care. Technical report, NHS England, Online. URL http://www.england.nhs.uk/wp-content/uploads/2013/06/e08-neonatal-critical.pdf.

[272] Bray BD, Ayis S, Campbell J, Cloud GC, James M, et al. (2014) Associations between stroke mortality and weekend working by stroke specialist physicians and registered nurses: Prospective multicentre cohort study. PLoS Medicine 11: e1001705.

Appendices

Appendix A

Data items collected in the EPICure study (1995)

The following text is copied verbatim from the supplementary appendix to the paper by Costeloe et al published in October 2000. [54]

> Data items included in the standardized form for all infants admitted to an NNU. Items requiring an answer yes or no are indicated y/n and ranges of options are given in parentheses. Throughout the record there were opportunities for the information to be amplified with free text.
>
> For all admissions: center number; 6-digit patient identifier; EDD by LMP; EDD by scan <20 weeks' gestation; maternal age; ethnic origin (white, black African, black Caribbean, black other, Indian, Pakistani, other-specified, not known); number of previous pregnancies; number of live births; any obstetric problems in this pregnancy, y/n; preeclampsia = hypertension (untreated diastolic >90 mm Hg) appearing in pregnancy with proteinuria, y/n; antepartum hemorrhage = any vaginal bleeding >20 weeks gestation after exclusion of local hemorrhage from the genital tract, y/n; prolonged rupture of membranes >24 hours (PROM) = membranes ruptured for >24 hours, y/n; chorioamnionitis suspected or with bacteriologic or histologic proof (these were combined for analysis); antenatal steroids, y/n; steroids started >24 hours before birth, y/n; maternal thyroid releasing hormone, y/n; tocolysis, y/n; mode of delivery (vaginal, cesarean section in labor, cesarean section not in labor); presentation (cephalic, breech, other);

hospital of birth; hospital providing continuing intensive care; age (hours) at admission to hospital providing continuing intensive care; plurality (singleton, twin, triplet, other); birth order; date and time of birth; sex; birth weight (g); occipito-frontal circumference (cm); maximum base-deficit in first 12 hours9; minimum appropriate fraction of inspired oxygen (Fio2) in first 12 hours6; maximum appropriate Fio2 in first 12 hours9; HR: >100 bpm at 5 minutes, y/n; congenital anomaliesfree text; first recorded temperature; time of first chest radiograph; radiograph score (0 = normal; 1 = fine reticulo-granular mottling, good lung expansion; 2 = mottling with air bronchograms; 3 = diffuse, mottling, heart borders just discernible, prominent air bronchograms; and 4 = bilateral confluent opacification of lungswhiteout); received surfactant, y/n; type (Survanta, Abbott Laboratories Ltd, Kent, United Kingdom; Curosurf, Serono Laboratories (UK) Ltd, Hertfordshire, United Kingdom; Exosurf, Wellcome UK, Middlesex, United Kingdom; ALEC, Britannia Pharmaceuticals Ltd, Surrey, United Kingdom; other); date of last tracheal intubation; date of last continuous positive airway pressure; last day of supplemental oxygen; systemic steroids for chronic lung disease, y/n; date of starting steroids; total number of days of steroids; pulmonary hemorrhage = acute onset of bloody tracheal secretions with acute deterioration requiring change of ventilator management, y/n; patent ductus arteriosus (PDA), y/n; indomethacin to treat PDA, y/n; ligation of PDA, y/n; insertion of abdominal drain for suspected perforation, y/n; laparotomy for necrotizing enterocolitis, y/n; other surgical proceduresfree text; medications at EDD and/or discharge (systemic steroids, diuretics, anticonvulsants, methyl xanthines, others excluding nutritional supplementsspecify); received total parenteral nutrition, y/n; age amino acids started; age lipids started; age enteral feeding started; received breast milk, y/n; weight, length, and occipito-frontal circumference at 4 weeks and EDD; developed any signs of ROP at any time, y/n; ROP right eye treated (cryotherapy, laser, date of first treatment); ROP left eye treated (cryotherapy, laser, date of first treatment); cerebral ultrasoundfirst after birth and closest to 1 week, 6 weeks, and EDD scoring each side separately:

hemorrhage (0 = none, 1 = subependymal or choroidal, 2 = intraventricular, and 3 = parenchymal), ventricular size (0 = no dilatation, 1 = <4 mm >97th percentile,5 and 2 = >4 mm >97th percentile), parenchymal cysts (0 = none, 1 = porencephalic cyst(s), and 2 = cystic leukomalacia); discharge before EDD, y/n; date of discharge; transfer(s) before EDD, y/n; destination(s) after transfers and/or discharge; death, y/n; date of death; active withdrawal of intensive care, y/n (asked only to tick yes if a formal decision had been made to withdraw care after appropriate discussion with family and staff and not to include occasions when the infant was extubated before death after an acute deterioration and unsuccessful resuscitation); postmortem examination, y/n (full, limited); and principal category of death (congenital anomaly + details; pulmonary immaturity = structural immaturity of the lung so gross as to render sustained ventilatory support unsatisfactory from the outset; RDS; RDS with intracerebral hemorrhage; RDS with infection; late sequelae of ventilation; intraventricular hemorrhage; other intracranial hemorrhage; necrotizing enterocolitis; other infections; and other, specify).

Throughout the form opportunities were given for free text to amplify the information.

For survivors to discharge: name at birth and discharge; National Health Service number; mother's National Health Service number; name and address of mother/principal carer; and name and address of general practitioner and of responsible pediatrician.

Appendix B

Data items collected in the EPICure 2 study (2006)

The text on the following pages is copied verbatim from the first appendix to the article by Costeloe et al reporting results from the EPICure 2 cohort, published in December 2012. [55] It was retrieved from the BMJ website on 14th August, 2014 from: `http://www.bmj.com/content/bmj/suppl/2013/01/07/bmj.e7976.DC2/cosk004134.ww1.pdf`

Appendix 1: EPICure 2 dataset [posted as supplied by author]

Items requiring an answer yes or no or not known are indicated y/n/nk, ranges of options are given in parentheses.

All births

Items 8-digit patient identifier.
Summary data submitted at 28days or death if sooner: case definition (late fetal loss, stillbirth, livebirth); timing of death for stillbirths and late fetal losses (antepartum, intrapartum, not known); legal abortion y/n/nk, outcome for livebirths (early death on labour ward, death <7d on NICU, death 7-27d, alive at 28d, not known)

Maternal items, pregnancy and delivery
Maternal date of birth, dd/mm/yy, or age if date of birth not known; Ethnic group (White, Black African, Black Caribbean, Black other, Indian, Pakistani, Bangladeshi, Chinese, mixed [specify], other [specify], not known); maternal height and weight, maternal BMI if available and weight and height not known; previous pregnancy outcomes: (spontaneous miscarriages, terminations of pregnancy, stillbirths, livebirths, caesarean sections, preterm births 24 – 36w); date of last menstrual period dd/mm/yyyy (certain, uncertain, not known); date of first booking appointment dd/mm/yyyy (certain, uncertain, not known, never booked); date of first scan dd/mm/yyyy or never scanned; gestational age at first scan ww+d; intended place of delivery at booking (hospital [specified], home, outside UK, not known); actual place of delivery (hospital [specified], outside hospital, not known); if hospital delivery (on labour ward, not on labour ward, not known); main reason for change between planned and actual place of delivery (no change, change of address during pregnancy, preterm labour and emergency admission to different hospital, planned in-utero transfer for clinical reasons, planned in-utero transfer for other reasons e.g. lack of capacity, not known); maternal smoking in pregnancy at time of booking (y/n/nk); maternal medical complications in pregnancy – indicate all that apply (none, pre-pregnancy diabetes type 1 or 2, gestational diabetes insulin dependent/non insulin dependent, essential hypertension on treatment at time of booking, epilepsy on treatment at first booking); obstetric complications – indicate all that apply (none, prolonged pprom >24h with date of membrane rupture dd/mm/yyyy, abruption, ante-partum bleeding after 20 completed weeks, pre-eclampsia, cervical suture); clinical suspicion of chorioamnionitis at any time before birth with date specified y/n/nk; maternal pyrexia >37.5^{O}C during 24h before birth yes with temperature specified/ no/ nk; chorioamnionitis noted at the time of birth y/n/nk; maternal antibiotics within 24h of birth (none, prophylaxis for PPROM, prophylaxis for known GBS carriage, treatment for suspected chorioamnionitis, treatment for other [specified], not known); evidence of fetal compromise >24h before birth (none, IUGR, oligohydramnios without PPROM); ante-partum CTG >24h before birth (none, normal, non-reassuring, pathological, not known); ante-partum Dopplers >24h before birth (not done, normal, evidence of redistribution with umbilical artery end diastolic flow present, umbilical artery absent or reversed end-diastolic flow, not known); ante-natal steroids (none, betamethasone, dexamethasone, not known + if given last dose more or less than 24h before birth, not known); tocolysis (none, atosiban, ritodrine, indometacin, nifedipine, other [specified]); labour (induced + main indication, spontaneous, spontaneous with augmentation, never in labour, not known); fetal monitoring in labour (none, continuous CTG, intermittent CTG, auscultation only, not known); CTG interpretation in labour (normal, non-reassuring, pathological, not known); epidural y/n/nk; presentation immediately prior to delivery (cephalic, breech, other, not known); mode of delivery – include all attempted – (spontaneous, instrumental,

caesarean section, not known); caesarean section indication if applicable – include all – (immediate threat to life of mother and/or fetus, maternal and/or fetal compromise not immediately life threatening); health professionals present immediately before birth – indicate all present – (none, qualified midwife, student midwife, consultant obstetrician, obstetric middle grade, obstetric SHO or F1/F2, ANNP, neonatal nurse, consultant paediatrician, paediatric middle grade, paediatric SHO or F1/F2); maternal supporters present during labour or delivery (none, partner, children, other family members, friends, lay supporter); fetus alive at admission to hospital: y/n/nk; fetus alive at onset of labour: y/n/nk; congenital anomaly suspected before birth (yes [specify], no, not known); congenital anomaly noted at delivery (none, present [specify], possible dysmorphism [specify], not known); was a plan for preterm birth discussed with an middle grade or consultant obstetrician (no opportunity, no, yes, not known); was a decision made not to perform a CS for fetal distress (y/n/nk); was paediatric counselling provided if so by whom (none, consultant, middle grade doctor, ANNP, SHO or F1/F2, nurse other than ANNP); did the parents express a choice about resuscitation and the provision of intensive care (no choice expressed, provide full intensive care for any live birth, withhold intensive care, assess and provide care at paediatric discretion, not known); was the possibility of withholding intensive care discussed (y/n/nk); date and time of birth.

Live births only

Trunk in occlusive wrapping at birth to avoid hypothermia y/n/nk; any heart rate at birth y/n/nk; signs of life in the first hour - indicate any that apply – (audible cry, spontaneous breathing, active body movements, heart beat); resuscitation – incluse any that apply – (not attempted, stimulation, oxygen, mask ventilation, ventilation via ETT + age at intubation, nCPAP, CPR, adrenaline, sodium bicarbonate); heart rate >100bpm at 5 minutes y/n/nk; surfactant given on labour ward y/n/nk

Babies admitted to Neonatal Unit for intensive care

If there was neither dating scan or certain LMP was the consultant confident that the GA was <27 weeks (y/n/nk); agreed gestational age at birth (ww/d); for what type of care was the baby admitted to the NNU (intensive, palliative, nk); name of hospital where baby was first admitted; date and time of admission; birth weight (n.nnn kg / never recorded); head circumference at birth (nn.n cm); sex (male, female, indeterminate); maximum base deficit in first hour (nn.n mm/l); maximum appropriate FiO_2 in first 12h (n.nn); minimum appropriate FiO_2 in first 12h (n.nn); maximum base deficit in first 12h (mm/l); temperature at admission (nn.n^{O}C); time temperature taken; surfactant given after admission to NNU (none, animal derived, synthetic, not known); prophylactic indometacin or ibuprofen (y/n/nk); transferred to another hospital within 24h of birth (no, yes – if yes where & time of admission to second hospital); TPN given (yes/no, if yes dates started amino acids and lipids) date enteral feed started (date / never fed); maternal breast milk at any time (y/n); date reached enteral feeding 150 ml/kg/day; any maternal breast milk at discharge (y/n/nk); at 36w pma still receiving mechanical respiratory support (yes/no); at 36w still receiving oxygen (no, ≥30% O_2, <30%O_2, low flow O_2 >0.1l/min, ≤0.1l/min); date last in supplemental O_2; home in O_2 (y/n); systemic steroid for BPD (none, dexamethasone, other [specify], date first given, starting dose mg/kg/day, weight when started, total days given steroid, number of separate courses, total dose mg); pulmonary haemorrhage (y/n); details of corticosteroid given other than for BPD; positive blood culture at first admission (none, GBS, E Coli, other [specify]; any other positive blood culture within 72h of birth (none, GBS, E Coli, other [specify]; positive blood culture >72h after birth (none, coagulase negative Staph, other{specify]); PDA treated with indometacin and/or ibuprofen (y/n/not applicable); ligation of PDA (y/n/not

applicable); were any suspected congenital anomalies confirmed (y/n/nk + details); were additional congenital anomalies detected & confirmed on NNU (y/n/nk); screened for ROP (y/n); date of first screen: worst stage of ROP in each eye (none, I, II, III, IV, V); plus disease (y/n); date of first treatment; method of treatment (laser/cryotherapy); cerebral ultrasound scan information requested if available – first scan, week 1, weeks 2-6, week 7-EDD, all scored on each side for haemorrhage (none, germinal layer, intraventricular), ventricular size (no dilatation, ventricular index ≤4mm over 97th centile, ventricular index >4mm over 97th centile), parenchymal injury (no evidence of injury, HPI echodense, HPI porencephalic cyst, PVL), extent of PVL if present (frontal, parietal, posterior)*; surgical procedures (none, abdominal drain for suspected perforation, Laparotomy for NEC, Laparotomy for perforation without NEC, intestinal resection, stoma, inguinal hernia repair, v-p shunt, Rickman reservoir, other[specify]); weight, OFC and length at 40w pma or discharge if sooner; transfers after 24h – dates and details of destination hospitals; did medical staff at any time recommend withdrawal of intensive care (y/n).

* The definition published in 1995 (1) of severe cerebral ultrasound scan abnormality was " showing unilateral or bilateral parenchymal cysts and/or hydrocephalus". Review of the 1995 database and analyses reveals this to be an error and scans showing echodense haemorrhagic parenchymal infarcts were included in that analysis as in 2006.

Late fetal losses and stillbirths

Date and time of death if known; gestation at which death confirmed; cause of IUD after admission (feticide, abruption, IUGR and hypoxia, other [specify], nk).

All post-admission deaths

Was intensive care electively withdrawn after discussion between family and staff (y/n); date, time and place of death; principal category of death (congenital anomaly [specify], pulmonary immaturity, RDS, IVH, RDS with IVH, RDS with infection, infection, NEC, late sequelae of ventilation, other [specify], not known).

All deaths

Post mortem examination (held or to be held, not offered, permission refused, coroner's PM, consent given but not performed, not known).

Babies discharged from hospital

Date of discharge; need for interpretation (y/n), parents' names and address, name and address of GP.

Reference

1. Costeloe K, Hennessy E, Gibson AT, Marlow N, Wilkinson AR. The EPICure study: outcomes to discharge he EPICure study: outcomes to discharge from hospital for infants born at the threshold of viability. Pediatrics. 2000; **106**(4): 659-71.

Appendix C

UK Neonatal Staffing Study questionnaire (1997)

The following page contains the questionnaire that was distributed as part of the UK Neonatal Staffing Study in 1997.

C. UK NEONATAL STAFFING STUDY QUESTIONNAIRE (1997)

Appendix 1 **: PHASE 1 CENSUS FORM**

Name of Nurse in Charge	Hospital

1. Please specify the time period used if other than the calendar year of 1996. ☐

2. Total number of deliveries in your hospital ? ☐

3. Does your unit provide only temporary support before safe transport for full intensive care elsewhere? ☐

Yes or No?

4. Total number of admissions to your NICU and/or SCBU? ☐

5. Total number of admissions to your NICU and/or SCBU <1500g? ☐

6. Total number of infants ventilated or given CPAP? ☐

Tick the box (√) if your total is only for those managed with an endotracheal tube ☐

7. Total number of ventilator days? ☐

Tick the box (√) if your total is only for those managed with an endotracheal tube ☐

8. Total number of cots? ☐

9. Total number of maximum intensive care cots (level 1)? ☐

10. Total number of consultants who contribute to the emergency on-call rota for neonatal care? ☐

11. Total number of consultants with 50% or more of their clinical sessions dedicated to neonatal care? ☐

12. Total number of fixed consultant-led business rounds per week? ☐

13. Total number of whole-time equivalent (wte) nurses on neonatal unit staff? ☐

14. Total number of whole-time equivalent (wte) nurses with a nationally-recognised qualification in neonatal intensive care? ☐

15. Does your unit provide a neonatal tertiary cardiac or surgical service? ☐

Yes or No?

16. Do you expect significant changes to these same questions in 1997? ☐

Yes or No?
(if Yes go to Q17)
↓

17. Please describe any planned significant changes for 1997/98 here:

(please continue on the back page if necessary)

82

Appendix D

Unit Profile Survey questionnaire (2006)

The next four pages contain a copy of the Unit Profile Study 2006 that was circulated by the EPICure study group to perinatal centres located in England.

EPICure 2
A study of extreme prematurity in 2006

Unit Profile Study

This form is being circulated to all hospitals in England that are contributing data to the EPICure 2 study. It forms part of the main EPICure data collection.

Its purpose is twofold:

- To collect information about workload and activity of each hospital
- To collect simple information about the approach of maternity and neonatal departments to extremely preterm births

The questions about activity, capacity and staffing are very little changed from a questionnaire that all hospitals completed in 1996 as part of the UK Neonatal Staffing Study. The information collected will be analysed using similar methodology to that which was used then so as to classify hospitals in terms of workload and the numbers of staff available. We have also agreed with the investigators from the staffing study that they will perform a comparison with the data reported by the same hospitals in 1996 that is held on their database.

When the EPICure investigators team was negotiating with the Medical Research Council (MRC) to secure funding for EPICure 2, the referees and the MRC were keen that we should try to study the impact on perinatal outcomes of both institutional and individual professional attitude towards the problems posed by extremely preterm birth. The questions on the final page of this questionnaire are a simple attempt to do this. At the same time work is going ahead, under the leadership of Dr Cris Glazebrook, Associate Professor in Health Psychology in Nottingham, to develop an instrument to study this area in more detail. Cris is seeking further funding to support this and, if successful, will be taking this forward after the end of EPICure 2 data collection in 2007.

This form is being sent by post to you as a designated Paediatric Co-ordinator for EPICure for your Trust. If you want it electronically, please let us know by email and we'll send it straight away. Completing this form will, in almost all cases, require consultation with medical and nursing or midwifery staff both within the maternity and neonatal services.

In particular, when it comes to completing Part 2 we are keen to try to get a consensus view. We do understand that this will not always be possible. For instance, if it is clear after discussion that the midwives and obstetricians within a unit cannot agree as to when they would usually expect to use continuous fetal monitoring please write this on your form, if possible with some information about the extent of the range of view – we have not put a specific cell giving you this option as we hope that you will only use it as a last resort.

The responses to this survey will be used to generate items that will be used in the final analysis of EPICure 2 to test their impact on clinical outcomes. It will not be possible to identify any institution in any presentation or publication arising from your responses to this form.

Please return your completed form by **31st July 2006** to the EPICure office at Homerton University Hospital. To do this you can use the Freepost address. All you need is to put your form in an envelope and write on it '**FREEPOST EPICURE 2'**. Please don't write anything else as otherwise the envelope might go astray! In case anything goes wrong please keep a photocopy of the form before posting it to us.

Very many thanks for your help.

Kate Costeloe
Kate.Costeloe@homerton.nhs.uk

Kerrie Montoute
Kerrie.Montoute@homerton.nhs.uk

Unit Profile Study Form

PART 1: DETAILS OF ACTIVITY AND STAFFING

This form is designed so that you can circle or tick items. Please read any comments in the right hand column before completing. In most cases the form is being sent to the named Paediatric EPICure Co-ordinator. It will usually be necessary to consult neonatal nursing, obstetric and midwifery colleagues in order to complete this form.

Please refer any queries to **Kerrie.Montoute@homerton.nhs.uk** ***or*** **Kate.Costeloe@homerton.nhs.uk**

Questions 1- 7 relate to ACTIVITY ***Please complete for births in the calendar year 2005***			
Item	**Response**		**Comments and Notes**
1. Is the activity data provided for the calendar year 2005?	Yes	NO	1 and 1a. Please give activity data for your unit for the calendar year 2005 if available. If not use the most recent 12-month period but please specify what period you have used.
1a. If the answer to Q1 was 'No' please specify the time period you have used.			
2. Total number of deliveries in your hospital.			2. Give total number of women delivered in your 12 month period, if necessary round to the nearest 100.
3. Does your unit aim only to provide temporary support for babies needing intensive or high dependency care (including CPAP) before safe transport elsewhere?	YES	No	3. If the policy of your unit is to provide only temporary support for all babies prior to safe transport to another unit for intensive or high dependency care, answer "yes".
4. Total number of admissions to your neonatal unit.			4. Give total number of admissions to your neonatal unit for all birth weights (please include transfers from other hospitals and re-admissions).
5. Total number of admissions to your neonatal unit <1500g.			5. Give total number of admissions to your neonatal unit for those with birth weight <1500g (please include transfers from other hospitals and re-admissions).
6. Total number of infants ventilated or given CPAP. If you collect your data in such a way that you cannot answer this question e.g. you can only provide numbers of babies ventilated via an endotracheal tube please give what data you can and indicate what it refers to ***e.g. 54 babies ventilated via an ETT, no information about CPAP available.***			6. Give total numbers of infants given mechanical ventilation or CPAP (by ET tube, or face mask or nasal prong), after admission to your neonatal unit i.e. do not include babies who receive IPPV at resuscitation at birth but who cease to require mechanical support before admission to the neonatal unit.
7. Total number of days of ventilation and /or CPAP. If you collect your data in such a way that you cannot answer this question e.g. you can only provide numbers days of ventilation via an endotracheal tube please give what data you can and indicate what it refers to ***e.g. 600 days of ventilation via an ETT, no information about CPAP available.***			7. Give total number of days all infants receive IPPV or CPAP (by endotracheal tube or face mask or nasal prong) after admission to your neonatal unit.

D. UNIT PROFILE SURVEY QUESTIONNAIRE (2006)

Questions 8 – 15: COT CAPACITY AND STAFFING ***Please complete for situation at 1st January 2006***			
8. Total number of cots			8. Give the total number of cots in your establishment.
8a. BAPM Categories of Care Do you categorise your cots as intensive or high dependency using BAPM 1992 or 2001 definitions – please circle.	1992	2001	8a. The system for classifying cots on Neonatal Units changed in 2001 with the publication of the 2nd edition of the BAPM Standards For Hospitals Providing Neonatal Intensive and High Dependency Care. The major change was that using the previous (1992) system all babies receiving any CPAP were classified as receiving intensive care whereas using the 2001 system most babies <1000g or receiving CPAP who fulfil no other criteria for intensive care are classified as receiving high-dependency care. If you are in any doubt as to which system you are using please seek clarification with your Clinical Director.
9a. Total number of intensive care cots			9a & b. Give total numbers of intensive care and high dependency cots categorised using the system you have indicated at 8a.
9b. Total number of high-dependency cots			
10. Total number of consultants who contribute to the emergency on-call rota for neonatal care.			10. Count each consultant on the on-call rota for the neonatal unit at nights and weekends. Please confirm answer to this question with your consultant in charge.
11. Total number of consultants with 50% or more of their clinical sessions or clinical and administrative PAs dedicated to neonatal care.			11. Count each consultant with 50% or more of their clinical sessions or clinical and administrative PAs dedicated to neonatal care (including clinics, postnatal ward rounds & neonatal administration). This may be zero in some neonatal units. Please confirm answer to this question with your consultant in charge. **e.g. *Example 1:* 4 consultants are on the on-call rota. 2 have 20% and 2 have 50% of the time in their job plan allocated for clinical and administrative work for neonatal care. This should be recorded as 2. *Example 2:* 5 consultants are on the on-call rota, all have 20% of the time in their job plan allocated for clinical and administrative work for neonatal care. This should be recorded as zero.**
12. Total number of fixed consultant led business rounds per week.			12. Count morning and evening consultant business rounds with junior staff Monday-Sunday. Please confirm answer to this question with your consultant in charge. **e.g. *Example 1:* 1 or more consultants do rounds in the morning and evening each day Mon-Sun, this should be recorded as 14 rounds. *Example 2:* 1 or more consultants do rounds on Mondays, Wednesdays and Fridays in the morning, this should be recorded as 3 rounds.**
13. Total number of Whole Time Equivalent (wte) nurses including ANNPs on your neonatal unit staff establishment who provide clinical care.			13. Give total number of budgeted whole time equivalent nurses who provide hands on clinical care, (this number should include current vacancies & those on maternity or sick leave and also time contributed to the nursing rota by ANNPs; it should not include time that ANNPs spend working on the medical rota. Time allocated for community or clinical facilitator / lecturer work should be excluded.
13a.How many of WTE of working time in 2005 were lost through vacancies, maternity leave, long term sick leave and secondments.			
14. How many of the nurses included in your response to Q13 have a nationally recognised qualification in Neonatal Intensive Care? – please give your answer in Whole Time Equivalents.			14. English & Scottish recognised qualifications e.g. ENB 402, 405 or 904 certificates and A19 Scottish Neonatal Nursing Certificates & the PSH module in Neonatal Critical Care.
15a. Does your unit provide a neonatal tertiary *cardiac* service?	YES	NO	15a & b A tertiary specialist centre for neonatal cardiac or surgical care is one that receives referrals from other hospital based consultants.
15b. Does your unit provide a neonatal tertiary *surgical* service?	YES	NO	

PART 2: POLICIES

Maternity unit policies: This section to be completed by Labour Ward staff ***Please discuss within the department so that a consensus view is presented describing the situation in your hospital at the present time***						
16. Does your maternity unit have written clinical guidelines for the management of extremely preterm labour and birth?	YES			NO		
17. Tick box that best describes your consultant Labour Ward cover.			More			
18. Circle the lowest gestational age (in completed weeks) that staff in your hospital would ever consider performing a Caesarean Section in a singleton pregnancy for fetal reasons.	22w	23w	24w	25w	26w	27w
19. Circle the lowest gestational age (in completed weeks) that staff in your hospital would usually expect to use continuous electronic fetal monitoring.	22w	23w	24w	25w	26w	27w
Paediatric policies: This section to be completed by Neonatal Unit staff ***Please discuss within the department so that a consensus view is presented***						
20. Do you have written departmental guidelines for the management of extremely preterm babies at birth?	YES			NO		
21. Does your policy specify which paediatric staff should be present at an extremely preterm birth?	YES			NO		
22. Circle the lowest gestational age (in completed weeks) that you would aim routinely to provide respiratory support by intubating or using nCPAP at birth.	22w	23w	24w	25w	26w	27w
23. Do you have a strategy in the hospital for preventing neonatal hypothermia?	YES			NO		
24. Does your strategy to prevent hypothermia involve putting the baby in plastic bag or wrapping?	YES			NO		

Name (please print):

Name of Hospital:

Contact Details:

THANK YOU VERY MUCH FOR YOUR HELP

PLEASE PHOTOCOPY THIS FORM & SEND THE ORIGINAL BY
31st JULY 2006 TO 'FREEPOST EPICURE 2'

Appendix E

Unit Profile Survey questionnaire (2011)

The following four pages show the Unit Profile Survey form that was mailed out to English neonatal units in early-mid November, 2011.

The same questions were also made available online using the Opinio 6.4.4 software provided by University College London. Some respondents had difficulty accessing this, and had not received a paper version in the post, hence a PDF version of the questionnaire was emailed to them.

Unit Profile Survey (2011)

This survey contains 4 pages of questions – 22 questions in total. Many are easy to answer, but several will need some additional research: either by asking colleagues for their input or by reviewing unit data such as admissions books (or electronic equivalent). The form is designed to be straightforward and easy to use. Boxes are provided by each of the questions for your answer: please tick or write your answer as appropriate. If there is insufficient space for additional comments, you should continue on another sheet. Notes are provided under each question to guide your answers. Please also remember to fill in the name of your unit/hospital at the top of each page.

EPICure
Population based studies of survival and health after extremely premature birth

BRITISH ASSOCIATION OF PERINATAL MEDICINE
Founded 1976

Contact information

Note: These details are requested in case there is a need to follow-up any of the questions.

Name: .....
Email: .....
Telephone: .....
Address: .....
.....
.....
.....
Post code: .....

NDAU
Neonatal Data Analysis Unit

Bliss
for babies born too soon, too small, too sick

NNAP

SECTION 1:
This section relates to *activity* during the calendar year 2010.

1 Are the data provided in this section based on the calendar year 2010?

Yes	No	If no, which dates?	
		From:	dd/mm/yyyy
		To:	dd/mm/yyyy

Please provide data about your unit for the calendar year 2010 if available. If not, use the most recent 12-month period but please specify what period you have used.

2 What is the total number of women who delivered in your hospital?

Provide the total number of women who delivered in your hospital during the period stated in question 1; multiple births in a single pregnancy should only count as one delivery.

3 Does your unit aim only to provide temporary support for babies needing intensive or high dependency care (including CPAP and high flow O_2) before safe transport elsewhere?

Yes	
No	

If the policy of your unit is to provide only temporary support for babies requiring high dependency or intensive care prior to transfer somewhere else, please answer "yes".

4 What is the total number of *babies* admitted to your neonatal unit?

All weights:	**<1500g:**

Please provide the total number of babies of any weight *and* the total number of babies who weighed <1500g at birth who were admitted during the stated period.
Only count *one* admission per baby in each category.

5 What is the total number of *admissions* to your neonatal unit?

All weights:	**<1500g:**

For the period stated in question 1, provide the total number of admissions to your neonatal unit for the first category. In the second, provide the total number of admissions of babies weighing <1500g. Babies may be counted more than once if they still weigh <1500g on a second or subsequent admission.

Return to: Dr. Laura McCormack, EPICure office, Institute for Women's Health, UCL, 74 Huntley Street, London, WC1E 6AU

6 What is the total number of babies receiving:

Ventilation	CPAP	High flow O_2	Any mode

Notes:

This section relates to the *total number of babies* receiving respiratory support during the specified period. **Each baby should be counted once only – for the highest mode of respiratory support received.** Only include respiratory support received on the neonatal unit (i.e. not in the delivery room). *Do **not** include babies who only receive supplementary (low flow or head box) oxygen*. If you are unable to provide a breakdown, please provide whatever information you can together with an explanatory note. For example, if babies receiving high flow oxygen are counted as CPAP, please indicate that in the notes section. Section (d) should equal the total of parts (a) to (c); please do **not** include babies who receive supplementary (low flow or head box) oxygen in this total.
Definitions: ventilation: invasive (endotracheal) support; CPAP: non-invasive support; High flow O2: vapotherm, optiflow or similar (*excluding* low-flow or headbox oxygen).

7 What is the total number of days in your unit of:

Ventilation	CPAP	High flow O_2	Any mode

Notes:

This section relates to the *total number of days* of respiratory support provided by your neonatal unit during the stated period. Only include respiratory support received on the neonatal unit (i.e. not in the delivery room).
*Do **not** include the number of days for babies receiving supplementary (low flow or head box) oxygen.*
If you are unable to provide a breakdown, please provide whatever information you can together with an explanatory note. For example, if days of high flow oxygen is counted by your unit as days of CPAP, please indicate that in the notes section. Section (d) should equal the total of parts (a) to (c); please do **not** include the number of days of supplementary (low flow or head box) oxygen that your unit provides.
Definitions: ventilation: invasive (endotracheal) support; CPAP: non-invasive support; High flow O2: vapotherm, optiflow or similar *(excluding* low-flow or headbox oxygen).

SECTION 2:
This section relates to *staffing* of the neonatal unit

8 How many consultants contribute to the emergency on-call rota for neonatal care?

Count each consultant on the on-call rota for the neonatal unit at nights and weekends. Please confirm the answer to this question with your Clinical Director.

9 What is the total number of consultants with 50% or more of their clinical sessions or clinical and administrative PAs dedicated to neonatal care?

Count each consultant with 50% or more of their clinical sessions or clinical and administrative PAs dedicated to neonatal care, including clinics, postnatal ward rounds and neonatal administration. This may be zero in some units. Please confirm the answer to this question with your Clinical Director.

- ***Example 1:*** *4 consultants are on the on-call rota. 2 have 20% and 2 have 50% of the time in their job plan allocated for clinical and administrative work for neonatal care. This should be recorded as 2.*
- ***Example 2:*** *5 consultants are on the on-call rota, all have 20% of the time in their job plan allocated for clinical and administrative work for neonatal care. This should be recorded as zero.*

10 How many fixed, consultant led business rounds are there per week?

Count morning and evening consultant business rounds with junior staff.
If a business round is conducted at the same time at two locations by different consultants (e.g. one on ITU and a second on HDU/SCBU), they should be counted as a single ward round.

- ***Example 1:*** *One or more consultants do rounds in the morning and evening each day Monday to Sunday; this should be recorded as 14 rounds/week.*
- ***Example 2:*** *One or more consultants do rounds every morning. During the working week (Monday to Friday), a further ward round is held in the evening. This should be recorded as 12 rounds/week.*
- ***Example 3:*** *One or more consultants do rounds on Mondays, Wednesdays and Fridays in the morning; this should be recorded as 3 rounds/week.*

Return to: Dr. Laura McCormack, EPICure office, Institute for Women's Health, UCL, 74 Huntley Street, London, WC1E 6AU

Unit Profile Survey (2011)

11 Provide the total number of the following in whole time equivalents (WTE) *actually* working or on annual leave on the medical rotas in your unit on 22nd November 2011:

a) Tier 1 (ST1-3/SHO)**:** ☐
b) Tier 2 (ST4-8/SpR)**:** ☐
c) ANNPs (in a medical role)**:** ☐

Give total number of whole time equivalent (WTE) staff on the medical rota who provide hands on clinical care. **This number should only include those currently working or on planned leave;** vacancies, those on maternity or sick leave, or posts covered by locums should *not* be counted. Advanced Neonatal Nurse Practitioners (ANNPs) working only in a nursing role, or time allocated for community or clinical facilitator/lecturer work, should also be excluded.

12 Provide the total number of the following in whole time equivalents (WTE) who *should be* working on the medical rotas in your unit on 22nd November 2011:

a) Tier 1 (ST1-3/SHO)**:** ☐
b) Tier 2 (ST4-8/SpR)**:** ☐
c) ANNPs (in a medical role)**:** ☐

Give total number of budgeted whole time equivalent (WTE) staff on the medical rota who provide hands on clinical care. **This number should include current vacancies and those on maternity or sick leave; it does *not* include those on annual leave.** Advanced Neonatal Nurse Practitioners (ANNPs) working exclusively in a nursing role, or time allocated for community or clinical facilitator/lecturer work, should be excluded.

13 On 22nd November 2011, did you have dedicated medical rotas at:

a) Tier 1 (ST1-3/SHO level)**?** **Yes** ☐ **No** ☐
b) Tier 2 (ST4-8/Registrar/middle-grade level)**?** **Yes** ☐ **No** ☐

In your hospital, is there a separate medical rota that only covers neonates at Tiers 1 and/or 2 – this is equivalent to ST1-3 (the old SHO) and ST4-8 (the old registrar) grades, respectively.

14 Provide the total number of the following in whole time equivalents (WTE) working on 22nd November 2011 on the nursing rotas in your unit who provide clinical care:

a) ANNPs/nurse consultants ☐
b) Trained nurses ☐
c) Nursery nurses/HCAs ☐

Give the total number of budgeted whole time equivalent nurses who provide hands-on clinical care; this number should include current vacancies and those on maternity or sick leave, as well as time contributed to the nursing rota by ANNPs. It should *not* include time that ANNPs spend working on the medical rota (this is covered in question 16). Time allocated for community or clinical facilitator/lecturer work should be excluded.

15 How many of the nurses included in your response to question 14 have completed a neonatal qualification (please provide your answer in WTE):

a) ANNPs/nurse consultants ☐
b) Trained nurses ☐
c) Nursery nurses/HCAs ☐

English and Scottish recognised qualifications, e.g. ENB 402, 405 or 904 certificates; A19 Scottish Neonatal Nursing Certificates; and the PSH module in Neonatal Critical Care. Please provide your answer using Whole Time Equivalents (WTE).

16 On 22nd Nov. 2011, how many of the following:

	ANNPs/Nurse Consultants	Trained nurses	Nursery nurses/HCAs
a) worked their normal, routine shift?			
b) worked an additional, or 'bank', shift?			
c) were external, agency staff?			
d) were absent through sick leave?			
e) were absent through maternity leave?			
f) were absent through secondment?			
g) were absent for any other reason?			

Include staff from all nursing shifts who were at work between 00:00 and 23:59 on Tuesday, 22nd November 2011.

Return to: Dr. Laura McCormack, EPICure office, Institute for Women's Health, UCL, 74 Huntley Street, London, WC1E 6AU

SECTION 3:
This section relates to the *facilities* provided by your neonatal unit

17 On 22nd November 2011, how was your unit designated within your network?

- **Special Care Unit** ☐
- **Local Neonatal Unit** ☐
- **Network NICU** ☐

Other:

The '*Toolkit for High-Quality Neonatal Services*' (DoH, 2009) defines categories of neonatal unit as listed in the question. These are broadly based upon the previous definitions of levels of care from '*Standards for Hospitals Providing Neonatal Intensive and High Dependency Care (2nd Edition)*' (BAPM, 2001), as follows:

Level 1 = Special Care Unit (SCU);
Level 2 = Local Neonatal Unit (LNU);
Level 3 = Network Neonatal Intensive Care Unit (NICU).

If you normally use a different definition, please provide information about the criteria you have used and how your unit is defined.

18 Please provide the nominal total number of cots in your unit on 22nd November 2011 allocated for:

- **a) Intensive care** ☐
- **b) High-dependency** ☐
- **c) Special care** ☐

TOTAL (excluding transitional care): ☐

For parts (a) to (c), give the total number of intensive care, high dependency and special care cots respectively in your neonatal unit. *This should **not** included any designated transitional care cots.* The final section should equal the total number of cots listed in the first 3 parts of the question.

19 In 2010, which BAPM classification did you use for categorising the level of care?

- **1992** ☐
- **2001** ☐

Other:

The system for classifying care on neonatal units changed in 2001 with the publication of the 2nd edition of the BAPM Standards For Hospitals Providing Neonatal Intensive And High Dependency. The major change was that, using the previous (1992) version all babies receiving any CPAP were classified as receiving intensive care, whereas using the current (2001) system most babies <1000g or receiving CPAP who are otherwise clinically stable are classified as receiving high dependency care. If your neonatal unit is *not* using the BAPM (2001) definitions, please indicate which classification system is being used instead. If there is any doubt as to which system you are using, please seek clarification from your Clinical Director.
Note: an updated BAPM classification was introduced in August 2011; this question relates to the classification of cots in 2010, hence the updated guidelines are not relevant.

20 How many rooms were available on 22nd November 2011 on your neonatal unit for parents to stay overnight?

State how many overnight rooms are available on or immediately adjacent to your neonatal unit for parents to "room in" with their babies prior to discharge. *This should **not** include any beds dedicated for maternal reasons.*

21 On 22nd November 2011, did your unit provide:

	Yes	**No**
• a neonatal tertiary *cardiac* service?	☐	☐
• a neonatal tertiary *surgical* service?	☐	☐

A tertiary specialist service is one that receives referrals from other hospital-based consultants.

22 What is the lowest gestational age limit (in weeks) at birth for the babies that are routinely cared for by your unit?

Please state the lowest gestational age at birth at which babies born at your hospital are routinely cared for on your neonatal unit for their whole postnatal stay (for example, some units care only for babies >26 weeks gestation). *Do **not** include babies who will normally be transferred elsewhere after birth.* If there is no lower limit in your hospital then state 'ALL'. Data should relate to current practice (i.e. November 2011).

Return to: Dr. Laura McCormack, EPICure office, Institute for Women's Health, UCL, 74 Huntley Street, London, WC1E 6AU

Appendix F

Hospital Episode Statistics

The variables shown in table F.1 were provided by the NHS Health and Social Care Information Centre for use in the Admissions Validation Study.

Table F.1: Variables from Hospital Episode Statistics available for data linkage with the EPICure cohorts of 1995 and 2006.

Original HES variable name	**Data description[a]**	**Availability**	
		1995	**2006**
HESYEAR	[Financial] year of HES data collection	✓	✓
study_id	Unique identification number for each row	×	✓
neodur	Age of baby (days) at admission	✓	✓
dob	Date of birth (baby)	✓	✓
dob_cfl	Date of birth check flag	✓	✓
ethnos	Ethnic category	✓	✓
newnhsno	NHS number	×	×
encrypted_hesid	Encrypted HES identification number	×	✓
homeadd	[Maternal] home post code	✓	✓
sex	Sex	✓	✓
admidate	Admission date	✓	✓
ADMI_CFL	Admission date check flag	✓	✓
admimeth	Admission method	✓	✓
admisorc	Admission source	✓	✓
disdate	Discharge date	✓	✓
dis_cfl	Discharge date check flag	✓	✓
disdest	Discharge destination	✓	✓
dismeth	Discharge method	✓	✓
classpat	Patient classification	✓	✓
resha	Strategic Health Authority of residence	✓	✓
soal	Super Output Area - Lower	×	✓
soam	Super Output Area - Middle	×	✓
imd04c	Index of Multiple Deprivation (2004) - Crime Domain	×	✓

Table F.1: (Continued.) Variables from Hospital Episode Statistics available for data linkage with the EPICure cohorts of 1995 and 2006.

Original HES variable name	Data description[a]	Availability 1995	Availability 2006
imd04ed	Index of Multiple Deprivation (2004) - Education Domain	×	✓
imd04hd	Index of Multiple Deprivation (2004) - Housing and Service Domain	×	✓
imd04hs	Index of Multiple Deprivation (2004) - Health and Disability Domain	×	✓
imd04i	Index of Multiple Deprivation (2004) - Income Domain	×	✓
imd04ia	Index of Multiple Deprivation (2004) - Income affecting Adults Domain	×	✓
imd04ic	Index of Multiple Deprivation (2004) - Income affecting Children Domain	×	✓
imd04le	Index of Multiple Deprivation (2004) - Living Environment Domain	×	✓
imd04rk	Index of Multiple Deprivation (2004) - Overall Rank	×	✓
RURURB_IND	Rural-urban indicator	×	✓
dobbaby_1	Date of Birth (baby tail 1)	✓	✓
dobbaby_2	Date of Birth (baby tail 2)	×	×
dobbaby_3	Date of Birth (baby tail 3)	×	×
dobbaby_4	Date of Birth (baby tail 4)	×	×
dobbaby_5	Date of Birth (baby tail 5)	×	×
dobbaby_6	Date of Birth (baby tail 6)	×	×
dobbaby_7	Date of Birth (baby tail 7)	×	×
dobbaby_8	Date of Birth (baby tail 8)	×	×
dobbaby_9	Date of Birth (baby tail 9)	×	×
birordr_1	Birth order (baby tail 1)	✓	✓
birordr_2	Birth order (baby tail 2)	×	×
birordr_3	Birth order (baby tail 3)	×	×
birordr_4	Birth order (baby tail 4)	×	×
birordr_5	Birth order (baby tail 5)	×	×
birordr_6	Birth order (baby tail 6)	×	×
birordr_7	Birth order (baby tail 7)	×	×
birordr_8	Birth order (baby tail 8)	×	×
birordr_9	Birth order (baby tail 9)	×	×
birweit_1	Birth weight (baby tail 1)	✓	✓
birweit_2	Birth weight (baby tail 2)	×	×
birweit_3	Birth weight (baby tail 3)	×	×
birweit_4	Birth weight (baby tail 4)	×	×
birweit_5	Birth weight (baby tail 5)	×	×
birweit_6	Birth weight (baby tail 6)	×	×
birweit_7	Birth weight (baby tail 7)	×	×
birweit_8	Birth weight (baby tail 8)	×	×
birweit_9	Birth weight (baby tail 9)	×	×
delmeth_1	Delivery method (baby tail 1)	✓	✓
delmeth_2	Delivery method (baby tail 2)	×	×
delmeth_3	Delivery method (baby tail 3)	×	×

Table F.1: (Continued.) Variables from Hospital Episode Statistics available for data linkage with the EPICure cohorts of 1995 and 2006.

Original HES variable name	Data description[a]	Availability 1995	Availability 2006
delmeth_4	Delivery method (baby tail 4)	×	×
delmeth_5	Delivery method (baby tail 5)	×	×
delmeth_6	Delivery method (baby tail 6)	×	×
delmeth_7	Delivery method (baby tail 7)	×	×
delmeth_8	Delivery method (baby tail 8)	×	×
delmeth_9	Delivery method (baby tail 9)	×	×
delplac_1	Delivery place (baby tail 1)	✓	✓
delplac_2	Delivery place (baby tail 2)	×	×
delplac_3	Delivery place (baby tail 3)	×	×
delplac_4	Delivery place (baby tail 4)	×	×
delplac_5	Delivery place (baby tail 5)	×	×
delplac_6	Delivery place (baby tail 6)	×	×
delplac_7	Delivery place (baby tail 7)	×	×
delplac_8	Delivery place (baby tail 8)	×	×
delplac_9	Delivery place (baby tail 9)	×	×
anasdate	First antenatal assessment date	✓	✓
anagest	First antenatal assessment gestation	✓	✓
gestat_1	Gestational age at delivery (baby tail 1)	✓	✓
gestat_2	Gestational age at delivery (baby tail 2)	×	×
gestat_3	Gestational age at delivery (baby tail 3)	×	×
gestat_4	Gestational age at delivery (baby tail 4)	×	×
gestat_5	Gestational age at delivery (baby tail 5)	×	×
gestat_6	Gestational age at delivery (baby tail 6)	×	×
gestat_7	Gestational age at delivery (baby tail 7)	×	×
gestat_8	Gestational age at delivery (baby tail 8)	×	×
gestat_9	Gestational age at delivery (baby tail 9)	×	×
birstat_1	Birth status (baby tail 1)	✓	✓
birstat_2	Birth status (baby tail 2)	×	×
birstat_3	Birth status (baby tail 3)	×	×
birstat_4	Birth status (baby tail 4)	×	×
birstat_5	Birth status (baby tail 5)	×	×
birstat_6	Birth status (baby tail 6)	×	×
birstat_7	Birth status (baby tail 7)	×	×
birstat_8	Birth status (baby tail 8)	×	×
birstat_9	Birth status (baby tail 9)	×	×
matage	Maternal age at delivery	✓	✓
motdob	Mother's date of birth	✓	✓
numbaby	Number of babies delivered (live and still born)	✓	✓
numpreg	Number of previous pregnancies	✓	✓
postdur	Postnatal stay duration	✓	✓
biresus_1	Resuscitation method (baby tail 1)	✓	✓
biresus_2	Resuscitation method (baby tail 2)	×	×
biresus_3	Resuscitation method (baby tail 3)	×	×
biresus_4	Resuscitation method (baby tail 4)	×	×
biresus_5	Resuscitation method (baby tail 5)	×	×
biresus_6	Resuscitation method (baby tail 6)	×	×

Table F.1: (Continued.) Variables from Hospital Episode Statistics available for data linkage with the EPICure cohorts of 1995 and 2006.

Original HES variable name	Data description[a]	Availability 1995	Availability 2006
biresus_7	Resuscitation method (baby tail 7)	×	×
biresus_8	Resuscitation method (baby tail 8)	×	×
biresus_9	Resuscitation method (baby tail 9)	×	×
sexbaby_1	Sex of baby (baby tail 1)	✓	✓
sexbaby_2	Sex of baby (baby tail 2)	×	×
sexbaby_3	Sex of baby (baby tail 3)	×	×
sexbaby_4	Sex of baby (baby tail 4)	×	×
sexbaby_5	Sex of baby (baby tail 5)	×	×
sexbaby_6	Sex of baby (baby tail 6)	×	×
sexbaby_7	Sex of baby (baby tail 7)	×	×
sexbaby_8	Sex of baby (baby tail 8)	×	×
sexbaby_9	Sex of baby (baby tail 9)	×	×
neocare	Level of neonatal care provided	×	✓
well_baby_ind	Well baby indicator flag	✓	✓
epikey	Record identifier	✓	✓

Notes:
[a] Taken from the HES data dictionary [123] unless otherwise indicated.

Appendix G

Systems Level Security Policy

Data with personal identifiable information were required for the investigations carried out as part of this thesis – particularly, for the Admissions Validation Study. Where possible, pseudonymous datasets were used for statistical analysis. At all stages, all data – including pseudonymised datasets – were stored on physical media that were encrypted at the level of the block device (e.g. hard drive or cdrom), with concurrent encryption of virtual (SWAP) memory using `cryptsetup` [192] packaged for Debian GNU/Linux. [193]

A systems-level security policy for the main computer located in the EPICure study offices at University College London was consequently produced for this, prior to commencing the research; this was approved by the National Information Governance Board in December 2011.[1] The approved version is reproduced on the following pages.

[1] Personal communication between Adam Goodwin, ECC Security Review team, and Rick Borges, NIGB Deputy Operations Manager at the NIGB Office, cc'd to myself on December 19th, 2011.

G. SYSTEMS LEVEL SECURITY POLICY

Institute for Women's Health

System Level Security Policy (SLSP)

EPICure Research System (ERS)

December 2011 Revision

Author:	Andrei Morgan
Created:	21/06/2011
Version:	1.1
Revised:	15/12/2011
Authorised by:	Neil Marlow
Implemented:	15/12/2011

List of individuals (15/12/2011):-

- Project Lead: Professor Neil Marlow (n.marlow@ucl.ac.uk 020 7679 6060)
- Head of Institute for Women's Health: Professor Peter Brocklehurst (p.brocklehurst@ucl.ac.uk 020 7679 6060)
- System Administrator: Dr. Andrei Morgan (andrei.morgan@ucl.ac.uk 020 7)
- UCL IS Network Group general: (nwg-ops@ucl.ac.uk 0207 679 7350)
- UCL Computer Security Team: (cert@ucl.ac.uk 020 7679 7338)

UCL is registered under the Data Protection Act for the purposes of research (Purpose 19):

- Registration number: Z6364106
- Expiry: 28 January 2012

The System shall be known as EPICure Research System (ERS, henceforth referred to as “The System”).

The System’s responsible owner shall be the Project Lead, Institute for Women’s Health, UCL.

The System’s Caldicott Guardian or Data Controller shall be Head of Department, Institute for Women’s Health, UCL.

Security of The System shall be governed by the corporate security policy of UCL (See Appendix A: UCL Information Security Policy; http://www.ucl.ac.uk/cert/swg/policy.html).

The System’s responsible security manager shall be the System Administrator, EPICure Research Group, Institute for Women’s Health, UCL. The System Administrator's duties shall include:-

- installing, maintaining and decommissioning all hardware excluding UCL-maintained switches and firewalls;
- implementing and updating operating systems, individual firewalls, required user applications, SQL databases;
- creating and maintaining user accounts for the ERS;
- assigning appropriate levels of access privileges;
- training staff: awareness of security and data protection;
- monitoring for breaches in ERS security.

The System shall be developed / provided by the System Administrator, EPICure Study Group, Institute for Women’s Health, UCL.

The System shall comprise a single workstation, implemented and maintained by the System Administrator following general principles laid out in UCL Security Policy and the Data Protection Act. All data on The System is stored securely using full disk-encryption with dm-crypt (with Linux Unified Key Setup - LUKS) – a UCL approved encryption software for GNU/Linux (see: http://www.ucl.ac.uk/isd/common/cst/good_practice/encrypt/encrypt/GuidanceStorageSensitiveData). The basic algorithm employed is built upon the Advanced Encryption Standard (AES) - a recognised standard that is continuously scrutinised and challenged by cryptologists around the world to ensure its strength and dependability - with a key size of 256 bits. A firewall is in place to prevent unauthorised network access to The System.

Personal data will be stored on a dedicated partition. Access to this data requires a password-protected dedicated workstation user account, only accessible via the workstation itself. Network access to the workstation has been completely disabled. Additionally, access to the databases containing identifiable personal data requires an additional (database) user logon and password (distinct from the user account workstation logon).

Any data stored on laptops, USB devices and other removable media will be anonymised or pseudonymised, and stored in encrypted containers in

accordance with the UCL guidance noted above.

Where equipment or media may contain personal data, following use the media will be either physically destroyed or the Gutmann method for destroying data will be utilised.

The System shall NOT be shared or used by any other organisations

System data may be obtained from the following organisations: consented trial volunteers; PCT/HA Age/Sex registers; NHS database extracts (NIGB; formerly NHS STS, HES [Hospital Episode Statistics] and ONS [Office of National Statistics]). Patient consent will be required for the disclosure of identifiable data or, where impracticable, permission will be sought from the Ethics & Confidentiality Committee of the NIGB.

Sensitive data may be transferred onto the system in one of two ways:

- (Indirect) transfer via physical media, e.g. CD or DVD; external storage device connected via USB. All media used for data transfer will be required to be encrypted in line with UCL Guidance on the Storage of Sensitive Data on Portable Devices and Media. Portable media will have all sensitive data securely deleted immediately following last use or, where this is not possible, will be securely destroyed.
- Network-based (direct) transfer: data may be downloaded onto The System via the internet. Sensitive data will only be downloaded from authenticated websites using Secure Socket Layer (SSL) encrypted connections; where possible, manual cryptographic checking of data integrity will also be carried out at the time of transfer.

The System shall process Clinical Research Trials data including:

- EPICure: Outcomes for births before 26 weeks gestation in the United Kingdom and Republic of Ireland (1995);
- EPICure-2: Outcomes for births before 27 weeks gestatation in England (2006);
- UK Neonatal Staffing Study (1996);
- Unit Profile Study (2006 and 2011);

as well as any other datasets required for those studies. Data analyses will be conducted using the smallest amount of sensitive data necessary.

Data collected may include:

- Personal details
- Medical details
- Family History details
- Social History details
- Surgical and Histopathology details
- Death certificate details

The System's authorised purpose(s) shall be the management, implementation and analysis of neonatal research trials. All trials managed on

the system have received ethical approval and have been registered with the UCL Data Protection Office (See Appendix B: UCL Data Protection Policy). UCL is currently registered under the Data Protection Act for the purposes of Research (Purpose 19); registration number: Z6364106; expiry: 28th January 2012.

The System's authorised users shall be members of the EPICure Study Group from the Institute for Women's Health, UCL, as well as other members of the EPICure Study Group as delegated by the Project Lead. In all instances, users will need to be present at the Institute for Women's Health in order to access The System.

Wireless access and remote usage of The System will not be permitted.

The System shall be risk assessed on an annual basis (See Appendix C: Risk Assessment Checklist For System Custodians). A risk management/security improvement plan shall be established to address all unacceptable risks.

The System shall benefit from the following security countermeasures:

- Physical Measures
  - All system hardware is located on 4th floor, Rockefeller Building (21 University Street). Access to the locked offices requires ID badge at manned reception and key entry to individual offices.
  - The System is configured to shut-down upon power failure.
  - Where required, data containing PID (personal identifiers) will be stored in encrypted containers – LUKS default (AES-CBC-ESSIV) algorithm, key size 256 bits.
- Logical Measures
  - Access requires user account and passwords.
  - Passwords expire every 6 months;
  - Passwords are subject to "John-the-Ripper" analysis.
  - Re-use of passwords is forbidden
  - User and database access passwords must meet the following criteria:
    - Contain characters from three of the following four categories
      - Uppercase alphabetic characters (A-Z)
      - Lowercase alphabetic characters (a-z)
      - Arabic numerals (0-9)
      - Nonalphanumeric characters, for example: ! $ # %
    - Not be based on the username.
    - Be at least 8 characters long.
    - Must not repeat any of your last 4 passwords.
    - Not contain spaces or commas.
  - System encryption passphrases must meet the following criteria:
    - All the criteria listed above under the "user and database access passwords criteria" section.

        - Be at least 25 characters in long

- Operating System
    - Debian GNU/Linux 'Stable' (currently Debian “Squeeze”, version 6.0.3)
- Firewall Software
    - Linux kernel 'netfilter/iptables' rules employed on The System to prevent unauthorised network access via the network intervace.
    - UCL Institutional Firewall separates UCL network from JANET network and the rest of the internet.
- Application level software
    - Only cryptographically verified free software from official Debian 'stable' repositories are permitted to be installed on The System.

The System shall benefit from the following internal/external audit arrangements:

- Monthly audit of user accounts
- Monitoring of logon success and failure events
- UCL Computer Security Department constantly monitors network for vulnerable machines.

The System shall benefit from the following resilience / contingency / disaster recovery arrangements:

- Backup Measures:
    - A full, encrypted data backup to an external hard drive will be initiated every time The System is booted.
    - Data will then be backed up incrementally while The System is operational. A version control system ("git") will be used to track changes to the data.
    - At the end of each work session, the external hard drive will be detached from the workstation and stored separately.
- Hardware failure
    - Incremental backups will be performed whenever The System is in use. Between these backups, the external hard drive will be unmounted from the main computer.
    - External hard drives will be rotated daily. It will be possible to recover a minimum of one (1) weeks data.

In the event of serious disruption or total system failure, research continuity shall be provided by the following means:

- Restoration of the ERS will be performed from the most recent, uncorrupted backup onto a newly provisioned machine. Estimate for purchase and restoration of essential hardware is 1-2 weeks. No data is considered "essential" to ongoing patient care and thus there will not be any deleterious effects from system down-time.

In the event of a security or confidentiality breach occurring the following procedure shall be followed:

- System Administrator, Project Lead, Senior Software Analyst or Head of Department will contact UCL Computer Security Team (cert@ucl.ac.uk; 020 7679 7338) (See Appendix D: User Guide to UCL Information Security Policy)

When the system or its data has completed its purpose/has become redundant or is no longer needed, the following methods will be adopted to dispose of equipment, back-up media or other stored data:

- Any research data still required will be stored according to the original protocol specified.
- The Gutmann method will be utilized for destroying data on workstations, external hard drives and any other media used.

This SLSP shall be the responsibility of the ERS System Administrator and shall be reviewed on an annual basis for its completeness and for relevant update.

The SLSP shall be available/distributed to all individuals mentioned in this document via e-mail attachment. The document will be stored separately by those groups and individuals named within.