MODELLING NEONATAL CARE PATHWAYS FOR BABIES BORN VERY PRETERM

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

Sarah Emma Seaton BSc MSc

Department of Health Sciences

University of Leicester

January 2018

ABSTRACT: MODELLING NEONATAL CARE PATHWAYS FOR BABIES BORN VERY PRETERM.

SARAH E SEATON.

Predicting length of stay in neonatal care is important for resource planning and the counselling of parents. However, it has received limited attention and two issues are:

1. Babies who die in neonatal care are not included appropriately and research should consider all babies simultaneously, irrespective of whether they live or die
2. The different levels of neonatal care (intensive, high dependency and special care) and how they contribute towards overall length of stay have not been considered

This thesis contains four inter-connected studies to investigate how statistical approaches can help to address these issues.

Firstly, a systematic review was conducted to identify factors commonly used to predict length of stay and mortality. Factors measurable at or around birth, such as gestational age and birthweight, were found to be important.

Secondly, competing risks methods were used to predict median length of stay in neonatal care for two competing events: babies who survive to discharge and babies who die before discharge. These estimates can be used by clinicians, with their clinical judgement, to counsel parents about the risk of mortality and about potential length of stay.

The third study develops this approach to account for the different levels of care received by the baby, using multistate modelling as a natural extension of the more limited competing risks approach. Mean lengths of stay at each level of care were estimated in order to facilitate commissioning of neonatal services.

Finally, the differences in length of stay between Operational Delivery Networks, (groups of neonatal units that work together) were investigated to determine if differences existed. These were examined to understand whether differences were due to varying levels of intensity of specific levels of care within a network or a difference in total length of stay.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my supervisors: Bradley Manktelow, Elizabeth Draper and Keith Abrams. Thank you for allowing me the independence I wanted to pursue my own research ideas, even those ideas that resulted in dead-ends, but the friendship, guidance and support that I needed when things went wrong.

This work was funded by a Doctoral Research Fellowship from the National Institute for Health Research (DRF-2013-06-011). The data for this study was supplied by the Neonatal Data Analysis Unit and I would like to particularly thank: Neena Modi, Richard Colquhoun and Eugene Statnikov. All neonatal units in England agreed for me to use their data in this study, and for that I was both overwhelmed and incredibly grateful, I hope the results of this work interest you. Thank you to Lisa Barker for providing clinical input, and David Jenkins for helping with my systematic review.

I was wonderfully supported throughout this work by the team at Bliss, especially Mehali Patel and Caroline Lee-Davey. Alongside Bliss, I was assisted by some wonderful parents who provided insight, reminded me of what mattered and shared their experiences of neonatal care so willingly.

This thesis would not have been possible without the support of so many people. Over the years my colleagues were not always my priority and I hope you never felt too neglected – thank you for helping and for putting up with me. My friends and family have been selflessly supportive throughout this work, often providing a listening ear when they had no idea what I was talking about – I saw your glazed eyes and I forgive you – thank you.

I must specifically thank Matt for proofreading and his support. Finally, I owe a great debt to my friends Alun and Kate: for (occasionally) knowing what I was talking about; for helping hunt out rogue commas in my code and, above all, for supporting me to get to this stage.

This thesis is for Hugo Christopher Dylan Parker, who lived only 35 days but whose legacy lives on, and his parents: Leigh Kendall and Martin, who have been tireless supporters of me and this work.

TABLE OF CONTENTS

TABLE OF ABBREVIATIONS

Acronym	Definition
AUC	Area Under Curve
ANN	Artificial Neural Network
AIC	Akaike Information Criterion
BAPM	British Association of Perinatal Medicine
BPD	Bronchopulmonary Dysplasia
CPAP	Continuous Positive Airways Pressure
EDD	Estimated Delivery Date
HD	High Dependency
HL	Hosmer-Lemeshow
HR	Hazard Ratio
IC	Intensive Care
MAIN	Morbidity Assessment Index for Newborns
NDAU	Neonatal Data Analysis Unit
NEC	Necrotising Enterocolitis
NICU	Neonatal Intensive Care Unit
NNRD	National Neonatal Research Database
NRES	National Research Ethics Service
OECD	Organisation for Economic Co-operation and Development
ODN	Operational Delivery Network
PMA	Post-Menstrual Age
QUIPS	Quality In Prognostic Studies [tool]
R&D	Research and Development
SC	Special Care
SGA	Small for Gestational Age
SNAP	Score for Neonatal Acute Physiology
SNAPPE	Score for Neonatal Acute Physiology – Perinatal Extension
TPN	Total Parenteral Nutrition

1 INTRODUCTION

1.1 PREDICTING LENGTH OF STAY IN NEONATAL CARE

Neonatal care is a highly technical and expensive speciality of medicine, with neonatal care costing approximately £420 million in England in 2006/07 (1). The total cost of very preterm birth and associated public sector costs, including education services, was estimated in 2009 to be £989 million in England and Wales (2). Babies born very preterm, defined as less than 32 weeks gestational age, and others requiring in-patient neonatal care are admitted to a neonatal unit following their birth. Around one in eight babies are admitted to neonatal care (3) and almost all babies born at less than 32 weeks gestational age who survive the initial period after birth receive care in a neonatal unit (4).

The length of stay of babies admitted to a neonatal unit can vary dramatically. Babies born at term, near their due date, form a heterogeneous group in terms of their care needs (5). Some may need a small amount of time in the neonatal unit, such as a few hours of monitoring. Others may have profound issues such as needing cardiac surgery for a heart defect. Babies born very preterm who survive often require respiratory support and their hospital stay is overwhelmingly driven by their prematurity. These babies will often need weeks to several months of specialist neonatal care (6).

There is currently little published research investigating how to estimate length of stay in neonatal care and this would be valuable to individual neonatal units and the healthcare service to aid the commissioning and allocation of resources. This has become increasingly important in recent years with major budget restrictions as healthcare services struggle more than ever to plan limited resources including staffing and availability of cots whilst also ensuring appropriate funding and provision of services (7).

Investigation of neonatal length of stay would also provide detail about the natural history of neonatal care provision which could be used to evaluate new service configurations and technologies. However, for commissioning and funding,

information about length of stay in neonatal care alone does not provide enough detail. Different levels of care exist within neonatal medicine, each with very different costs, staffing and resource needs. Estimates of length of stay would be more informative if they could be broken down to provide information about the levels of care required.

Beyond the benefit that estimating length of stay would have for the commissioning of services, the use of estimates would help parents and families to understand the potential care pathway for their baby. The impact of having a baby requiring specialist neonatal care on parents and families should not be underestimated, with parents frequently asking healthcare professionals, *"When can my baby come home?"* throughout their baby's stay in neonatal care. Currently, answering this question can be difficult. Anecdotally, rough estimates are given such as *"They'll be home by their due date"* or *"They'll go home once they can keep themselves warm and feed."* This may be a good estimate, with many surviving babies probably being discharged home around the date they were due to be born. However, it is unclear if this is true and estimates of length of stay would provide more accurate information for counselling parents and preparing them for the possibility that their baby may be in hospital for a long time.

The estimation of length of stay in neonatal care is not straightforward for two reasons: 1) babies who die in neonatal care are often excluded or included without appropriate adjustment and it is unclear how to account for this; 2) current estimates of length of stay do not consider the levels of care a baby requires whilst in the neonatal unit. To date the limited research which has been undertaken to provide current estimates is not clinically informative. Without consideration of these issues, it is impossible to fully understand the needs of these babies and the impact their stay in neonatal care has on both the healthcare service and their families.

1.2 ADMISSION TO THE NEONATAL UNIT

1.2.1 NEONATAL CARE ADMISSION AND PRETERM BIRTH

The World Health Organisation refers to a birth before 28 weeks of gestation as extremely preterm; from 28 to 31 weeks as very preterm; from 32 to 36 weeks as moderate to late preterm and births after 37 weeks as being term (Table 1-1) (8). Gestational age is conventionally reported in weeks^{+days}, although the days are often not denoted when completed weeks are used, as is the case in this thesis. There are known to be minor inaccuracies in the calculation of the expected date of delivery due to biological variability in the timing of the fertilisation of the egg and implantation of the blastocyst (the inner layer of the blastocyst forms the embryo, whilst the outer layer forms the placenta) (9). Therefore, gestational age is measured in completed weeks in this thesis as the measurement of weeks and days is likely to be inaccurate.

Around 60 percent of babies admitted to neonatal units are born at term (3) and there are a variety of reasons that these babies are admitted, such as:

- Respiratory problems including: pneumonia, chronic lung disease and respiratory distress syndrome.
- Infections including: sepsis, meningitis and group B streptococcus.
- Hypoglycaemia (low blood sugar).
- Jaundice (build-up of the waste product bilirubin in the blood).
- Asphyxia (deprived of oxygen at birth).

Another group of babies admitted for neonatal care are those born preterm which is defined as any birth which occurs before 37 weeks of pregnancy and many admissions born before 32 weeks gestational age. These babies are generally admitted to neonatal care for reasons of prematurity and predominantly need respiratory support, assistance maintaining their body temperature and help with feeding. Very preterm babies who survive have the longest length of stay and as such are the focus of this thesis. The phrase "**very preterm**" will be used to describe all babies born at 24 to 31 weeks gestational age throughout this thesis.

The term "**post-menstrual age**" (PMA) refers to the gestational age that the baby would now be if the pregnancy had continued (9). For example, a baby born at 31 completed weeks of gestational age who is now two weeks old is now 33 weeks PMA.

Table 1-1: Definitions of preterm birth according to the World Health Organisation.

Weeks of gestational age	**Category**
Less than 28 weeks	Extremely preterm
28 to 31 weeks	Very preterm
32 to 36 weeks	Moderate to late preterm
37 weeks and above	Term

Wide variation exists in the survival and management of babies born at 22 to 24 weeks gestational age between countries (10) and individual centres (11). There are suggestions that for babies born at less than 25 weeks, birthweight should also be considered alongside gestational age for making decisions about intensive treatment (10).

In the United Kingdom, there is no legal requirement to register fetal deaths (i.e. babies that die before birth) that occur before 24 weeks of pregnancy and there are regional differences in whether births are reported as live born (12). Active resuscitation of babies born at 24 weeks is recommended (13). Therefore, this thesis investigates babies born at and after 24 weeks gestational age, where legal registration of all births is required.

Survival of preterm babies has improved, particularly over the last 20 years, as seen in Table 1-2 which is reproduced from Manktelow, Seaton et al (14). This decline in mortality has been seen following the introduction of routine use of surfactant (15) and the development and improvement of neonatal intensive care services (16, 17).

1.2.2 Deaths in neonatal care and the calculation of length of stay

Historically there has been limited research that has investigated length of stay within neonatal care. Research that has been undertaken has focussed on babies who survive to discharge from neonatal care, i.e. excluding those who die during their stay (18). This focus is likely to be because babies who survive to discharge account for the

majority of neonatal care workload and because parents are interested in time to discharge home.

Babies who die often only live for a short time, which is very different from the surviving babies who can require many months of care in hospital. Very preterm babies can have a high rate of in-unit mortality, particularly for those born at less than 28 weeks gestational age (14) and any analysis which is only based on survivors does not fully describe neonatal care requirements particularly when considering workload. Often it is statistically easier to exclude deaths as the distributions of the time spent in neonatal care are very different for babies who die compared to babies who survive, and combining these groups creates a distribution which is complex to model. Alternatively, treating all babies the same, irrespective of whether they live or die, in an analysis can introduce bias to the estimate of length of stay due to the different distributions of their lengths of stay (19).

Table 1-2: Observed percentage of singleton White babies admitted to neonatal care surviving in 1994 to 1997 and 2008 to 2010 in the East Midlands and Yorkshire (reproduced from Manktelow, Seaton et al (14)).

Gestational age (weeks)	**1994-1997**		**2008-2010**	
	Survival % (male)	**Survival % (female)**	**Survival % (male)**	**Survival % (female)**
23	18.7	13.0	28.6	34.8
24	20.0	17.9	48.1	55.6
25	45.1	43.6	73.4	67.2
26	55.7	65.1	77.4	83.6
27	75.8	82.3	83.3	90.0
28	80.4	82.1	88.7	93.1
29	89.5	94.8	93.0	96.1
30	95.0	95.5	95.3	98.9
31	97.6	98.4	98.8	98.6
32	97.8	98.5	98.8	98.6
Overall	85.5	86.9	91.2	92.6

1.2.3 OUTCOMES AFTER PRETERM BIRTH

Preterm birth has both short and long-term complications and consequences. Whilst survival has increased for babies born at extremely preterm gestational ages (Table 1-2), the occurrence of morbidities has remained the same with only around 41% being discharged from hospital with no major morbidity (20). Of babies born at 22 to 26 weeks gestational age in the EPICure study (a study based in England to measure survival and morbidity after extremely preterm birth) who survived to 36 weeks PMA, 22% had treatment for retinopathy of prematurity, 61% had severe bronchopulmonary dysplasia and 13% had a severe abnormality recorded on a cerebral ultrasonography. At three years of age the surviving children had increased risk of developmental or cognitive impairment (21).

1.3 LEVELS OF NEONATAL CARE

1.3.1 LEVEL OF CARE GIVEN TO THE BABY

Whilst the overall estimation of length of stay in the neonatal unit is of interest it does not provide a full picture of care, even if babies who die and babies who survive are both considered. Within neonatal medicine there are different levels of care that a baby might require. Each level of care has its own associated costs and impact for both families and the NHS. For example, very preterm babies are at the highest risk of needing intensive care and these babies are also at the highest risk of mortality.

The current definitions of the levels of care within neonatal medicine were defined by the British Association of Perinatal Medicine (BAPM) in 2011 (22). Broadly, there are currently three levels of care provision on a neonatal unit: intensive care; high dependency care and special care. Currently these are defined by the type of treatment required by a baby which differs from older definitions that related to their physical location or demographic characteristics (23). For each calendar day of care, each baby is classified to the highest level of care they received, even if they receive a lower level of care for part of the day. Staffing levels for each level of care are specified by NHS England (23) to ensure that the provision of neonatal care relies *"on having an adequate and appropriate workforce."*

Intensive care is given to the sickest or most unstable babies. Typically it involves the highest demands from staff and NHS England recommends a nurse to baby ratio of 1:1 (24). Interventions classified as intensive care include:

- Mechanical ventilation via a tracheal tube.
- Dialysis of any kind.
- Presence of a chest drain.
- Presence of a silo for gastroschisis (covering of exposed abdominal organs).
- Presence of replogle tube (a tube used to drain saliva).
- Epidural catheter (pain relief placed in around the spinal cord).
- Exchange transfusion (removal of blood replaced with donated blood).
- Therapeutic hypothermia (lowering of temperature to reduce tissue injury from lack of blood flow).
- Receiving non-invasive ventilation (e.g. nasal continuous positive airways pressure (CPAP) therapy or oxygen) and parenteral nutrition (intravenous feeding that bypasses the process of eating and digestion).

The next level of care is **high dependency care,** which still requires highly skilled staff, but at a lower recommended ratio of one nurse to two babies. Examples of high dependency care include:

- Receiving non-invasive ventilation without parenteral nutrition.
- Parenteral nutrition (intravenous feeding that bypasses the process of eating and digestion).
- Tracheostomy (an opening in the neck to insert a tube to assist with breathing).
- Continuous infusion of drugs.
- Barrier nursing (infection control used to protect staff and the baby from infection).

Babies who require additional care that is not considered intensive care or high dependency care will generally receive **special care**, for which a nurse to baby ratio of 1:4 is recommended. Examples include:

- Nasal oxygen.
- Phototherapy (light therapy typically used to treat jaundice).
- Four hourly observations.
- Stoma in situ (an artificial opening into the surface of the abdomen).

In addition to the three levels of neonatal care some neonatal units also offer **transitional care**, where the mother provides the majority of the care within hospital with support from a healthcare professional such as a midwife. Examples of babies who may be offered transitional care include:

- Babies requiring antibiotics.
- Low birthweight babies who are otherwise healthy.

This level of care is offered in different physical locations, with some hospitals requiring an admission to the neonatal unit and others offering it on the postnatal ward. A recent survey in England highlighted the inconsistency of provision of this care with only 50% of units having clarity around the funding of this care (25). Throughout this thesis recorded transitional care is amalgamated with special care, which accommodates similar needs of a baby.

During their time in the neonatal unit, a baby can receive different levels of care until discharge from the unit or dying whilst in hospital. In this thesis this series of movements between the different levels of care is defined as the **neonatal care pathway**. In this work, the data related to neonatal care is manipulated so that the *neonatal care pathway* occurs hierarchically, i.e. all intensive care is considered first, followed by all high dependency care and then all special care.

1.3.2 NEONATAL UNIT LEVELS OF CARE

In addition to the levels of care provided to the individual baby, which do not depend on a physical location, neonatal units are designated at different levels. The different types of unit are defined as:

- **Special Care Baby Unit:** Level 1, a unit which can provide only special care.

- **Local Neonatal Unit:** Level 2, a unit that can provide high dependency care and short periods of intensive care, generally to babies born after 27 weeks of gestation.
- **Network Neonatal Unit:** Level 3, also known as Neonatal Intensive Care Unit, a unit able to provide all levels of care.

Neonatal units are organised into **Operational Delivery Networks** (ODNs), which are groups of units within a geographical region that work together to provide the full range of neonatal services (26). Every ODN will have at least one Neonatal Intensive Care Unit, and can therefore provide intensive care to any individual baby in a given geographical region. All neonatal units in England that provided data to this study are presented in Figure 1-1, with the designated level of the unit and the ODN they belong to represented.

1.4 AIMS OF THIS THESIS

This thesis has four proposed aims:

1. To identify the factors that predict mortality and length of stay in the neonatal unit, focussing on babies born very preterm (less than 32 weeks gestational age).
2. To investigate the length of stay of very preterm babies (24 to 31 weeks gestational age) admitted for neonatal care to inform parental counselling regarding the risk of mortality and length of stay for a baby of given characteristics.
3. To examine the different levels of care, the neonatal care pathway, required whilst a very preterm baby (24 to 31 weeks gestational age) is in neonatal care to inform commissioning of specialist neonatal services by providing estimates of the number of days of levels of care required.
4. To compare the levels of care provided by different Operational Delivery Networks to babies born at 24 to 31 weeks gestational age in order to identify and investigate differences in care provision.

1.5 OVERVIEW OF PHD THESIS

This chapter has provided an overview of the background to the risk of mortality and the estimation of length of stay within neonatal care. The issues with current research have been identified and the aims of this thesis have been outlined.

Chapter 2 will provide a systematic review of the literature for the prediction of mortality and length of stay for very preterm babies in neonatal care. This systematic review was undertaken to identify the factors that are important to include in an analysis that involves prediction of mortality or length of stay. This review focusses on length of stay and addresses the first aim of this thesis. This work was subsequently published in *BMJ Open* (18) and a copy of the article can be found in Appendix 2.

Chapter 3 introduces the National Neonatal Research Database (NNRD) from which data were extracted for use throughout this thesis. Data from all babies born at 24 to 31 weeks gestational age and admitted to neonatal units in England were selected. These data are used throughout the thesis to illustrate the methods and to investigate length of stay and levels of neonatal care required. The NNRD has approval to be used for service evaluation and its use for research requires new approvals. In addition to those approvals, agreement was also required from all 162 neonatal units in England that contributed data to the NNRD. The process of obtaining ethical approval, R&D and neonatal unit agreement is described in Chapter 3.

Survival analysis, the foundation of all methodologies used in this thesis, is introduced in Chapter 4. This approach allows the measurement of time until an event of interest occurs. The Cox proportional hazards model is introduced and a preliminary analysis is undertaken. Although this method allows for the estimation of length of stay in a neonatal unit, it only allows consideration of one endpoint and is therefore unable to distinguish between babies who die in neonatal care and those who survive to discharge. To model these two endpoints simultaneously an extension of survival analysis known as competing risks is introduced in Chapter 5. Competing risks methodology, which is already well established, allows for two or more endpoints where the occurrence of one prevents the other(s) from happening. In this thesis the competing events are death in neonatal care and discharge. The methodology for

competing risks is introduced by extending the Cox model outlined in Chapter 4. An alternative approach of the flexible parametric modelling is explored in Chapter 5. The strengths and limitations of these approaches are discussed before a final model is presented to address the second aim of this thesis.

The extension from survival analysis to competing risks analysis allows the modelling of length of stay for all babies irrespective of their outcome. However, this still provides little detail on the time spent in hospital, which is often a prolonged period of time for very preterm babies who survive their neonatal stay. To investigate this further, an extension of competing risks known as multistate modelling is introduced in Chapter 6. This approach allows the modelling of 'intermediate' events before death or discharge. For example, the different levels of care: intensive care; high dependency care and special care could be included as the intermediate events. The strengths and limitations of this approach are discussed before a final model is proposed to address the third aim of this thesis.

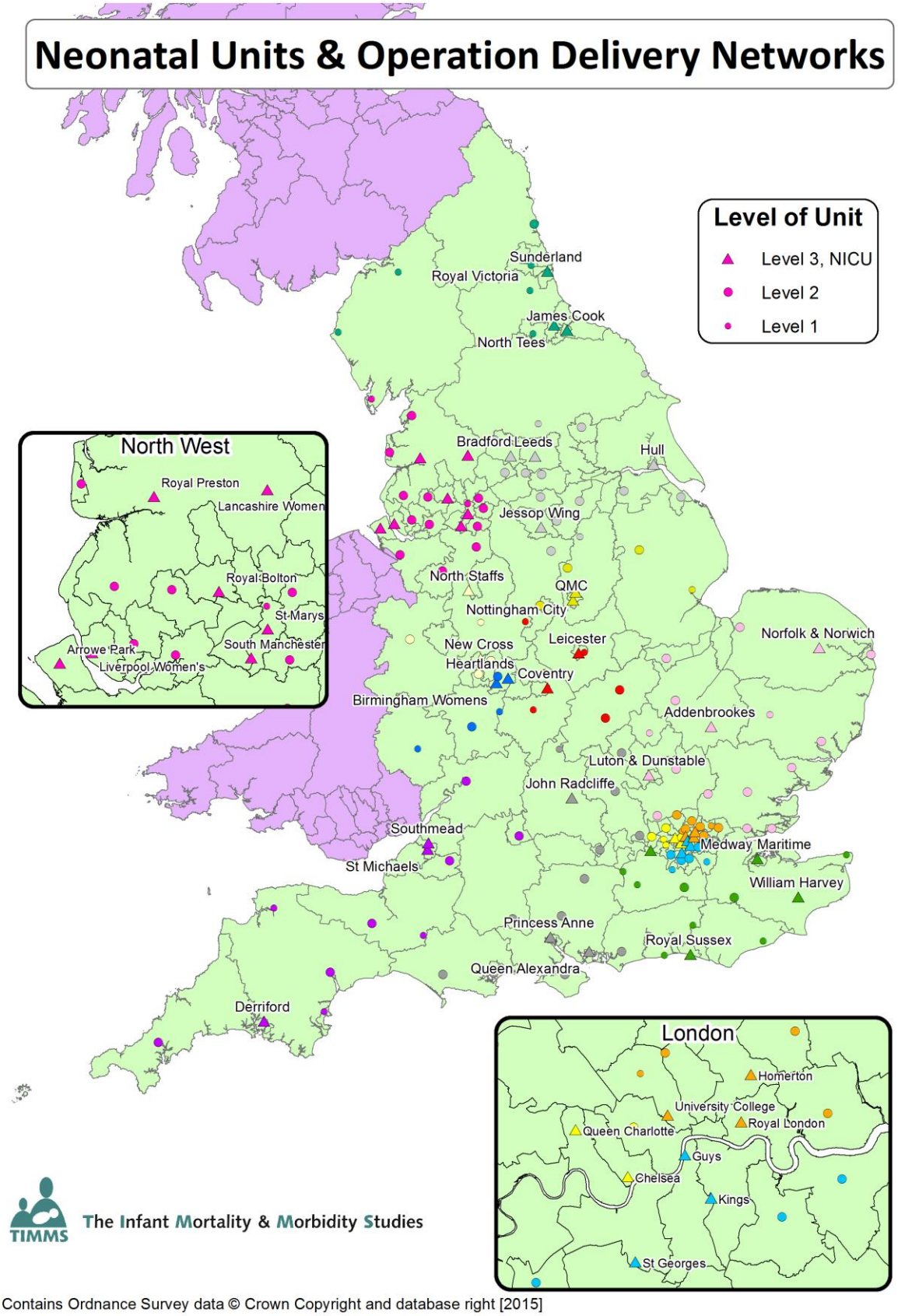
Details of the observed data related to levels of care required by very preterm babies has never been investigated before and therefore an article was written for clinicians and published in *Infant* (27). Results from Chapter 6 have also been published in *PLOS ONE* (28). Copies of both these articles can be found in Appendix 6.

Chapter 7 provides a comparison of two ODNs to the rest of England to demonstrate the use of multistate modelling to aid investigation of both expected length of stay in neonatal care and the differences in the provision of the different levels of care. This analysis addresses the fourth and final aim of the thesis. Results can be used by clinicians and commissioners to consider how care in one ODN differs, and whether this difference is meaningful or important.

A critical discussion of the methods and results in this thesis is provided in Chapter 8. The strengths and limitations are discussed and the results are critiqued. The clinical impact of the results from this thesis are presented, along with plans for future research before final conclusions are then drawn.

Throughout the undertaking of this thesis, a variety of dissemination and public involvement activities have been undertaken. These activities and the impact of this thesis are presented in the Appendix 8.

Figure 1-1: Map of neonatal units in England with the Operational Delivery Network indicated by the colour and the level of unit by the shape of the marker.



2 SYSTEMATIC REVIEW: FACTORS THAT PREDICT NEONATAL MORTALITY AND LENGTH OF STAY

2.1 CHAPTER OVERVIEW

The ability to predict length of stay and mortality in neonatal care is important for healthcare resource planning and to aid counselling parents. Whilst some research has investigated the prediction of mortality in this population, limited research has investigated what predicts length of stay in the neonatal unit. It is likely that factors that predict length of stay may also predict the risk of mortality for some groups of babies. This chapter presents a systematic review of the research which has been undertaken to predict mortality and/or length of stay. The analyses undertaken by individual studies will be investigated and details of the factors included within the analyses of those studies will be extracted and discussed. Factors that are considered important, defined as those included by multiple research studies, will be used to inform the analyses in later chapters of this thesis.

The length of stay element of this review was published in *BMJ Open* (18) and a copy of the article is provided in Appendix 2.

2.2 INTRODUCTION

Modelling the neonatal care pathway, including the need for different levels of care and the risk of mortality, is a complex issue. All analyses require appropriate adjustments to be made to account for all factors which may impact on the outcome, such as characteristics of the baby, to ensure appropriate estimation of survival and length of stay. It is also important to balance this with the analyses needing to be medically relevant and useful with the need for simplicity if used in a clinical setting. Statistical models that account for a small number of factors, which can be objectively and routinely measured, potentially offer a more informative use in a clinical setting (29). This also reflects the concept of statistical parsimony (using the simplest statistical model with the least variables to explain the outcome). In some areas, such

as the prediction of mortality, complex risk scores such as the Score for Neonatal Acute Physiology (SNAP) which contained 34 items have been simplified to allow ease of use (30, 31). This indicates the importance of conducting the simplest appropriate analysis and adjustments to allow a score to be clinically relevant and useful. The findings of this review will be used to make appropriate adjustments throughout this thesis.

2.2.1 AIMS AND OBJECTIVES OF THIS REVIEW

The aim of this systematic review was to investigate the factors used in previous research, usually within a multivariable statistical model, to predict neonatal length of stay or mortality. This review was registered with PROSPERO (32), the international prospective register of systematic reviews (registration number: CRD42013006020).

The objectives of this review were to:

1. Undertake a systematic search strategy to identify all literature investigating the prediction of neonatal mortality and/or length of stay in the neonatal unit.
2. Extract details related to the variables included in the final multivariable analysis of each of these studies to identify important factors.
3. Critically appraise the studies to identify potential sources of bias and poor study quality.
4. Select variables to be used for the analysis in this thesis.

2.3 METHODOLOGY OF THE SYSTEMATIC REVIEW

2.3.1 SEARCH STRATEGY

Medline, Embase and Scopus were searched systematically for peer-reviewed articles published from 1 January 1994 to 31 Dec 2013, with a subsequent update to 31 May 2016. Small variations of the same search strategy were used for each database, depending on the subject headings of that database. The search was focussed on four issues: consideration of the patient (i.e. the preterm baby); the setting (neonatal unit); the outcome (mortality or length of stay) and the type of analysis of interest (prediction model). This is similar to the PICO (P: population; I: intervention; C: comparison; O: outcome) approach recommended for formulating research questions

(33). The full search strategy is provided in Table 2-1. Reference lists of articles included in the review were scrutinised to identify other potentially relevant articles and these were added to the review ('hand searching') as appropriate.

2.3.2 INCLUSION CRITERIA

Studies were included that reported risk factors for length of stay in the neonatal unit or mortality, using a multivariable model (for example: logistic regression, linear regression). The statistical significance of the risk factors included in the model by study authors was not considered. The population of interest was very preterm babies or very low birthweight babies (often used by American studies in lieu of a measure of gestational age) as these had the highest risk of mortality and were likely to have longer lengths of stay. However, if the study population included both very preterm births and term births then they were also included in the research.

Survival of preterm babies improved dramatically from the 1980s to the early 1990s with the increasing use of antenatal steroids and surfactant and therefore 1994 was chosen for the start of the search period (15). Prior to 1994, predictors were likely to be different, or their effect would have now changed due to improvements in medical intervention. Studies using data from prior to 1994 were excluded whilst those with data that encompassed the time period (i.e. included data from pre-1994 and later) were included. Finally, to be included studies needed to have been undertaken in a human population and have been published in English.

2.3.3 EXCLUSION CRITERIA

Exclusion criteria were determined in advance and comprised:

- Non peer-reviewed conference proceedings, although efforts were made to investigate if the work was subsequently written into a peer reviewed publication.
- Articles not containing original research: reviews, letters and editorials.
- Countries which were not members of the Organisation for Economic Co-operation and Development in 1994 (34), as these countries were likely to have very different healthcare systems to those of developed nations.

- Clinical trials, as the selection of the trial population would be unlikely to be representative of all babies in neonatal care.
- Inappropriate study population, for example investigation of a paediatric or maternal population, or inappropriate outcome, for example predicting readmission to hospital.
- Specific disease areas (for example: E-Coli outbreaks or infections) as these babies have very different risk profiles to other babies in neonatal care.
- Work that was subsequently updated or validation studies.

Titles, abstracts and full articles were assessed to see if the study met the inclusion/exclusion criteria. Data related to the overall study and variables included in author's statistical analyses were extracted using a pre-designed form (Appendix 2). A random 10% subset of this review, focussing on the length of stay aspects, was screened by a second reviewer (David Jenkins, a statistician in the Department of Health Sciences, University of Leicester) to ensure papers were selected systematically.

2.3.4 DATA EXTRACTION

Each article was screened and, for those in scope, the data were extracted related to which factors had been accounted for in the final multivariable analyses predicting length of stay or mortality. Identified factors were grouped together to form similarly themed groups which were defined as:

- **Inherent factors**: details about the baby, typically known at birth, which are fixed, for example, gestational age and birthweight.
- **Antenatal treatment or maternal factors**: details about the baby or the mother known prior to birth, for example, whether antenatal steroids were administered.
- **Conditions of the baby**: medical concerns or symptoms being experienced by the baby including their blood pressure and body temperature.
- **Treatment of the baby**: the care or treatment which the baby was given including surgery.

- **Organisational factors**: the hospital of care and organisational issues, such as any transfer between units experienced by the baby.

The categories were defined *a priori* and where it was unclear which category a risk factor belonged to, a discussion was held with the supervisors of this thesis until consensus was reached.

2.3.5 ASSESSMENT OF STUDY QUALITY

Prognostic studies and observational research are often considered to be poor quality, and prone to bias, mainly due to the study design rather than systematic issues introduced by the research team within the study set-up (35). No studies were excluded for poor study quality although quality was considered using the areas of: study participation; study attrition; measurement of predictors; outcome measurement and statistical analysis and reporting. This is a framework based on the Quality in Prognostic Studies (QUIPS) tool although no formal scoring was undertaken (36).

2.4 RESULTS OF SYSTEMATIC REVIEW: OVERVIEW

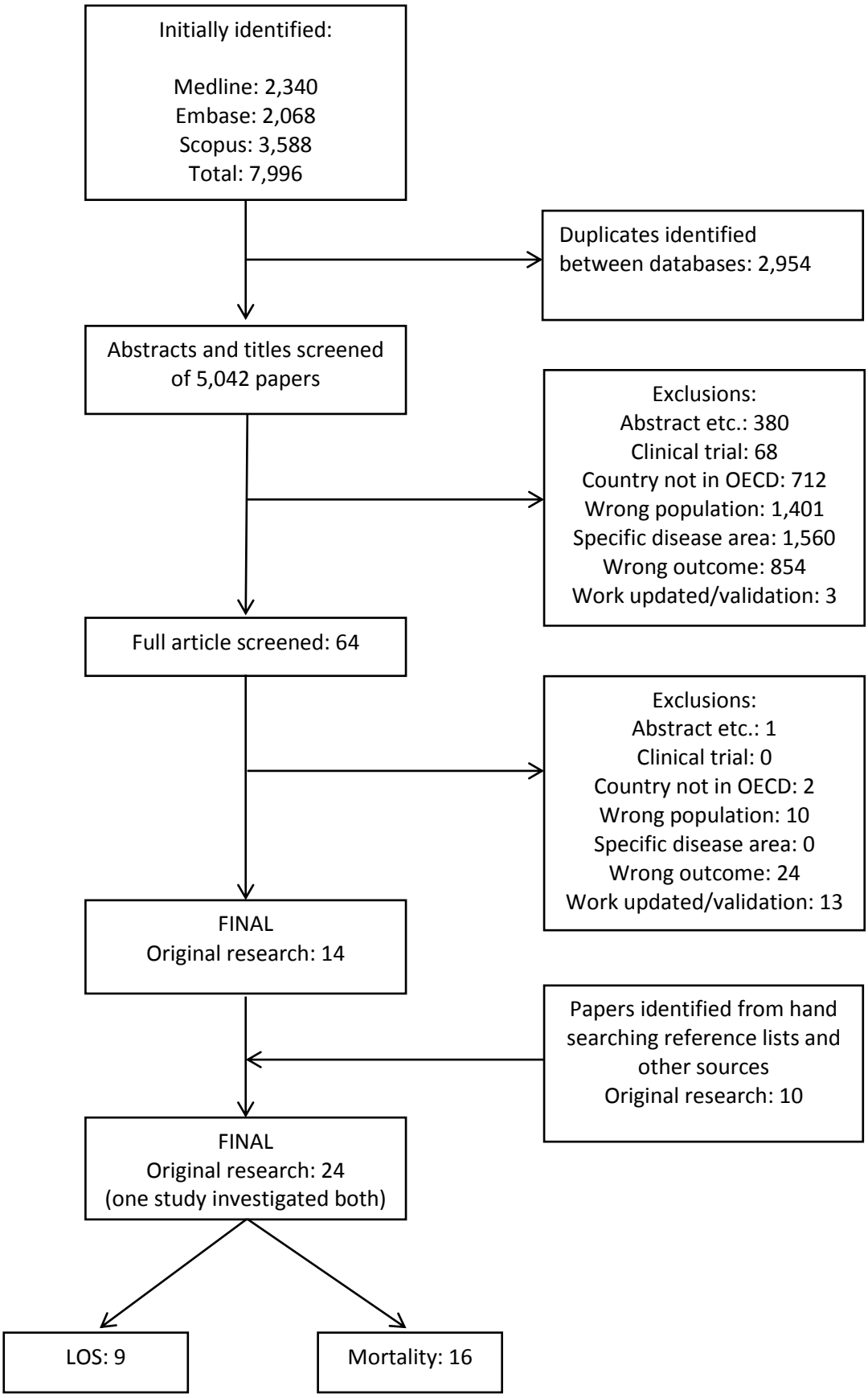
A total of 7,996 studies were identified from Medline, Embase and Scopus. Figure 2-1 details the exclusions in a flow chart as recommended by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (37). A total of 64 articles were fully screened by reading the full article and 24 articles were ultimately included in the review. Of the 24 articles, 16 investigated the prediction of mortality and nine investigated predicting length of stay. One study investigated both mortality and length of stay. Whilst this systematic review investigated the factors that predict mortality and/or length of stay, the discussion here focusses more on length of stay, where the evidence has been more limited.

Table 2-1: Search strategy for the systematic review in each of the three databases: Medline, Embase and Scopus.

Medline		Embase		Scopus	
1	exp Intensive Care Units, Neonatal/ or exp Intensive Care, Neonatal/ or neonatal care.mp	1	"intensive care unit".mp. or exp intensive care/ or exp intensive care unit	1	(TITLE-ABS-KEY("intensive care" OR "special care" OR "NICU" OR "high dependency" OR "standard care" OR "nursery care")
2	("intensive care" or "special care" or "NICU" or "high dependency" or "standard care" or "nursery care").mp	2	"neonatal care".mp. or exp newborn care	2	TITLE-ABS-KEY("gestational age" OR "infant" OR "premature" OR "preterm" OR "baby")
3	exp Gestational Age/ or exp Infant, Newborn/ or exp Premature Birth/ or exp Infant, Premature/ or preterm.mp. or baby.mp. or exp Infant, Premature, Diseases	3	("intensive care" or "special care" or "NICU" or "high dependency" or "standard care" or "nursery care").mp	3	TITLE-ABS-KEY("determinant$" OR "risk factor$" OR "clinical predictor$" OR "predictor$" OR "prognostic" OR "indicator$" OR "prediction" OR "probability")
4	exp Infant, Low Birth Weight/ or very low birthweight.mp. or exp Infant, Very Low Birth Weight	4	exp prematurity/ or "preterm".mp. or exp gestational age	4	TITLE-ABS-KEY("length of stay" OR "mortality" OR "survival"))
5	("determinant$" or "risk factor$" or "clinical predictor$" or "predictor$" or "prognostic" or "indicator$" or "prediction" or "probability").mp	5	"low birthweight".mp. or exp low birth weight	5	1 and 2 and 3 and 4
6	exp Risk Factors	6	"very low birthweight".mp. or exp very low birth weight	6	5 and PUBYEAR > 1993 AND NOT ALL(animal$ OR rat OR rats OR cat OR cats OR bovine OR sheep)

Medline		Embase		Scopus	
7	"length of stay".mp. or exp "Length of Stay"	7	"risk factor".mp. or exp risk factor	7	LIMIT 5 - TO(LANGUAGE, "English")
8	exp Infant Mortality/ or exp Perinatal Mortality/ or mortality.mp. or exp Hospital Mortality/ or exp Mortality/ or "neonatal mortality".mp	8	("determinant$" or "risk factor$" or " clinical predictor$" or "predictor$" or "prognostic" or "indicator$" or "prediction" or "probability").mp	8	-
9	exp Survival/ or survival.mp	9	"length of stay".mp. or exp "length of stay"	9	-
10	1 or 2	10	"infant mortality".mp. or exp infant mortality	10	-
11	3 or 4	11	"perinatal mortality".mp. or exp perinatal mortality	11	-
12	5 or 6	12	"hospital mortality".mp. or exp mortality	12	-
13	7 or 8 or 9	13	"neonatal mortality".mp. or exp newborn mortality	13	-
14	10 and 11 and 12 and 13	14	1 or 2 or 3	14	-
15	limit 14 to (english language and humans)	15	4 or 5 or 6	15	-
16	limit 15 to yr="1994 -Current"	16	7 or 8	16	-
17	-	17	9 or 10 or 11 or 12 or 13	17	-
18	-	18	14 and 15 and 16 and 17	18	-
19	-	19	limit 18 to (human and english language)	19	-
20	-	20	limit 19 to yr="1994 - Current"	20	-

Figure 2-1: Flow chart of search results and the screening process for the systematic review.



2.5 RESULTS OF THE SYSTEMATIC REVIEW FOR MORTALITY

Sixteen studies (14, 30, 38-51) investigated the prediction of mortality, one of which also investigated length of stay (40). A variety of patient populations were included, and this diversity demonstrates the lack of a clear definition of a high-risk population that would potentially benefit most from mortality prediction (Table 2-2). All studies included preterm populations or babies born with a low birthweight and whilst there was heterogeneity in the selection of the study populations, most studies were based in neonatal intensive care units. Although definitions of intensive care will differ between countries, babies treated in intensive care units will generally be the sickest in the population. However, it is impossible to assess if an individual baby within an intensive care unit required intensive care. For example, in the United Kingdom care is defined according to treatment rather than physical location (22). Whilst there were differences in study populations, research was focussed on the group of babies most likely to benefit from a mortality prediction because they were at the highest risk of adverse outcomes.

Within individual studies, exclusions comprised: deaths in the delivery room (30, 38); major/lethal congenital anomalies (14, 41-44, 47, 48, 50, 51); higher order multiple births (three or more) (50); implausible birthweight or gestational age (14, 47, 48); received palliative or comfort care (30, 43, 47); released to a nursery in less than 24 hours (30); admitted to intensive care more than 48 hours after birth (30); indeterminate sex (14, 43); deaths not attributed to care (for example: complex cardiac cases) (46); step down care (40); missing data (mentioned specifically by study (43) but likely to have been used by more studies); hydrops fetalis (accumulation of fluid) (42) and deaths before labour (41). Many of these exclusions represent the babies with unusual or different survival rates compared to their peers. However, some exclusions such as major/lethal congenital anomalies are subjective, as no clear internationally accepted definition exists.

Most studies aimed to predict survival or mortality using data that was available at or around the time of birth (14, 41, 42, 44, 47, 48, 50, 51). Others predicted survival or mortality using data available very soon after birth, typically in less than 24 hours (38,

39, 46, 49). Four studies did not indicate the time frame within which data were available (30, 40, 43, 45). Therefore, the most common approach was to be able to make a prediction of mortality either around the time of birth or within 24 hours. This provides an informative, and accessible estimate early in the neonatal care pathway. Factors used in the analyses of the sixteen studies can be found in Table 2-2.

2.6 SELECTED CLINICAL FACTORS: MORTALITY

2.6.1 INHERENT FACTORS OF THE BABY

All studies accounted for inherent information about the baby, particularly gestational age at delivery (87.5%, 14/16 of studies) and birthweight (75%, 12/16) (Table 2-3). All identified inherent conditions were easily measurable except for congenital anomalies and SNAPPE-II (a mortality risk score: Score for Neonatal Acute Physiology – Perinatal Extension II).

There is no internationally agreed definition of major congenital anomaly (52), with some research defining major anomalies as incompatible with long-term survival (e.g. anencephaly, the absence of major parts of the brain) whereas others class major anomalies as anything that has functional or cosmetic consequences (e.g. cleft lip). No studies in this review defined what conditions were considered congenital anomalies and seven studies (43.8%) used this as part of their exclusion criteria.

SNAPPE-II was considered as both an inherent characteristic and a characteristic related to the condition of the baby as the score comprises a mixture of several components including birthweight (inherent), blood pressure, temperature and presence of seizures (conditions). As SNAPPE-II requires information related to the baby's condition it cannot be measured as quickly or objectively as other inherent factors.

Ethnicity is often considered predictive of mortality (53) and was included by three studies (38, 39, 51), although most of these studies were based in predominantly White populations and therefore it is unsurprising that few accounted for ethnicity. One study acknowledged that maternal ethnicity was highly predictive, but that the

level of missing data they encountered caused it to be excluded from the final analyses (42).

Inherent factors were also used in various studies' inclusion and exclusion criteria, particularly birthweight and gestational age, indicating their importance when defining a population at risk of mortality.

2.6.2 CLINICAL CONDITIONS OF THE BABY

Nine studies (56.3%) (Table 2-3) accounted for clinical conditions and details related to the baby including body temperature and blood pressure. However, the choice of these conditions varied widely, with some occurring early and others potentially a long time after birth.

In total, 17 different clinical conditions were considered, indicating a lack of consensus over which conditions were most predictive of mortality. The only factors measured by more than one study were Apgar score (31.3%, 5/16. Apgar score is a measure of the physical condition of a baby at birth (54)) and temperature of the baby on admission to the neonatal unit (12.5%, 2/16). Several of the conditions that were accounted for occur early in the clinical pathway, for example: condition on admission to the neonatal unit. Other conditions could occur at any point throughout a baby's stay, for example Necrotising Enterocolitis (NEC: a complication where a portion of the bowel experiences tissue death). Therefore a mortality risk score which could be revised in light of later occurring conditions could be helpful.

One study commented that, particularly in the very low birthweight babies, diagnosis of a medical condition was unlikely to improve prediction as they were already a homogeneous group in terms of their prematurity, and that diagnosis of specific conditions was likely to be more useful in the larger more mature babies (30).

2.6.3 ANTENATAL OR MATERNAL FACTORS

Nine studies (56.3%, 9/16) accounted for antenatal treatment or maternal factors. However, again there was a wide range of factors considered by the studies, with many only being considered by one study (Table 2-3).

The most common characteristic accounted for in analyses was the provision of antenatal steroids (50% of studies, 8/16), which are known to significantly improve survival, particularly in the preterm population (55). All factors identified and used could be measured during or before labour, and were therefore known at the birth of the baby, allowing for a quick prediction of the risk of mortality if these were considered important enough for inclusion. However, one study highlighted that it did not account for antenatal events because *"this may obscure the ill effects of improper treatment"* (30).

2.6.4 Organisational factors and treatment of the baby

Less than half of the studies accounted for organisational factors (Table 2-3: 31.3%, 5/16), or factors related to the treatment of the baby (12.5%, 2/16). This indicates that the general consensus was that the majority of a baby's risk of mortality could be explained by factors specifically related to the baby, inherent or otherwise. Care should be taken when accounting for organisational or treatment factors as this may obscure improper treatment.

2.7 Discussion: mortality

2.7.1 Other published research

A recent systematic review conducted to identify the factors which predict neonatal mortality (56) found similar conclusions to this one. However, the restriction in this review to OECD countries has potentially provided a more representative sample of populations to be compared. Birthweight and gestational age predict survival well, and generally other factors only provide a small improvement in model prediction (30). Nevertheless, Medlock et al noted that the differences between the sexes and ethnic groups are of interest to clinical decision makers (56).

2.7.2 Conclusions

There has been interest in the area of predicting neonatal mortality, with 19 studies published over the time period investigated. Studies generally focussed on providing predictions of mortality risk based on the state of the baby at birth, or within 24 hours.

A minimal number of factors are required to make a prediction quick and simple to use. A measure of the early condition of the baby should also potentially be included such as Apgar score (54). Mortality scores have been used to provide a risk adjustment when comparing hospitals and therefore it is important to not include variables that can be affected by differences in the provision of neonatal care. Care should be taken when accounting for treatment of the baby to ensure that improper treatment effects are not obscured (30, 57).

Table 2-2: Summary characteristics of the studies identified for inclusion in the systematic review of mortality prediction.

	Country	**Population investigated**	**Study Setting**	**Model selection**	**Statistical methods**	**Model fit methods**
Ambalavan (2001) (38)	USA	Birthweight <1000g	Tertiary care centre (likely to be intensive care due to birthweight)	Literature and then forward selection used to shorten the list	Logistic model and Artificial Neural Network (ANN)	Area under the curve (AUC)
Ambalavan (2005) (39)	USA	Birthweight 401 to 1000g	Network of intensive care units	Previous literature	Logistic model and ANN	AUC (with t-tests to compare between ANN and Logistic approaches)
Berry (40)	Canada	All births	Neonatal intensive care	Previous literature and clinical knowledge	Logistic model	None
Bolisetty (51)	Australia	23 to 31 weeks gestation	All tertiary and specialist nurseries in New South Wales	Backwards stepwise elimination	Logistic model	None
Cole (41)	UK	22 to 31 weeks gestation	Probably intensive care (all births <32 weeks in UK usually admitted to NICU)	Previous knowledge	Logistic model	AUC Hosmer Lemeshow (HL) test

	Country	**Population investigated**	**Study Setting**	**Model selection**	**Statistical methods**	**Model fit methods**
Evans (42)	Australia and New Zealand	Birthweight<1500g or <32 weeks gestation	Network of 29 intensive care units	Univariate analysis. If significant, entered into multivariable and removed sequentially using Likelihood Ratio tests	Logistic model	AUC HL Test
Ge (43)	Canada	23 to 30 weeks gestational age	Level III intensive care units	Preselected clinically important factors entered into stepwise model	Multinomial model	Briers Score HL Test (extension) Calibration plots
King (50)	Canada	23 to 28 weeks gestational age	Tertiary centre	Unclear	Logistic model	None discussed
Locatelli (44)	Italy	Birthweight <750g	Neonatal intensive care unit	Forward selection from a long list of variables	Logistic model	Adapted R^2
Manktelow (14)	UK	<33 weeks gestational age	Neonatal intensive care unit	Selected in advance from previous knowledge	Logistic model	AUC Coxs Calibration Briers Score Farringtons Test

	Country	**Population investigated**	**Study Setting**	**Model selection**	**Statistical methods**	**Model fit methods**
Moro (45)	Spain	Birthweight <1500g	Neonatal unit (likely to be intensive care)	Not explicit but appears that variables significant in univariate are put into multivariable.	Cox model	Proportional hazards assessed
Parry (46)	UK	<33 weeks gestational age	Neonatal intensive care	List of variables selected by experts and then AIC to select best model	Logistic model	AUC HL Test Coxs Calibration
Richardson (30)	Canada	All births	Neonatal intensive care unit	Variables from previous work, removed if unreliable (feedback from experts) or no association with mortality	Logistic model	AUC Goodness of fit
Shah (47)	Canada	22 to 32 weeks gestational age	Neonatal intensive care unit	Stepwise procedure	Logistic model	AUC HL Test
Tyson (48)	USA	22 to 25 weeks gestational age	Neonatal intensive care unit	Variables selected in advance	Logistic mixed model	Bootstrapped coefficients AUC HL Test

	Country	**Population investigated**	**Study Setting**	**Model selection**	**Statistical methods**	**Model fit methods**
Zernikow (49)	Germany	Birthweight <1500g or <32 weeks gestational age	Neonatal intensive care unit	Forward selection model	Logistic model and ANN	AUC Predicted versus observed

Table 2-3: Factors identified by the statistical analyses within each study when predicting mortality. Factors are presented by individual study and overall.[1]

	Ambalavan (2001)	Ambalavan (2005)	Berry	Bolisetty	Cole*[1]	Evans	Ge	Locatelli	King	Manktelow*[2]	Moro	Parry*[3]	Richardson*[4]	Shah*[5]	Tyson	Zernikow*[6]	Number of studies
Inherent factors																	
Birthweight (including z score, weight for gestational age, small for gestational age)	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		✓	12
Congenital anomalies			✓	✓							✓					✓	4
Ethnicity/nationality	✓	✓		✓													3
Gestational age	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	14
Multiplicity	✓	✓							✓						✓		4
Sex		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓		11
SNAPPE-II*[7]			✓														1
Any inherent factor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	16
Antenatal treatment and maternal factors																	
Antepartum haemorrhage				✓													1
Antenatal steroids	✓	✓		✓			✓	✓	✓		✓				✓		8
Attended prenatal care		✓															1
Breech birth				✓													1

	Ambalavan (2001)	Ambalavan (2005)	Berry	Bolisetty	Cole*[1]	Evans	Ge	Locatelli	King	Manktelow*[2]	Moro	Parry*[3]	Richardson*[4]	Shah*[5]	Tyson	Zernikow*[6]	Number of studies
Emergency delivery																✓	1
Gravida		✓															1
Hypertension/pre-eclampsia		✓				✓											2
Marital status of mother		✓															1
Parity		✓															1
Prenatal antibiotics		✓															1
Prepartum haemorrhage		✓															1
Presence of labour		✓															1
Surfactant							✓				✓						2
Tocolytic agents given		✓															1
Any antenatal or maternal factor	✓	✓		✓		✓	✓	✓			✓				✓	✓	9
Conditions of the baby																	
Apgar score	✓	✓						✓					✓			✓	5
Base excess												✓					1
Capillary pH																✓	1
Condition on admission																✓	1
Intraventricular Haemorrhage (IVH)											✓						1
Lowest temperature													✓				1
Lowest serum pH													✓				1

	Ambalavan (2001)	Ambalavan (2005)	Berry	Bolisetty	Cole*[1]	Evans	Ge	Locatelli	King	Manktelow*[2]	Moro	Parry*[3]	Richardson*[4]	Shah*[5]	Tyson	Zernikow*[6]	Number of studies
Mean blood pressure													✓				1
Necrotising enterocolitis (NEC)											✓						1
PaO_2/FIO_2 (fraction of inspired oxygen)													✓				1
Pneumothorax											✓						1
Respiratory Distress Syndrome	✓																1
Seizures													✓				1
SNAP II*[8]							✓										1
SNAPPE-II			✓														1
Temperature on admission to unit												✓				✓	2
Urine output													✓				1
Any condition of the baby	✓	✓	✓				✓	✓			✓	✓	✓			✓	9
Treatment of the baby																	
Mechanical ventilation							✓										1
Surgery whilst in hospital			✓														1
Any treatment of the baby			✓				✓										2
Organisational factors																	

	Ambalavan (2001)	Ambalavan (2005)	Berry	Bolisetty	Cole*[1]	Evans	Ge	Locatelli	King	Manktelow*[2]	Moro	Parry*[3]	Richardson*[4]	Shah*[5]	Tyson	Zernikow*[6]	Number of studies
Centre (random effect)															✓		1
Mortality rate of centre		✓															1
Outborn/inborn							✓									✓	2
Transferred			✓														1
Any organisational factor		✓	✓				✓								✓	✓	5

[1]: *[1] This is known as the PREM (prematurity risk evaluation score) score. The *"birth model"* is presented which was described as being *"almost as reliable before delivery as after."*

*[2] Also known as the '*Draper grid',* named after one of the co-authors.

*[3] Also known as the Clinical Risk Score for Babies (CRIB) score.

*[4] This is the Score for Neonatal Acute Physiology (Perinatal Extension): SNAP II/SNAPPE II. Presented here is SNAPPE II, which was designed as a measure of mortality risk. SNAP is made of the same components, minus birthweight, small for gestational age and Apgar score and measures illness severity.

*[5] This is the model presented in Table 4 of the paper. The more complicated model offered little improvement in model performance (0.35% improvement in AUC).

*[6] This is the ANN model, which performed better but is unsuitable for individual treatment decisions. However, this is a weakness of all population based estimation. ANN model had 13 items compared to logistic models which had six.

*[7] SNAPPE II comprised: mean blood pressure; lowest temperature; PaO_2/FIO_2 ratio (a measure of lung function); lowest serum pH; multiple seizures; urine output; birthweight; SGA; Apgar score. This was counted as both the condition of the baby and inherent factors as the score combines both.

*[8] SNAP II was comprised of: mean blood pressure; lowest temperature; PaO_2/FIO_2 ratio; lowest serum pH; multiple seizures and urine output. SNAP II is comprised solely from conditions of the baby.

2.8 RESULTS OF THE SYSTEMATIC REVIEW FOR LENGTH OF STAY

Nine studies (5, 6, 19, 29, 40, 58-61) were identified that investigated length of stay including one that also investigated mortality (19). Characteristics of the studies are presented in Table 2-4.

Studies were based in the USA (5, 19, 29, 59); Sweden (58); Canada (40); the UK (6, 60) and Germany (61). A variety of patient populations were included: all gestational ages (19, 40); 23 to 36 weeks gestational age (61); 23 to 32 weeks (60); 24 to 28 weeks gestational age (6); 401 to 1000g birthweight (59); less than 27 weeks gestational age (29); 22 to 29 weeks gestational age or 401g to 1500g birthweight or larger babies meeting a specific criteria (5), and 30 to 34 weeks gestational age (58) (Table 2-4). Although the search was for papers from 1994, little research was identified until the late 2000s indicating the recent increased interest in this area, potentially correlating with improvements in neonatal survival leading to more babies with long lengths of stay in the neonatal unit and thus an increased pressure on neonatal services.

Six studies (66.7%) were based within intensive care (5, 6, 19, 40, 59, 60) and it was unclear where the other studies were based, although in two studies it was likely to be intensive care given the prematurity of the population (29, 61). Therefore, although there were differences with regards to gestational age, these babies were likely to be the sickest of the population and have the longest length of stay, if they survived.

Exclusions within specific studies included: (major) congenital anomalies (5, 19, 29, 58, 59); deaths in hospital (5, 58-60) or before admission to intensive care (19); previously discharged babies subsequently readmitted (5); babies that were admitted for comfort/terminal (palliative) care (19); step down care (40); surgery (5, 58, 59); ambiguous sex (6); implausible birthweight (6); unusual care pathways (60); in hospital longer than one year (29) and transfers to facilities not considered neonatal care (29). Some of these exclusions may be subjective and therefore not consistently measured between studies.

Table 2-4: Summary characteristics of the studies identified for inclusion in the length of stay systematic review.

	Country of study	**Year of publication (data)**	**Number of patients in study**	**Population investigated**	**Physical location of study**	**Model selection**	**Statistical methods**	**Model fit methods**
Altman (58)	Sweden	2009 (2004-2005)	2388	30 to 34 weeks gestational age	Neonatal units of varying levels of care	Univariate analysis then significant ($p<0.2$) entered into stepwise	Linear regression	R^2
Bender (19)	USA	2013 (1999 and 2002)	293 (validated on 615)	All gestations	Neonatal intensive care unit	Prior knowledge	Accelerated failure time parametric models	Cross validation R^2
Berry (40)	Canada	2008 (2002)	604	All gestations	Neonatal intensive care unit	Prior knowledge	Logistic regression	None, but validation in other centres recommended
Hinchliffe (6)	UK	2013 (2006-2010)	2723	24 to 28 weeks gestational age	Neonatal intensive care unit	Prior knowledge	Competing risks	None (acknowledged as weakness)
Hintz (29)	USA	2010 (2002-2005)	2254	<27 weeks gestational age	Unclear but likely to be neonatal intensive care due to	Prior knowledge	Logistic regression	R^2

	Country of study	**Year of publication (data)**	**Number of patients in study**	**Population investigated**	**Physical location of study**	**Model selection**	**Statistical methods**	**Model fit methods**
					gestational age			
Lee (2013) (59)	USA	2013 (2008-2010)	2012	401 to 1000g birthweight	Neonatal intensive care unit	Stepwise selection	Linear mixed model	R^2
Lee (2016) (5)	USA	2016 (2008-2011)	23,551	All babies 401g to 1500g or 22 to 29 weeks gestational age plus larger babies meeting specified criteria	Neonatal Intensive Care Units	Prior knowledge then minimum AIC	Negative binomial model with hospital as random effect	Root Mean-Square Error (RMSE)
Manktelow (60)	UK	2010 (2005-2007)	4702	23 to 32 weeks gestational age	Neonatal unit	Prior knowledge and then change in deviance to decide how to model variables	Quantile regression	Observed versus predicted comparison
Zernikow (61)	Germany	1999 (1989-1996)	2144	23 to 36 weeks gestational age	Unclear but single centre	Forward stepwise	Artificial neural networks Multiple linear regression	Observed versus predicted comparison

Table 2-5: Factors included by the statistical analyses within each study when predicting length of stay. Factors are presented by individual study and overall.[2]

	Altman et al	Bender et al*[a]	Berry et al	Hinchliffe et al	Hintz et al	Lee et al *[b] (2013)	Lee et al (2016)	Manktelow et al	Zernikow et al	Number of studies
Inherent factors										
Birthweight (modelled in multiple ways including categorised, small for gestational age, z score)	✓	✓		✓	✓	✓	✓	✓	✓	8
Congenital anomalies			✓				✓		✓	3
Date/year of birth							✓		✓	2
Ethnicity/race/nationality						✓	✓		✓	3
Gestational age	✓	✓		✓				✓	✓	5
Head circumference									✓	1
Length of baby at birth									✓	1
Multiplicity	✓						✓			2
Sex		✓		✓		✓	✓	✓		5
SNAPPE-II *[1]			✓							2
Any inherent factor	✓	✓	✓	✓	✓	✓	✓	✓	✓	9
Antenatal treatment and maternal factors										
Antenatal steroids						✓	✓			2
Diabetes							✓			1
Emergency delivery									✓	1
Fetal distress						✓	✓			2

	Altman et al	Bender et al*[a]	Berry et al	Hinchliffe et al	Hintz et al	Lee et al *[b] (2013)	Lee et al (2016)	Manktelow et al	Zernikow et al	Number of studies
Hypertension						✓	✓			2
Maternal age	✓						✓			2
Mode of delivery							✓			1
Other maternal/obstetric condition							✓			1
Received prenatal care							✓			1
Any antenatal treatment or maternal factor	✓					✓	✓		✓	4
Conditions of the baby										
Admission reason								✓		1
Apgar score						✓	✓			2
Bronchopulmonary Dysplasia (BPD)					✓					1
Hyperbilirubinaemia	✓									1
Hypoglycaemia	✓									1
Infection	✓									1
Respiratory Distress Syndrome	✓									1
Retinopathy of Prematurity (stage 3 or higher)					✓					1
Sepsis episode or NEC					✓					1
Severe morbidity*[2]	✓									1
SNAP*[3]		✓								1
SNAPPE-II			✓							2
Any condition of the baby	✓	✓	✓		✓	✓	✓	✓		7
Treatment of the baby										

	Altman et al	Bender et al*[a]	Berry et al	Hinchliffe et al	Hintz et al	Lee et al *[b] (2013)	Lee et al (2016)	Manktelow et al	Zernikow et al	Number of studies
Surgery whilst in hospital			✓							1
Surgery for Patent Ductus Arteriosus, Necrotising Enterocolitis, or Retinopathy of Prematurity					✓					1
Umbilical vein catheter									✓	1
Ventilation									✓	1
Any treatment of the baby			✓		✓				✓	3
Organisational factors										
Centre (random effect)					✓	✓		✓		3
Domiciliary care	✓									1
Fixed discharge criteria	✓									1
Level 3 centre	✓									1
Transferred/Outborn status			✓				✓			2
Any organisational factor	✓		✓		✓	✓	✓			5

[2] *[1]: The calculation of the SNAPPE-II score includes: mean blood pressure; lowest temperature; PaO_2/FiO_2 ratio; lowest serum pH; multiple seizures; urine output; birthweight; small for gestational age and Apgar score. These are a combination of inherent and conditions of baby factors and so SNAPPE II appears in both categories.

*[2]: Severe morbidity is defined as: any of: Intraventricular Haemorrhage (IVH) Grade three or four; Retinopathy of Prematurity (ROP) greater than or equal to Stage three; Bronchopulmonary Dysplasia (BPD)

*[3]: This is the original SNAP score, devised in 1993, and comprised of 34 items, largely related to the condition of the baby. Examples of items belonging to the score include: heart rate, blood pressure and platelet count.

*[a] The final model is taken to be the SNAP one as this model was validated.

*[b] This study stratified analyses by birthweight, and different variables were used for each stratification. All variables from all models are listed here.

Table 2-6: Quality assessment of the included studies investigating length of stay. These are assessed using a modified version of the QUIPS tool.

	Domains of quality				
	Study participation	**Study exclusion/attrition**	**Outcome measurement (for example: definition and measurement)**	**Risk adjustment and clinical predictors* (for example: missing data)**	**Statistical analyses and reporting (for example: was validation considered)**
Altman et al	Study is population based (and included 21/34 units in Sweden) but babies were excluded if moved to a hospital not included in study.	Babies discharged to other clinics were excluded.	Continuous post-menstrual age at discharge.	Detailed information about how factors were measured.	None mentioned
Bender et al	Single centre study.	Transfers were included in the analysis and their length of stay in other facilities was included in the total length of stay. Sensitivity analyses excluded them.	Continuous length of stay (days).	Made use of mortality scores with large number of elements included. Potential issues if there was missing data.	Split sample.
Berry et al	Study based in two hospitals. Data extracted from ward	Length of stay days after transfer to another centre were not included.	Length of stay categorised into: <21 days or ≥21 days. No	Made use of mortality scores with large number of elements included.	Acknowledgement that future validation required.

	Domains of quality				
	Study participation	**Study exclusion/attrition**	**Outcome measurement (for example: definition and measurement)**	**Risk adjustment and clinical predictors* (for example: missing data)**	**Statistical analyses and reporting (for example: was validation considered)**
	registers, charts and patient records.		justification for these cut points.	Potential issues if there was missing data.	
Hinchliffe et al	Population based study covering a region of hospitals. Data is extracted from medical records and stored in a routine database used for research purposes.	Minimal loss to follow up when discharged out of region covered by study. Included in analysis as censored observations.	Proportion of deaths and discharges - continuous length of stay (days).	Detailed information about how factors were measured.	Acknowledged that further work required to assess model.
Hintz et al	Population based study within a large network containing multiple hospitals. Data extracted from a routine database set up for research.	Attrition of babies transferred out of the region covered by study.	Early (lowest quartile of age at discharge) or late discharge (high quartile of age at discharge). No justification for these cut points.	Variables clearly defined. Some factors subjective in measurement including Bells staging for NEC.	Split sample
Lee et al (2013)	Population based study of a large number of intensive care units.	Attrition from transfers to lower levels of care	Continuous length of stay in days (log transformed).	Limited details about variables but most could be measured objectively.	Split sample

	Domains of quality				
	Study participation	**Study exclusion/attrition**	**Outcome measurement (for example: definition and measurement)**	**Risk adjustment and clinical predictors* (for example: missing data)**	**Statistical analyses and reporting (for example: was validation considered)**
		(acknowledged as causing bias).			
Lee at al (2016)	Population based study in 90% of intensive care units in large American state	Only babies inborn or transferred to unit in study within one day of life.	Continuous length of stay (days).	Variables clearly defined and objectively measured. Missing data not discussed.	Split sample
Manktelow et al	Population based study covering a region of hospitals. Data is extracted from medical records and stored in a routine database used for research purposes.	Minimal attrition: when discharged out of region covered by study.	Continuous length of stay (days).	Some factors subjectively measured including reason for admission to intensive care.	Acknowledged that future validation needed.
Zernikow et al	Single centre study	Transfers excluded from the study.	Continuous length of stay (days).	Limited information about variables but most objective to measure.	Split sample

2.9 SELECTED CLINICAL FACTORS: LENGTH OF STAY

The nine studies investigating length of stay in neonatal care presented a total of 39 factors which were accounted for in analyses. Details of prognostic factors selected and included in the final analysis by each study are given in Table 2-4.

2.9.1 INHERENT FACTORS OF THE BABY

All nine studies accounted for some form of inherent factor, with the most common being birthweight (88.9%, 8/9), gestational age (55.5%, 5/9) and sex (55.5%, 5/9) (Table 2-5). Birthweight was accounted for by most studies but there were differences in the approaches to measuring it including use of z-scores or birthweight centiles (62). Birthweight z-score is a measurement of how many standard deviations a given birthweight is from the average for certain characteristics, usually gestational age and sex.

All inherent factors were available shortly after birth and were simple and objective to measure. A first day of life prediction for length of stay can be made using these factors, although this estimate may change over time depending on the clinical progress of the baby during their stay.

Other factors, identified at or before birth, such as congenital anomalies were only accounted for by three studies (Table 2-5: 33.3%, 3/9) but often comprised part of the exclusion criteria (55.5%, 5/9). As stated in the previous section, congenital anomalies are hard to account for in any analysis as there is no internationally recognised list of what constitutes a major anomaly. Some congenital anomalies are unlikely to impact on length of stay at all, whereas more severe anomalies or those that can require surgery, for example gastroschisis (baby born with their intestines outside the abdomen), may have a substantial impact on length of stay. Consequently, even when studies exclude or adjust for major anomalies it can never be guaranteed that it is a comparison of *"like with like"*. Thus, while congenital anomalies may have an impact on length of stay, it is likely too broad a term to include in a prediction model.

2.9.2 CLINICAL CONDITIONS OF THE BABY

Whilst information from the first day of life was useful (61), prediction was improved when perinatal factors (29) or severity of illness (19) were considered and it may be appropriate to amend an initial length of stay prediction with the changing clinical condition of the baby. Seven studies attempted to capture this by adjusting for some clinical condition of the baby, but there was little consensus on what this factor should be, with included factors ranging from early occurring conditions (for example: reason for admission to intensive care), to those that occurred potentially much later in the care pathway (for example: Bronchopulmonary Dysplasia, BPD, which is not diagnosed until at least 36 weeks PMA). Therefore, whilst it is potentially important to account for the condition of the baby in some way, there was little agreement over what form this should take. It seems fair that the choice should probably be an early occurring condition, independent of quality of care. However, it is difficult to account for specific clinical conditions when not all babies will survive long enough to potentially experience it.

For the very preterm baby, medical conditions may be experienced prior to discharge (for example: retinopathy of prematurity, ROP), but prematurity may well still be the main factor that determines the overall time in the neonatal unit. Lee et al alluded to this stating that normal birthweight babies and those born closer to term are likely to have varied reasons for their length of stay, making predictions complex (5).

2.9.3 ANTENATAL OR MATERNAL FACTORS

Four studies accounted for antenatal or maternal factors but there was little agreement between studies over what factors should be included. Three factors were accounted for by more than one study: use of antenatal steroids (Table 2-5: 2/9); hypertension in pregnancy (2/9) and maternal age (2/9).

2.9.4 ORGANISATIONAL FACTORS AND TREATMENT OF THE BABY

Organisational factors were considered in five studies (55.5%, 5/9), with most relating to the setting of the care including transfers between units (40). This demonstrates the potential importance of considering the varying levels of care, and how length of stay

can potentially be strongly influenced by organisational factors. However, organisational factors are likely to vary between countries. Similarly, even within a country, the designated level of the unit may not indicate the type of care given to the baby. In spite of this, organisational factors were seen by some authors to be equally or even more important than perinatal risk factors (58).

2.10 OTHER ISSUES IN ESTIMATING LENGTH OF STAY

2.10.1 DEATHS IN NEONATAL CARE

Whilst most papers excluded babies who died in hospital, two papers included deaths in the calculation of length of stay. One paper undertaken by Hinchliffe, Seaton et al accounted for this in the methodology implemented (6) and another acknowledged *"mortality rates may have introduced bias, since non-survival truncates observed length of stay"* (19). A third study that excluded deaths acknowledged that: *"accounting for deaths in length of stay measurement may be particularly complex…"* and that the *"length of stay of deaths could vary widely depending on the clinical trajectory…"* (5). No other study incorporated deaths in neonatal care into the analysis and exclusion of this group overlooks the workload and care required by these babies.

2.10.2 THRESHOLDS FOR DISCHARGE

Thresholds for determining the timing of discharge exist informally within neonatal medicine in the UK. Clinicians regularly state that a baby will be discharged at or around the date they were due to be born. Babies are rarely discharged before they gain the ability to suck and feed at around 34 to 35 weeks PMA. Irrespective of the clinical conditions experienced by the baby, it is thought that many preterm babies, particularly those born at less than 32 weeks of gestation, are likely to have matured and recovered enough to be discharged at or near term, their prematurity having been the overwhelming reason for their long length of stay (18). For some babies, later occurring conditions such as late occurring sepsis may cause an increase in their stay which had not been initially anticipated. However, these clinical conditions may not be identifiable until a long time after birth. Therefore, estimates of length of stay should be adapted in light of these conditions where appropriate using clinical judgement.

Whilst the stay in hospital of preterm babies is largely determined by their prematurity, normal birthweight babies (5) and those born closer to term are likely to have varied reasons for their length of stay, making predictions complex. These babies should be considered separately or an adjustment or stratification made in any prediction model (5).

2.11 STUDY QUALITY: LENGTH OF STAY

The adapted form of the QUIPS tool (36) was used to investigate the quality of the studies. A summary is presented in Table 2-6.

2.11.1 STUDY QUALITY: STUDY PARTICIPATION

All studies included in this review made use of data from routine sources, for example medical records, and none actively recruited participants. As such there were no issues related to recruitment of study participants. Six of the studies (5, 6, 29, 58-60) were population based covering large geographical areas, in one instance 90% of an American state (5), and therefore these covered a range of hospital services. Three studies were based in one (19, 61) or two (40) centres with length of stay in other facilities not being included in two studies (58, 61), potentially leading to underestimation of length of stay as the number of transfers between units in neonatal care can be high.

2.11.2 STUDY QUALITY: EXCLUSION AND ATTRITION

Attrition due to babies being transferred out of the area covered by the hospital or study was a potential issue in all studies except one (19) which included other facilities in their length of stay calculations. However this study (19) was based in a single centre, and although they lost no babies to attrition, the specific details about the population they recruited only included care received whilst within that hospital site.

Studies acknowledged attrition as a potential source of bias (59), although for the population based studies this issue was likely to be minimal as these collected data from most hospitals within a geographical region.

2.11.3 STUDY QUALITY: OUTCOME MEASUREMENT

Seven studies estimated continuous length of stay or PMA at discharge (5, 6, 19, 58-61) whilst two studies categorised length of stay, one by dichotomising into <21 days and ≥21 days (40) and the other by classifying discharge as early or late (defined as the lowest and highest quartile of PMA) (29). The decision of how to model length of stay was often based on the statistical analysis being implemented. It was likely that there were minimal issues with the measuring of length of stay as it is an objective measure, although it is of course possible to incorrectly record a date. The decision to dichotomise a continuous variable into a categorical variable is often considered a poor choice in statistics as it potentially leads to a loss of data and power (63). However, the use of continuous length of stay is difficult to model statistically, and this is also often a limitation of the methods implemented.

2.11.4 STUDY QUALITY: RISK ADJUSTMENT

The variables used for risk adjustment are identified here as being the factors which are considered predictive of length of stay and are detailed in Table 2-5. Several studies chose objective measures, which were easily measured and defined, to include as important factors when predicting length of stay (5, 6, 58, 59, 61). Two studies made use of mortality scores including MAIN (morbidity assessment index for newborns) and SNAPPE II (19, 40). These scores contained a large number of factors, leading to potential issues if one or more data items were missing, although this was not discussed for either study. Some factors were subjectively measured, for example reason for admission to the neonatal unit (60) and Bells staging (the severity scoring) for NEC (29).

2.11.5 STUDY QUALITY: STATISTICAL ANALYSES AND REPORTING

Five studies had validated their results by splitting the sample during the initial analysis and reserving data for validation purposes (5, 19, 29, 59, 61). Two studies acknowledged that further validation was needed before results could be generalised (40, 60) and one acknowledged that further work was needed to assess their modelling techniques (6). One study, as part of their analyses, had conducted a pre-planned

external validation but subsequently concluded that the non-validated model was statistically superior (19). However, here the variables used in the validated model are presented in Table 2-5. Only one study made no mention of the validation of the results (58). Therefore, a strength of these studies was that most addressed the issue of validation in some way, either by implementation or discussion.

Whilst no formal scoring of study quality has been undertaken here, all studies had a level of quality which appeared good given the constraints of the study design.

2.12 STRENGTHS AND LIMITATIONS OF THIS REVIEW: LENGTH OF STAY

There is a paucity of evidence investigating the prediction of neonatal length of stay and this review has investigated the limited evidence for the first time (18). Clinical judgment is an important and potentially informative factor for predicting length of stay, although this was not possible to investigate here. However, well-developed prediction models, such as those identified here, are useful because they can provide estimates that can be used to inform the knowledge which clinicians have from their own diagnostic abilities and are more accurate than clinical judgement and assessment alone. Clinicians would like to be able to predict outcomes perfectly, but using statistical models to aid clinical judgement introduces error and uncertainty to better represent a real-world scenario. Prediction models should be developed with the help of clinicians who are able to help distinguish between relevant and irrelevant factors (64).

2.12.1 SYSTEMATIC NATURE OF THIS REVIEW

To ensure that the search was conducted using a truly systematic approach, a random 10% of articles, which included the identified articles, were extracted and provided to a second reviewer. This second review was undertaken independently and eight of the nine articles were identified by both reviewers independently. The ninth article was discussed between both reviewers who subsequently agreed it should be included.

2.12.2 CHOICE OF POPULATION

A weakness of this review was that a variety of settings and gestational groups were considered in the studies in this review. It is likely that different gestational ages will require different prediction models, incorporating very different factors. Creating a prediction model for babies of all gestations, such as that proposed and attempted by Bender (19) or Berry (40), is unlikely to perform well. Babies born near term may have very different reasons for being in the neonatal unit to those born preterm, who have mainly been admitted due to their prematurity and underdevelopment. This was discussed by Lee et al (5) who stratified their analyses by different birthweight groups to group babies of a similar clinical condition together and acknowledged that babies born at a normal birthweight, or near term, may need further stratification by the reason for their admission, for example: sepsis or respiratory disease (5). This approach has potential, and future length of stay predictions should focus on groups of babies that are clinically similar, for example those born very preterm or with a very low birthweight, or analyses should be stratified by clinical condition or disease.

This review focussed on countries that were members of the OECD from 1994, as these countries were likely to be similar in terms of their healthcare services, populations and survival rates of preterm babies. Research has been conducted into the prediction of length of stay in other contexts and countries but not included in this review, for example work by Pepler et al in South Africa (65) and Shah et al in Eritrea, Africa (66). Whilst these studies were in different contexts similar variables were accounted for in their analyses. In the study by Pepler et al, the prediction of length of stay was adjusted for birthweight, Apgar score and maternal ethnicity (65). Similarly, Shah et al accounted for gestational age, birthweight (including small for gestational age), Apgar score, sex, caesarean section birth, maternal age and the baby acquiring pneumonia or hypothermia (66).

2.12.3 INDIVIDUAL PATIENT META-ANALYSIS

A meta-analysis of the data presented in this review was not undertaken, due to the varying analyses and adjustments made in each study. Theoretically, an individual patient data meta-analysis could have been undertaken in order to overcome these

issues, however this is known to be problematic, particularly acquiring the necessary data (67), so this was not conducted. Similarly, it was not possible to investigate publication bias due to the varying analyses and this could be an issue.

2.13 DISCUSSION: LENGTH OF STAY

This chapter provides a systematic review investigating the factors that predict mortality and length of stay in neonatal care. Factors were identified if they were included by study authors in a multivariable analysis which aimed to predict mortality or length of stay. A total of 24 studies were identified, of which 16 investigated the prediction of mortality and nine articles investigated the estimation of length of stay (one investigated both).

The prediction of mortality is well established and has received attention over the years including a systematic review of the factors that predict mortality published in 2011 (56). Although differences existed between study populations, most research focussed on babies born preterm, or with a low birthweight, indicating the attempt to identify a group most likely to benefit from risk prediction. Models generally comprised factors known about immediately at birth or within 24 hours. Twelve studies accounted for birthweight and 14 accounted for gestational age, agreeing other published research which has identified these as the most important factors for predicting mortality and other neonatal outcomes (56). Other factors are thought to offer minimal improvement in prediction models, although sex of the baby and whether antenatal steroids were given are likely to also be clinically important.

The prediction of length of stay had received little attention in the literature, with the earliest study having been published in 1999 (61), and the other eight from 2008 onwards despite this search extending to 1994. This increase in research probably correlated with improvements in neonatal survival, increasing the importance of the prediction of length of stay in the neonatal unit. Eight studies included birthweight, five studies included gestational age and five studies included the sex the baby. It was apparent that inherent factors (those known about at the time of birth) were most useful for predicting length of stay and allowed for an early prediction to be made. There was evidence that use of an early measure of the sickness of the baby might also

be useful, although the ability to revise this estimate over time using clinical knowledge, as a baby developed other medical complications, would probably be more informative.

From this review gestational age, birthweight and sex were identified as being important factors to consider when modelling the prediction of length of stay and mortality and these will be considered throughout the thesis. Gestational age was the factor most often accounted for in the prediction of mortality and the second most popular when predicting length of stay. Gestational age can be objectively measured and has a consistent meaning, whereas birthweight has different meanings at different gestational ages. Additionally, gestational age is considered more fundamental to the physiology of the baby and more valid for populations (30). Therefore, this thesis will consider gestational age initially in analyses.

This review identified the lack of research which has been undertaken in predicting length of stay, indicating the need for this research project. A limitation identified in one study was that their approach of including deaths in the estimation of length of stay caused a truncation of observed length of stay (19). This thesis aims to address this issue.

The ability to predict length of stay would be useful for clinicians and service providers, as well as aiding clinical discussions with parents. Inherent factors appear to be the most important to account for, particularly gestational age, birthweight and sex. This information from the first day of life is informative for predicting length of stay in a simple model and estimates from these models could provide a useful early indicator of likely length of stay in neonatal care. Inclusion of additional factors may be useful, although there is a lack of consensus on what this should be. An early predictor of morbidity may be useful, for example, Apgar score or the reason for admission to neonatal care. However, in the data source for this thesis (Chapter 3) Apgar score was poorly recorded with 20% missing data at one and five minutes. Similarly, for 85% of babies in this thesis the reason for their admission to neonatal care was “need for intensive care” which was not sufficiently detailed for use in this thesis.

Whilst inclusion of another appropriate factor may have improved prediction, models with a small number of factors offered easier clinical use in practice (29) and this review indicates they should include a small number of inherent factors which are easy to measure. The model should potentially also include an early occurring factor as a proxy of the baby's morbidity or clinical condition, although there is no clear agreement over what this factor should be. Alternatively, future models could be dynamic and account for later occurring factors. An appropriate compromise may be to use length of stay estimates alongside clinical judgement to revise them accordingly.

2.14 CHAPTER CONCLUSION

In this chapter, a systematic review was performed to investigate the variables that are considered important when predicting mortality and/or length of stay in neonatal care. It was anticipated that the prediction of these outcomes would share many common factors. Indeed, for both prediction of mortality and length of stay, inherent factors, which were simple to measure on the first day of life were identified as being the most important, with a high proportion of studies accounting for gestational age, birthweight and sex. There was a lack of consensus on what other variables would be useful for adjustments, although it seems that it would help to include a measurement of morbidity.

Gestational age was used in 19 studies and birthweight in 20 studies in this review. The use of gestational age was identified as being more important for physiology (30) and therefore will be used for initial adjustments in this thesis. Additional work in this thesis will also incorporate birthweight and sex.

3 THE NATIONAL NEONATAL RESEARCH DATABASE

3.1 CHAPTER OVERVIEW

Data for this thesis were obtained from the National Neonatal Research Database (NNRD). This chapter provides a brief overview of the process of obtaining permission to access the data including ethics, R&D approval and gaining agreement from neonatal units. A brief introduction to the NNRD and summary statistics will be presented.

3.2 DATA SOURCE

Neonatal units in Great Britain (England, Wales and Scotland) routinely provide data regarding day-to-day care of babies admitted to their unit via a system known as 'Badger.net' or 'Neonatal.net'. The collective group of neonatal units which contribute data is known as the Neonatal Collaborative. The data collected include information about maternal and baby demographics, clinical interventions, diagnoses and the outcomes. Badger.net has existed since 2004 and is used to plan services and record neonatal care activity for use in the calculation of payments by NHS England (68). It is also used by clinicians to facilitate aspects of clinical care including the writing of discharge letters. Badger.net is managed by Clevermed Ltd.

Key variables are extracted from the Badger.net system and these data are transferred to the Neonatal Data Analysis Unit (NDAU, collaborators in this work) who perform data quality checks, clean and anonymise the data, and produce the National Neonatal Research Database (NNRD). The NNRD is managed and maintained at Imperial College London and Chelsea and Westminster NHS Trust. The NNRD is available for a fee for local, regional and national research projects and service evaluations. As of 2015, all neonatal units in England, Wales and Scotland, provide data to the NNRD. No neonatal units in Northern Ireland provide data. This thesis focuses on data from England as data from the other countries were not available for all years of this analysis (2011 to 2014) (69).

The NNRD has ethical approval for the establishment of the database and the collection of data (Research Ethics Committee, REC Reference: 10/H0803/151) and submits an annual review to the National Information Governance Board (NIGB), now the Confidentiality Advisory Group (CAG), under the Health Service (Control of Patient Information) Regulations 2002 (Reference: ECC 8-05(f)/2010). The Caldicott Guardians and Lead Clinicians (dedicated member of staff involved in the data collection) of all Trusts have approved the use of the NNRD data for health service evaluation, but new research projects, such as this PhD, require separate ethical approval and agreement from all participating neonatal units.

Key variables were identified from the NNRD and a data extraction was completed for this study by NDAU. Subsequent updates to the data were provided as and when data downloads became available.

3.3 ETHICS AND RESEARCH & DEVELOPMENT (R&D) APPROVALS

Approvals were required to use the data held within the NNRD for the secondary analysis of this research project. As this was a secondary analysis of previously collected data, this project was deemed to represent minimal risk to participants and the National Research Ethics Service (NRES) recommended a proportionate review was required. Details regarding the study including a protocol, agreement form and study information leaflets were produced. These were submitted to NRES and reviewed within 14 days (standard for proportionate review). Approval was granted on 22 May 2014 from Lancaster, North West Committee (REC Reference: 14/NW/0349). A copy of the ethical approval letter is in Appendix 3.

For this secondary analysis, R&D approval was only required from the NHS Trust that physically held these data: Chelsea and Westminster NHS Foundation Trust. R&D approval was granted on 17 June 2014 (R&D Reference: C&W14/043) and a copy of the R&D approval letter is in Appendix 3. The University of Leicester provided professional indemnity insurance for this study (Appendix 3).

3.3.1 WRITING TO ALL NEONATAL UNITS

All neonatal units that contribute data to the NNRD have granted permission for their data to be used for service evaluation. However, as this study was research, explicit agreement to use each unit's data was required from all units contributing to the NNRD. As of July 1st 2016 the NNRD began using an 'opt-out' approach, where units are informed about studies and given the opportunity to not participate. However, when this study started a full 'opt-in' approach was required.

All 162 neonatal units in England as of June 2014 were written to in order to seek agreement to use their data. A study information leaflet, approved by the REC, was provided alongside a letter and form for completion (Appendix 3) to each neonatal unit. It was required that the form be signed by the Lead Clinician (the named contact for the NNRD) at each neonatal unit, or equivalent, indicating their agreement to take part in this study. Initially all units were contacted on 30 June 2014 with a request for response by 1 August 2014. This was followed up with a second email with an extended deadline, before attempts to personally contact units began in September 2014. The personal direct contact improved response rates, although a variety of issues continued to cause slow response rates, which are outlined with their resolution in Chapter 3.3.2. Ultimately, by December 2014, 100% of units in England had provided agreement for their unit's data to be used in this research. A list of all participating neonatal units and their corresponding Lead Clinician can be found in Appendix 3.

3.3.2 ISSUES AND RESOLUTIONS EXPERIENCED WHILST GAINING STUDY AGREEMENT

The process of obtaining agreement from the 162 neonatal units took approximately six months. Whilst this did ultimately result in agreement from 100% of units, there were issues throughout this six month period which required resolution. A brief outline of the main issues and resolutions is provided in this section.

Whether local R&D required (communication with R&D departments)

Issue: The favourable opinion letter from the REC stated that *"management permission ("R&D approval") should be sought from all NHS organisations involved in*

the study in accordance with NHS research governance arrangements…" This permission was granted when R&D approval was granted from Chelsea and Westminster and no further approvals were required. However, many Trusts did not understand that it was not necessary to obtain local R&D approval for their organisation.

Resolution: This problem required direct communication with clinicians and Trust R&D offices. An email template was drafted, using details from the Standard Operating Procedures for Research Ethics Committees, to send when the query was raised. Subsequent emails were amended to explain this issue briefly, and all units were provided with copies of the ethical approval and R&D approval letters.

Trust management governance or internal review required

Issue: Some Trusts required an internal management review to be undertaken.

Resolution: These were generally straightforward and all information was provided to R&D departments to facilitate this. Trusts that required this were: Northampton General Hospital NHS Trust; Royal Berkshire NHS Foundation Trust; Surrey and Sussex Healthcare NHS Trust; Sandwell and West Birmingham Hospitals NHS Trust; The Shrewsbury and Telford Hospital NHS Trust; Northern Lincolnshire and Goole NHS Foundation Trust and Southend University Hospital NHS Foundation Trust.

Confusion over the favourable opinion letter from the REC

Issue: Even after providing clarification, the (standard) wording on the REC letter still created confusion, with some Trusts remaining unconvinced this was the final ethical approval.

Resolution: North-West Lancaster REC were approached to seek advice over the wording of the favourable opinion letter. Whilst they were unable to change the letter, which was standard, they did provide an email which stated that *"the site concerned [for R&D] is the one where the study is being conducted, not where the data was obtained"* (personal communication with Lancaster REC) and another which stated clearly: *"this is the final opinion. You need to ensure the conditions mentioned are met*

[i.e. that R&D approval should be obtained from Chelsea and Westminster]… but you don't need to inform us of this."

3.4 DATA INCLUSION AND EXCLUSION

Data were obtained from the NNRD and included in this thesis on singleton babies born at 24 to 31 weeks gestational age who were admitted to neonatal units in England on the first day of life and discharged between 2011 and 2014.

The selection of the population by year of discharge means that babies who die towards the end of 2014 are more likely to be included than babies that survive as their care will have continued into 2015. However, this issue is counter-balanced by the babies included from the start of 2011 who are more likely to be those remaining in hospital from care which commenced in late 2010. The selection of this population may have created issues if temporal trends were to be investigated, but this was not of interest within this thesis.

A total of 21,631 babies met the inclusion criteria. Key predictors of length of stay were identified from the systematic review in Chapter 2 as: gestational age, sex and birthweight and initial analyses were undertaken considering gestational age alone. In analyses which only make use of gestational age, babies are retained regardless of their value of birthweight or sex.

3.4.1 EXCLUSIONS DUE TO IMPLAUSIBLE LENGTH OF STAY

The NNRD collects daily data related to each baby and therefore total length of stay was calculated by summing each individual day of care for each baby. Duplicated days of care recorded due to transfers were removed. Those babies who were discharged home before 34 weeks PMA were removed (n=205) as this was felt to be clinically unlikely to have occurred with 34 weeks considered the earliest point at which babies learn to suck and feed (70). Similarly, an upper limit for total length of stay of six months was applied, leading to 199 babies being excluded. Although some babies do stay in neonatal care beyond six months, the numbers are small. A total of 21,227 babies remained in the analysis at this stage.

3.4.2 EXCLUSIONS DUE TO UNUSUAL PATTERNS OF CARE

It is difficult to define a typical care pathway within neonatal medicine. Babies are stepped up and down levels of care according to their clinical needs. Throughout this thesis, care is amalgamated and assumed to occur in a step down fashion (i.e. intensive care followed by high dependency followed by special care) irrespective of the actual order in which the care was received. Although this assumption is acceptable when considering the care received as a whole, it should not be used as a day-by-day prediction of what level of care is needed for an individual baby.

Babies with unusual patterns of care were excluded. This is defined as babies discharged from neonatal care after receiving intensive care but no other lower levels of care (n=57) and babies discharged from neonatal care immediately after receiving high dependency care (n=132) having never received any lower level of care. Although these care pathways are possible, they are likely to represent babies being discharged for palliative care which represents a different care pathway to that under investigation. After exclusion the total number of babies included in analyses was 21,038. Summary information of the included babies is provided in Table 3-1.

3.4.3 UNOBSERVED AND MISSING DATA

Daily data were available for babies throughout their time in neonatal care, indicating the level of care they received on each day of life. Occasionally a baby would be transferred from neonatal care for other specialist care which did not provide data to the NNRD, for example a specialist surgical centre, and then subsequently be transferred back into a neonatal care unit covered by the NNRD. It is not possible to know what happened on the days not reported to the NNRD, but it is likely that these babies were undergoing and recovering from surgery or other specialist treatment, and therefore intensive care days were imputed for the missing days of care. Babies who were discharged to specialist care or a hospital outside of the geographical area covered by the NNRD who subsequently did not return to an NNRD hospital were considered to have been discharged alive from neonatal care, as long as they were not excluded due to a very short length of stay, and no days were imputed for the care after transfer. This assumption forms a sensitivity analysis throughout this thesis.

On a small number of days the level of care was not recorded. For these days special care was imputed as the assumed level of care (n=151 days of care, <0.01% of the total days of care). This assumption was not tested further as the small number of days on which this occurred means this will not have impacted on results but the imputed data prevents potential needless exclusion due to an incorrect length of stay.

3.5 SUMMARY STATISTICS

3.5.1 SUMMARY STATISTICS FOR BABIES INCLUDED IN THIS THESIS

Table 3-1 provides summary statistics for all babies (n=21,038) included in this thesis. Broadly similar numbers of singleton babies were discharged from neonatal care across the four year period. The population had similar case-mix with regards to gestational age and birthweight over the four years. Around 40% of babies were born at 30 and 31 weeks gestational age and more males were admitted for neonatal care than females.

Details about some of the interventions and treatments received by the very preterm babies in this study are given in Table 3-2 and the percentage indicates the percentage of days on which that care was given. For example, in 2011 8.5% of all care days given to preterm babies included some ventilation support. Non-invasive respiratory support, ventilation support, total parenteral nutrition (TPN) and phototherapy are the treatments given most frequently to very preterm babies.

The level of care received by an individual baby is determined by the treatment they receive. When a baby receives several treatments in a single day of varying levels of care, the level is the highest provided that day. However, an individual baby may be receiving more than one type of intervention and therefore the total days of treatments or interventions each year could equal more than the total number of days of care provided in a year.

Table 3-1: Summary statistics of singleton babies born at 24 to 31 weeks gestational age included in this thesis.

	Year of discharge/death from neonatal care			
	2011	**2012**	**2013**	**2014**
Total babies admitted, n	5,368	5,343	5,228	5,099
Gestational age (weeks), n (%)				
24	284 (5.3)	287 (5.4)	276 (5.3)	268 (5.3)
25	327 (6.1)	336 (6.3)	316 (6.0)	325 (6.4)
26	464 (8.6)	465 (8.7)	417 (8.0)	437 (8.6)
27	537 (10.0)	579 (10.8)	480 (9.2)	468 (9.2)
28	690 (12.9)	707 (13.2)	702 (13.4)	685 (13.4)
29	758 (14.1)	791 (14.9)	807 (15.4)	748 (14.7)
30	983 (18.3)	937 (17.5)	976 (18.7)	944 (18.5)
31	1,325 (24.7)	1,241 (23.2)	1,254 (24.0)	1,224 (24.0)
Sex of baby, n (%)				
Male	2,953 (55.0)	2951 (55.2)	2937 (56.2)	2756 (54.1)
Female	2,411 (44.9)	2389 (44.7)	2287 (43.7)	2334 (45.8)
Indeterminate	4 (0.01)	3 (0.01)	4 (0.01)	9 (0.01)
Birthweight (g) Mean (SD)	1224.3 (375.4)	1213.8 (369.3)	1232.1 (372.2)	1215.9 (367.8)
Died in neonatal care, n (%)	492 (9.2)	487 (9.1)	416 (8.0)	367 (7.2)
Total days of care, n	305,150	306,267	295,828	298,177
Days of intensive care	60,995	63,040	60,348	62,058
Days of high dependency care	77,707	83,726	81,346	86,789
Days of special care	166,448	159,501	154,134	149,330

Table 3-2: The number of days of care (%) provided for selected treatments for the babies included in this thesis.

	Year of discharge/death from neonatal care			
	2011	**2012**	**2013**	**2014**
Intensive care, n (%)				
Ventilation	25,836 (8.5)	26,749 (8.7)	23,828 (8.1)	23,123 (7.8)
Chest drain present	1,074 (0.4)	931 (0.3)	845 (0.3)	891 (0.3)
Full exchange transfusion	87 (0.03)	141 (0.1)	115 (0.04)	130 (0.04)
Day of major surgery	693 (0.2)	505 (0.2)	672 (0.2)	606 (0.2)
Presence of replogle tube	221 (0.1)	230 (0.1)	283 (0.1)	176 (0.1)
High dependency, n (%)				
Tracheostomy	352 (0.1)	130 (0.04)	135 (0.1)	267 (0.1)
Non-invasive respiratory support	70,104 (23.0)	80,426 (26.3)	79,828 (26.9)	87,930 (29.5)
Total parenteral nutrition (TPN)	57,044 (18.7)	58,289 (19.0)	56,513 (19.1)	57,892 (19.4)
Special care, n (%)				
Stoma in situ	5,233 (1.7)	4,668 (1.5)	4,687 (1.6)	5,033 (1.7)
Phototherapy	12,855 (4.2)	14,414 (4.7)	14,856 (5.0)	13,805 (4.6)

Common medical conditions experienced by the very preterm babies admitted for neonatal care are shown in Table 3-3. Sepsis and jaundice are highly prevalent amongst preterm populations, with more than half of these babies experiencing them at some stage during their care. An individual baby can experience a number of clinical diagnoses whilst in hospital.

The final discharge destinations of the 21,038 babies in this thesis is described in Table 3-4. In this analysis, all babies who did not die within neonatal care are considered to have been discharged alive from neonatal care including those discharged to other locations for continuing care. Sensitivity analyses will be used to examine those discharged to other services in more detail (Chapters 5.7.5 and 6.6.4).

Table 3-3: The number (%) of diagnoses recorded within the NNRD for the babies included in this thesis.

Clinical condition or diagnosis	**Number of cases (%)**
Sepsis suspect	14,889 (70.8)
Prematurity (28 to 31 weeks)	12,733 (60.5)
Jaundice	11,159 (53.0)
Respiratory Distress Syndrome	10,050 (47.8)
Respiratory Distress (newborn)	13,710 (65.2)
Patent Dutus Arteriosus	4,834 (23.0)
Prematurity (24 to 27 weeks)	4,091 (19.4)
Early risk of infection	3,865 (18.4)
Gastro-oesophageal reflux	3,143 (14.9)
Anaemia of prematurity/Anaemia	4,618 (22.0)
Intrauterine Growth Restriction (IUGR)	2,582 (12.3)
Apnoea	2,565 (12.2)
Necrotising Enterocolitis (NEC)	2,497 (11.9)

Table 3-4: The number (%) of the final discharge destinations for the babies included in this thesis.

Discharge destination	**Number of babies (%)**
Died	1,762 (8.4)
Home/Foster care	18,542 (88.1)
Ward	267 (1.3)
Another hospital	219 (1.0)
Another specialist hospital	74 (0.4)
Surgery	141 (0.7)
Cardiac care	24 (0.1)
Unknown location	9 (0.04)

3.5.2 SUMMARY STATISTICS FOR BABIES WITH UNUSUAL CARE PATHWAYS

Summary statistics are provided for babies excluded from this thesis (n=189) in Table 3-5. The demographics of this group changed across the years of the study demonstrating the heterogeneity of this group. However, the number of babies within this group was small.

Table 3-5: Summary statistics of babies born at 24 to 31 weeks gestational age excluded from this thesis.

	Year of discharge/death from neonatal care			
	2011	**2012**	**2013**	**2014**
Total babies admitted, n	50	45	52	42
Gestational age (weeks), n (%)				
24	4 (8.0)	5 (11.1)	9 (17.3)	13 (31.0)
25	10 (20.0)	6 (13.3)	8 (15.4)	5 (11.9)
26	8 (16.0)	7 (15.6)	3 (5.8)	7 (16.7)
27	6 (12.0)	6 (13.3)	4 (7.7)	3 (731)
28	5 (10.0)	6 (13.3)	7 (13.5)	1 (2.4)
29	5 (10.0)	10 (20.2)	5 (9.6)	5 (11.9)
30	4 (8.0)	1 (2.2)	5 (9.6)	2 (4.8)
31	8 (16.0)	4 (8.9)	11 (21.2)	6 (14.3)
Sex of baby, n (%)				
Male	38 (76.0)	22 (48.9)	38 (73.1)	29 (69.1)
Female	12 (24.0)	23 (51.1)	14 (26.9)	13 (30.9)
Indeterminate	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Birthweight (g) Mean (SD)	1088.8 (425.8)	944.5 (295.1)	1061.3 (462.9)	1062.1 (673.4)
Died in neonatal care, n (%)	0	0	0	0
Days of care, n	2,746	2,730	4,053	2,574
Days of intensive care	1,626	1,697	2,261	1,525
Days of high dependency care	1,120	1,033	1,792	1,049
Days of special care	0	0	0	0

The final discharge destinations of the babies with unusual care pathways is provided in Table 3-6. For those babies discharged to other care facilities the final outcome is not known. A small number of babies are recorded as being discharged home. These may be incorrectly recorded data, or they may have been sent home for palliative care. A report by the charity *Together for Short Lives* suggests that a small number of palliative neonatal cases die in children's hospices or at home (71). However, there are

limited guidelines produced by Neonatal Networks (72) and most refer to life-limiting conditions, for example Edwards Syndrome (73), that are likely to require management beyond the neonatal period.

Table 3-6: The number (%) of the final discharge destinations for the babies excluded from this thesis.

Discharge destination	**Number of babies (%)**
Home/Foster care	28 (14.8)
Ward	23 (12.1)
Another hospital	50 (26.5)
Another specialist hospital	26 (13.8)
Surgery	55 (29.1)
Cardiac care	7 (3.7)

The babies with unusual care pathways were excluded from the work in this thesis as it is impossible to determine if their unusual care was due to missing data due to discharge outside the NNRD area or potential palliative care cases. The number of these cases was too small to warrant a sub-analysis and so they were not investigated further.

3.5.3 Operational delivery networks

Neonatal units are grouped together to form regional Operational Delivery Networks (ODNs) and work together to provide the full range of neonatal services. The ODNs used in this thesis are those that existed in 2013. The North West was the largest ODN in terms of the number of admissions, with 13.9% of the total in England. As of 2016, this ODN had divided into three separate ODNs: Cheshire and Merseyside, Greater Manchester and Lancaster and South Cumbria. The ODN of first admission for each baby is described in Table 3-7.

Table 3-7: The number (%) of babies admitted to neonatal units within each ODN: England 2011 to 2014.

Operational Delivery Network	**Number of admissions (%)**
East of England	1,647 (7.8)
Midlands South West	1,200 (5.7)
North Central and North East London	1,968 (9.4)
North West London	1,060 (5.0)
North West	2,928 (13.9)
Northern	1,083 (5.2)
Peninsula and Western	1,400 (6.7)
South East Coast	1,482 (7.0)
South London	1,392 (6.6)
Staffordshire, Shropshire and Black County	972 (4.6)
Thames Valley and Wessex	1,744 (8.3)
Trent	1,764 (8.4)
Yorkshire and Humber	2,252 (10.7)
Home or in Transit[3]	36 (0.2)
Out of area (including non-NHS) and unknown location[3]	110 (0.5)
Total	**21,038**

3.6 LENGTH OF STAY IN NEONATAL CARE

3.6.1 OBSERVED LENGTH OF STAY OF SURVIVORS AND DEATHS

The total length of stay was calculated for each baby. A baby was analysed as having died if a death was recorded in neonatal care and otherwise they were considered to have been discharged alive from neonatal care, although this may not have been a discharge home. A small proportion of discharges were to specialist services including surgical centres (n=141, 0.7%); cardiac centres (n=24, 0.1%) and other hospitals (n=219, 1%) and there was no subsequent re-admission to a neonatal unit providing data to the NNRD (Table 3-4). These babies were not excluded as they had data recorded until they reached at least 34 weeks PMA and received at least one day of special care. However, the total length of stay, and potentially the proportion of

[3] These babies are included in this thesis as they were admitted to a neonatal unit within the NNRD area of this study (i.e. one of the 162 units in England) on the first day of life, although they were born elsewhere.

deaths, is underestimated for these babies. Sensitivity analyses are used to investigate the robustness of this assumption throughout this thesis. In this study 8.4% of babies died during their neonatal stay which is similar to estimates from other UK studies (14) for this population (Table 1-2).

Figure 3-1 provides the observed length of stay for those babies who died or survived to discharge from neonatal care. The length of stay for the babies who died was shorter, with most deaths occurring in the initial days of life, and nearly all having occurred by 50 days after birth. Conversely, babies who survive were rarely discharged from neonatal care in the early days of life. The babies discharged earliest are those born at later gestational ages, at 30 and 31 weeks gestational age.

The observed median length of stay is provided in Table 3-8 for all babies and by the outcome of the baby. For the earlier gestational ages, particularly less than 28 weeks gestation, where the risk of mortality is highest, the median length of stay and the median length of stay for those who survive are different. For those babies born after 28 weeks gestational age, where the risk of mortality is lower, the median length of stay for those who survive is similar to the overall median length of stay.

Figure 3-1: Observed length of stay for babies who survive to discharge and those who die in neonatal care.

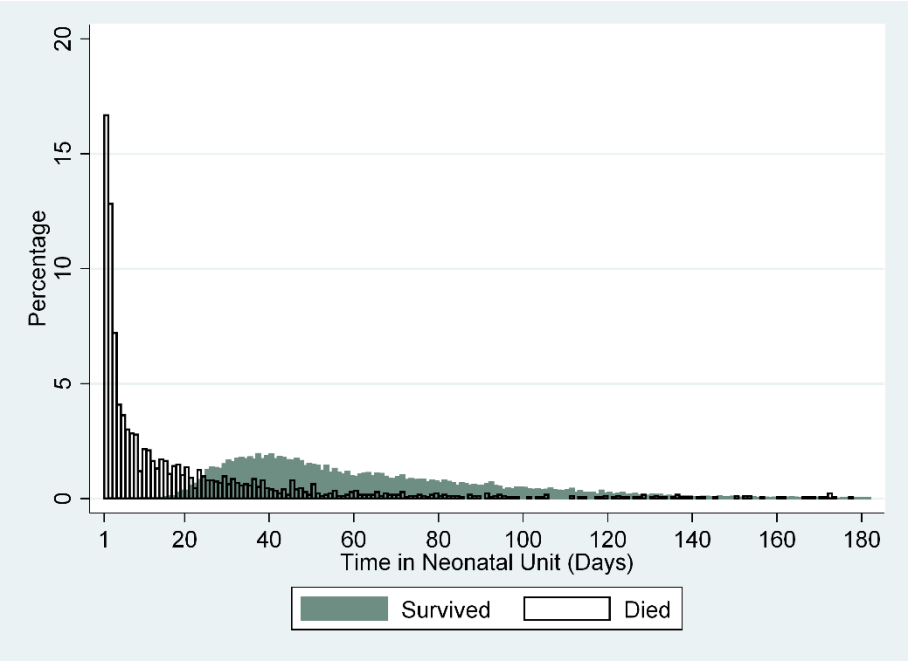


Table 3-8: Observed median length of stay overall and by babies who survived to discharge and babies who died.

Gestational age (weeks)	**Median length of stay (died) (25^{th}, 75^{th})**	**Median length of stay (survived) (25^{th}, 75^{th})**	**Median length of stay (overall) (25^{th}, 75^{th})**
24	8 (2, 23)	119 (105, 135)	102 (16, 125)
25	11 (3, 31)	104 (91, 120)	97 (76, 115)
26	8 (2, 29)	89 (77, 105)	85 (71, 102)
27	3 (10, 32)	77 (66, 91)	74 (62, 89)
28	6 (2, 26)	65 (55, 78)	63 (53, 77)
29	3 (1, 17.5)	52 (44, 64)	51 (44, 64)
30	3 (1, 13)	42 (35, 51)	41 (35, 50)
31	5 (2, 17)	32 (27, 40)	32 (26, 40)

It is possible to separate the cohort of babies who survive into those who were discharged home and those discharged to another hospital (Figure 3-2). These groups are combined together here to form a 'discharge from neonatal services' group. Babies are discharged to other hospitals not captured by the NNRD throughout the days after birth and there appears to be no time points where this is more likely.

Length of stay and the proportion of babies who die varies according to gestational age with babies born earlier having a higher proportion of deaths and longer lengths of stay (Figure 3-3). For all babies, most deaths occurred in the first 50 days of life, with the majority occurring in the first week. Early discharges began to occur for babies born at 24 weeks at around 80 days, for babies born at 28 weeks discharges began to occur at around 40 days and for babies born at 31 weeks, discharges occurred from around 20 days after birth.

Figure 3-2: Observed length of stay for all babies separated by their final discharge destination: home, death or another hospital.

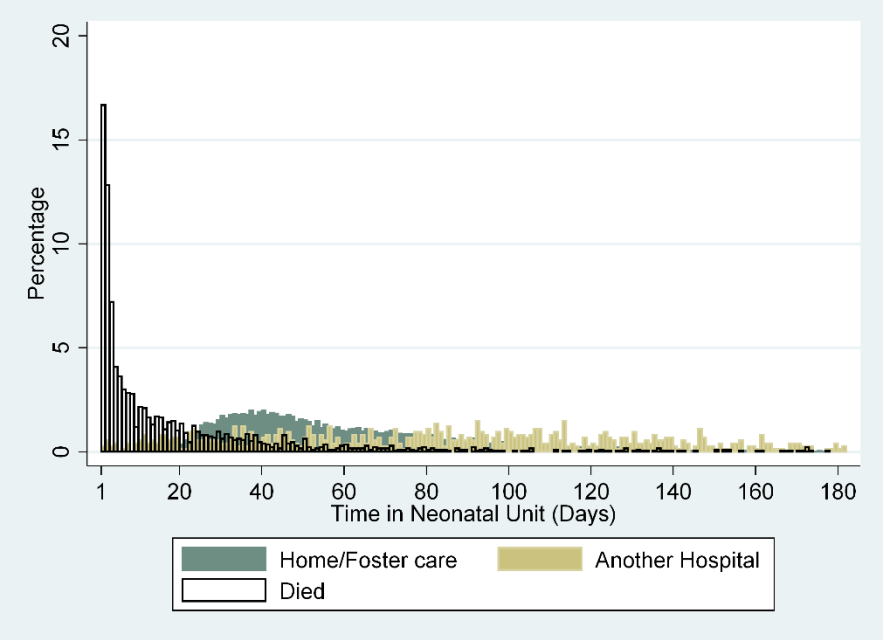


Figure 3-3: Observed length of stay in neonatal care, by babies who survived to discharge or who died in neonatal care, by week of gestational age.

3.7 DATA QUALITY

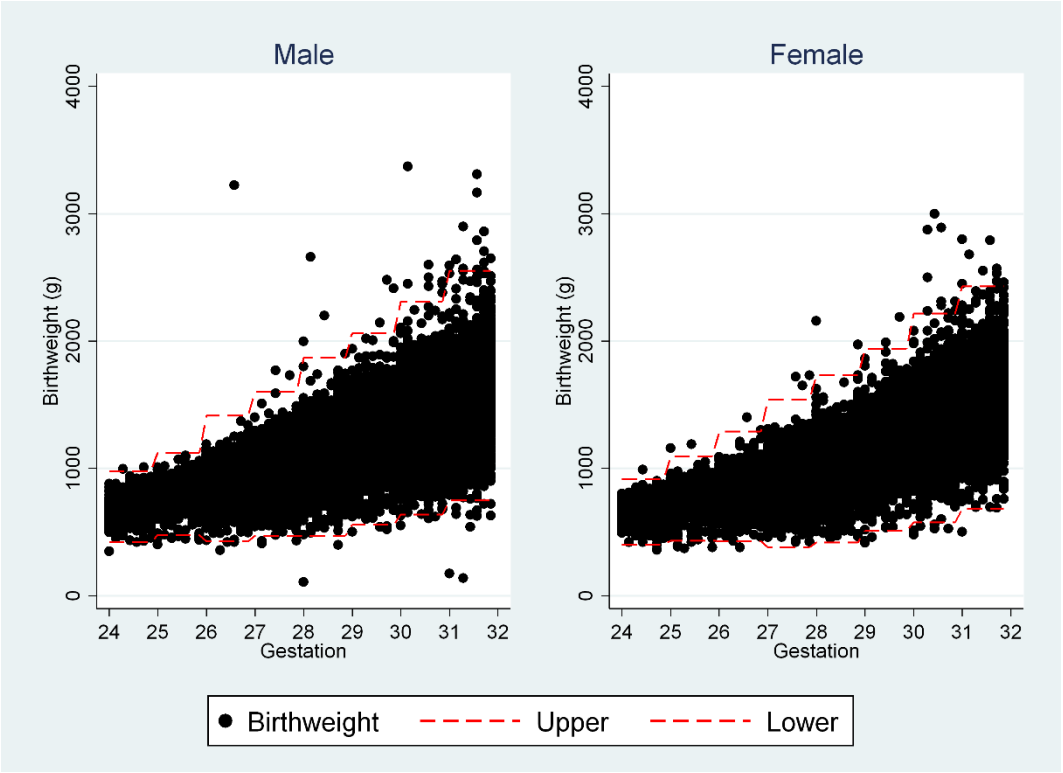
The NNRD is a data source created from routine medical sources via the Badger.net system. Data are entered by clinical staff and the NNRD contains approximately 400 data items making it one of the most detailed population level neonatal datasets available (68). However, as it was never created for research purposes, it may potentially suffer from data quality and completeness issues.

3.7.1 MISSING DATA AND PLAUSIBLE VALUES

In this thesis babies born at 24 to 31 weeks gestational age were considered. Overall data completeness for key demographic variables was good with no missing data related to the sex of the baby (n=20 babies were of unknown or ambiguous sex), less than 0.5% missing information related to days of care and only one case of missing data for birthweight.

Data quality checks and assurances are imposed on the NNRD data to ensure that results are not only complete but also reliable (68). This is facilitated by identification of missing and inconsistent values which are fed back to neonatal units for correction. However, even following these checks, certain variables may have extreme values which may be due to incorrect data entry, or because they are a correct and valid results which happens to be extreme or unusual. For example, Figure 3-4 provides a scatter plot of birthweight against gestational age for males and females. When birthweight is considered in this thesis those babies with missing data (n=1) or a birthweight more than three standard deviations from the observed median for a given week of gestational age (n=117) by sex are excluded (74). Similarly, the 20 babies with ambiguous sex are excluded when sex is considered in the analyses.

Figure 3-4: Birthweight for gestational age (plotted in weeks and days to aid reading of the graph) with three standard deviations highlighted (upper/lower).



3.8 CHAPTER CONCLUSION

This chapter has provided an overview of the NNRD dataset used throughout this thesis. A brief overview of the ethical and R&D approvals obtained to use the NNRD data for this research project was provided. Agreement was obtained from 100% of neonatal units in England by December 2014 to allow subsequent transfer of the data in early 2015, followed by a data update in summer 2015. Summaries of the data used throughout this work and explanations for the relevant exclusions were also provided.

4 SURVIVAL ANALYSIS FOR THE ESTIMATION OF NEONATAL LENGTH OF STAY

4.1 CHAPTER OVERVIEW

Survival analysis considers the modelling of time until an event of interest, often death, occurs. In this chapter, the theory and concepts of survival analysis will be introduced, notably the survival and hazard functions. The most commonly used statistical method, the Cox proportional hazards model, will be described in detail and a preliminary analysis investigating the time babies spend in the neonatal unit will be undertaken. Alternative approaches to undertaking survival analysis, including use of flexible parametric models, will be explored. The strengths and limitations of these approaches to measure length of stay in neonatal care will be discussed before potential extensions are investigated in subsequent chapters.

4.2 INTRODUCTION

Survival analysis is a branch of statistics that describes a range of statistical methods concerned with the measurement of time until an event occurs. As this form of analysis is frequently used to investigate mortality it is known as 'survival' analysis although it is also known as time-to-event analysis, reliability analysis (engineering) and event history analysis (social sciences). The occurrence of a particular event of interest is often described as a 'failure' although there is no requirement for the event to be a negative outcome. Survival analysis can be applied to outcomes other than mortality, for example time to reoccurrence of an infection or time to discharge from hospital.

In a conventional survival analysis time is measured between two events. One example which relates to neonatal care would be to set the time origin (start of the analysis time period) as the time point when a baby is born (Figure 4-1) and measure time until a potential event occurs, such as death. If the participant does not experience the event during the follow up time then their observation is censored (see Chapter 4.3).

Survival analysis offers advantages over other methods as these approaches can handle time-to-event data which are usually skewed distributions, often with many early events and few late ones. If there is no censoring, then other statistical approaches, for example using a logistic regression to predict survival to certain time points (e.g. survived to 30 days: yes/no), could be used although these may still not answer the question of interest. The probability of experiencing the event as a function of time is often of more interest than simply whether an event occurs or not and survival analysis allows this investigation of time.

Survival analysis methods can be used to investigate the time that passes before a baby leaves neonatal care. Leaving neonatal care could be via a discharge home or dying in neonatal care, or a combined outcome of both. These approaches can be used to help answer questions including: *"How long will a preterm baby spend in neonatal care?"* or, if death is considered in the analysis: *"What is the probability a preterm baby will die during their time in neonatal care?"*

Figure 4-1: A pictorial representation of a standard survival analysis measuring time to event: death or discharge.

4.3 Censoring

Not all subjects will necessarily experience the event of interest during the time window observed by the study. Alternatively, they may experience the event at a time not observed by the research team. The approach to handle this issue is known as **censoring**.

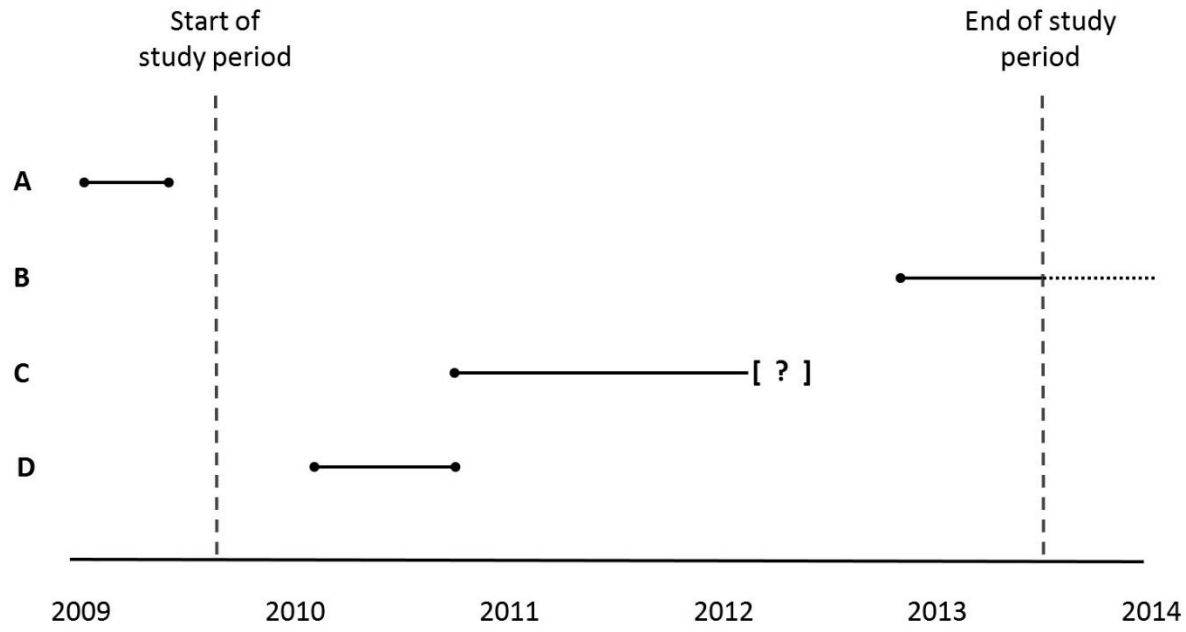
Overall, there are three types of censoring: left-, right- and interval-censoring. Left-censoring is when the event occurs before follow-up time has begun (Patient A in Figure 4-2). Right-censoring occurs when the event of interest happens after the conclusion of the study's data collection. For example, if the final event was discharge

from the neonatal unit, and a baby still remained in hospital when the data collection was concluded, the observation would be right-censored (Patient B in Figure 4-2). This is also known as administrative censoring.

Interval-censoring occurs when the event happens between two points in time (Patient C in Figure 4-2), but it is unclear exactly when it has occurred. For example, in a study of infection rates a patient may experience reinfection between two appointments and it is unclear when this has happened. Patient D in Figure 4-2 is an example of a person who is part of the study and experienced their event during the study and therefore is not censored.

The censoring mechanisms are usually unknown and therefore assumptions are made about censoring (75). Censoring is assumed to be non-informative in this thesis, i.e. the survival times provide no information about the distribution of the censoring times and vice versa. An alternative way of considering this is that an individual who is censored is no healthier or sicker than those who remain in follow up and they have not dropped out of the study because of any impact on them because of the study. The assumption of non-informative censoring is usually considered valid for instances of right-censoring, but not if participants in a study are lost to follow up (76) as this may be an indicator of a different population, for example if all losses to follow up are patients who are discharged to hospices.

Figure 4-2: Pictorial representation of left- (A), right- (B) and interval- censoring (C) and no censoring (D).



4.4 SURVIVAL AND HAZARD FUNCTIONS

The two main components of interest in a survival analysis are the **survival function** and **hazard function**. The survival function $S(t)$ measures the proportion of subjects who have not experienced the event as a function of time, i.e. the proportion still alive over time if the outcome is mortality. The rate of the decline of the survival function depends on the risk of experiencing the event and this is known as the hazard function.

4.4.1 SURVIVAL FUNCTION

The time point T is a non-negative random variable that denotes the time when the event will occur and t is some specified time such that T occurs at a point in time after t. The survival function $S(t)$ represents the probability that a patient has not had an event at time t and is defined as:

$$S(t) = P(T \geq t)$$

Equation 4-1

At time 0, no subjects will have experienced the event and therefore $S(0) = 1$ (i.e. everyone is event-free at the start of follow up).

The probability density function $f(t)$, which is the probability of experiencing an event at time t is:

$$f(t) = \lim_{\delta \to 0} \frac{P(t \leq T \leq t + \delta)}{\delta}$$

Equation 4-2

where δ is a small increase in time and the distribution of event times $F(t)$ can be written as:

$$F(t) = P(T < t) = \int_0^t f(u)du$$

Equation 4-3

The survival function, and the event times relate by:

$$S(t) = P(T \geq t) = 1 - P(T < t)$$

or alternatively

$$S(t) = \int_t^\infty f(u)du$$

Equation 4-4

It is generally most informative to produce a plot of the survival function against time to aid interpretation.

4.4.2 HAZARD FUNCTION AND RATIO

The rate at which the survival function decreases depends on the risk of experiencing the event at time t and this is known as the hazard function: $\alpha(t)$. The hazard function is the instantaneous rate of failure at time t, given that the participant has not

experienced the event so far. The hazard function is an estimate of the incidence rate and the unit is the number of events per unit of time. This can be written as:

$$\alpha(t) = \lim_{\delta \to 0} \frac{P(t \leq T \leq t + \delta | T \geq t)}{\delta}$$

Equation 4-5

It is possible to relate the hazard and survival function to each other mathematically. The hazard is the probability of experiencing the event in the next instant given that it has not been experienced so far:

$$\alpha(t) = \frac{f(t)}{S(t)}$$

Groups can be compared using **hazard ratios** in a survival analysis, which is the ratio of the hazard rates between different levels of an explanatory variable. The hazard ratio $HR(t)$ between two groups A and B is denoted:

$$HR(t) = \frac{\alpha_A(t)}{\alpha_B(t)}$$

Equation 4-6

The ratio indicates if the hazard is higher (hazard ratio > 1) or lower (hazard ratio < 1) in one group than the other. For example, a hazard ratio of two can be interpreted as that at any time point, twice as many participants in group A compared to group B experience the event. The hazard ratio between the groups is reported as one number and is assumed to be constant over time. This is known as the proportional hazards assumption (see Chapter 4.7.1). However, even when the proportional hazards assumption is relaxed it is possible to estimate hazard ratios, but as these will vary over time the values should be presented graphically against time.

4.4.3 CUMULATIVE HAZARD FUNCTION

A final useful statistic is the **cumulative hazard function** $\mathrm{A}(t)$ which is defined as:

$$\mathrm{A}(t) = \int_0^t \alpha(u)du$$

Equation 4-7

This is a measurement of how much 'hazard' a person has been exposed to by time t, or how much hazard they have accumulated. Statistical approaches that model on the cumulative hazard scale will be introduced in Chapter 5. The survival function and the cumulative hazard are related by:

$$S(t) = \exp(-A(t))$$

Equation 4-8

4.5 APPROACHES FOR MODELLING TIME-TO-EVENT DATA

There are several statistical approaches that can be used in survival analysis. Broadly, these fall into four groups: non-parametric, semi-parametric, parametric and flexible parametric analyses.

The Kaplan-Meier estimator (77) is a non-parametric statistic used to estimate the survival function. The estimator can be plotted and this appears as series of steps which decrease each time there is an event. With a large sample size, or minimal censoring, the Kaplan-Meier estimator approaches the observed survival. This will be introduced in Chapter 4.6.

The use of semi-parametric models makes no assumption about the shape of the baseline hazard function. The baseline hazard function is the value of the hazard function when all covariates are set to zero. This approach is known as Cox proportional hazards modelling and this is the most common survival model used in medical statistics (78). The estimation of the Cox model does not directly model the baseline hazard, although it is possibly to subsequently estimate it if required, and therefore it is only possible to estimate the survival function at the times when events

are observed. This leads to a step function for the survival function. The Cox model will be discussed further in Chapter 4.7.

Parametric methods assume a functional form for the shape of the baseline hazard, for example a Weibull distribution. Whilst this approach can be flexible, it forces the hazard to be monotonic (increasing or decreasing with no turning points) leading to a lack of flexibility in the shape of the baseline hazard which may not be sensible to assume (79). However, parametric methods have advantages over semi-parametric approaches as the modelling of the hazard means it is possible to model time points where no events occurred, and extrapolation beyond the final time point is also possible. This approach is discussed further in Chapter 4.8.

Finally, flexible parametric models are an extension of parametric models, and overcome the limitation imposed by the modelling of the baseline hazard using a parametric shape (80). There are a range of flexible parametric models available and one approach extends Weibull (parametric) modelling via use of restricted cubic spline functions to model the baseline cumulative hazard. Models are estimated on the log cumulative hazard scale as this is generally a stable function. A further discussion is found in Chapter 4.8.

4.6 NON-PARAMETRIC ESTIMATION

The Nelson-Aalen estimator (81) is the non-parametric estimate of the cumulative hazard rate and is denoted:

$$\hat{A}(t) = \sum_{i=1}^{n} \frac{N_i}{Y_i}$$

Equation 4-9

where N_i is the number of events and Y_i is the number of participants at risk at time t_i. The shape of the Nelson-Aalen estimator provides an indication of the shape of the hazard rate.

The Kaplan-Meier estimator (77) is the non-parametric estimate of the survival function. The estimator of the survival function $\hat{S}(t)$ is denoted:

$$\hat{S}(t) = \prod_{i=1}^{n} \left(1 - \frac{N_i}{Y_i}\right)$$

Equation 4-10

where t is the duration of the study, N_i is the number of events up to t_i and Y_i is the number of participants at risk just prior to t_i.

4.7 COX PROPORTIONAL HAZARDS MODEL

One approach for modelling survival data is the Cox proportional hazards model (78). A Cox model can be considered as two components: the hazard function and the effects of the covariates. The hazard rate can be estimated, given the covariates, as:

$$\alpha(t|\mathbf{Z}) = \alpha_0(t)\exp(\boldsymbol{\beta}'\mathbf{Z})$$

Equation 4-11

Written in this form, $\alpha_0(t)$ is the baseline hazard and $\boldsymbol{\beta}'\mathbf{Z}$ is the linear predictor ($\boldsymbol{\beta}'$is the regression coefficients and $\mathbf{Z}$ is the value of the covariates) which describes how the hazard varies in response to the values of the covariates. The baseline hazard function is found when all covariates have been set to zero and the estimates of the regression coefficients are obtained by maximising the partial likelihood (78) which can be written as:

$$\prod_{i=1}^{n} \left(\frac{\exp(\boldsymbol{\beta}'\mathbf{Z_i})}{\sum_{i \epsilon R(t_i)} \exp(\boldsymbol{\beta}'\mathbf{Z_i})} \right)^{d_i}$$

Equation 4-12

where i represents the individuals still at risk of event d, t_i are the events times and $R(t_i)$ is the risk-set, the group of individuals still alive and uncensored, prior to time t_i. The partial likelihood is independent of the baseline hazard, as can be seen from

Equation 4-12 which does not rely on $\alpha_0(t)$. Therefore the regression coefficients can be estimated with no information about the baseline hazard. If the baseline hazard is required then estimators developed by Breslow or Kalbfleisch and Prentice can used (82, 83). The Breslow estimator maximises the baseline hazard by substituting the estimates of the regression coefficients into the full likelihood function. Full details can be found in the references: (82, 83).

The survival function and cumulative hazard for the Cox model, given covariates, are defined as:

$$S(t|\boldsymbol{Z}) = S_0(t)^{\exp(\boldsymbol{\beta}'\mathbf{Z})}$$

Equation 4-13

$$A(t|\boldsymbol{Z}) = A_0(t)\exp(\boldsymbol{\beta}'\mathbf{Z})$$

Equation 4-14

The Cox model is a semi-parametric model as no assumption is made about the shape of the baseline hazard as this is not directly estimated (84). This property can be useful when hazards are likely to have an unusual shape not easily captured by other distributions. However, a consequence of this is that proportional hazards, that the hazard ratio is constant over time, must be assumed. This is not a property isolated to the Cox model, and can be seen when considering the calculation of the hazard ratio. If the model only contains a binary variable $x_1$which takes values 0 or 1 then the hazard rate when $x_1 = 0$ is $\alpha_0(t)$ and when $x_1 = 1$ the hazard rate is $\alpha_0(t)\exp(\beta)$. Therefore, the hazard ratio, comparing one group to the other, is:

$$\frac{\alpha_0(t)\exp(\beta)}{\alpha_0(t)}$$

The baseline hazards $\alpha_0(t)$ cancel out and β is the log hazard ratio which can be exponentiated to give the hazard ratio, which does not depend on time. This approach and many other survival analysis methods are modelled on the log hazard scale.

There are situations when there can be ties in the time that events occur, particularly when multiple events can occur at the same time. This is often the case in neonatal care as the measurement of time for the commissioning of care is in whole days. This is

not because these data are interval-censored (see Chapter 4.3), but because commissioning occurs on a day-to-day basis and therefore irrespective of how much of a day occurs, a whole day is counted for cost purposes. Additionally, the deaths in neonatal care often occur in the early days of life, adding more ties to the data. Ties in the data can create issues with the calculation of the partial likelihood and to accommodate this methods are proposed to break the tie, although most software uses a default of the Breslow method (85). The Breslow approximation to the partial likelihood sums the covariate patterns for each subject experiencing the event at the same time point and raises the result to a power equal to the number of events (d_j) which have tied at that time point (84, 86). This can be applied to Equation 4-12 to give an adjusted partial likelihood for participant i at the j^{th} failure time:

$$\frac{\prod_{i \in R_j} \exp(\boldsymbol{\beta}' \mathbf{Z_i})}{\left\{\sum_{i \in R_j} \exp(\boldsymbol{\beta}' \mathbf{Z_i})\right\}^{d_j}}$$

Equation 4-15

Other approaches to handle tied event times are the Efron approximation, the Kalbfleisch and Prentice exact expression or by subtracting a tiny random amount from each tied survival time (86). In this thesis, I use the Breslow estimator to correct for tied times in the data.

4.7.1 Proportional hazards assumption for the Cox model

If proportional hazards are assumed then this can be formally tested using Schoenfeld residuals and the Therneau-Grambsch test. Schoenfeld residuals are the observed covariate values for an individual minus the expected value, and there is a residual for every individual who experienced an event, for every covariate in the model. These residuals can be plotted against a function of time (87) and investigated for time trends. To aid examination of whether there is a trend a smoothed line is usually fitted. If there are no issues of proportional hazards then the residuals will be randomly scattered and the smoothed line will be relatively straight (see example in Figure 6-12, babies born at 24 weeks). If this line is increasing, decreasing or changing substantially

over time this indicates that the hazard may not be constant over time (see example in Figure 6-14). Thus, the proportional hazards assumption may be violated.

A formal significance test of the smoothed line used with the Schoenfeld residuals is the Therneau-Grambsch test (85). This test fits a generalised linear model to the residuals and tests if the slope is non-zero. A significant result, here taken to be $p<0.001$, indicates the slope is not flat and that there are potential issues with proportionality. It is not unusual with large datasets to detect very small departures from proportional hazards, or with uncommon events, to over-fit the test. Therefore, the Therneau-Grambsch test should be used alongside the Schoenfeld residual plots and a lower threshold for significance is used in this thesis.

The relaxation of the proportional hazards assumption can be implemented by incorporating time-dependent effects into the statistical model. These are computationally intensive and can be time consuming with standard software in large datasets (88), although they have become much more accessible in standard survival analysis in recent years. This will be discussed further in Chapter 5.

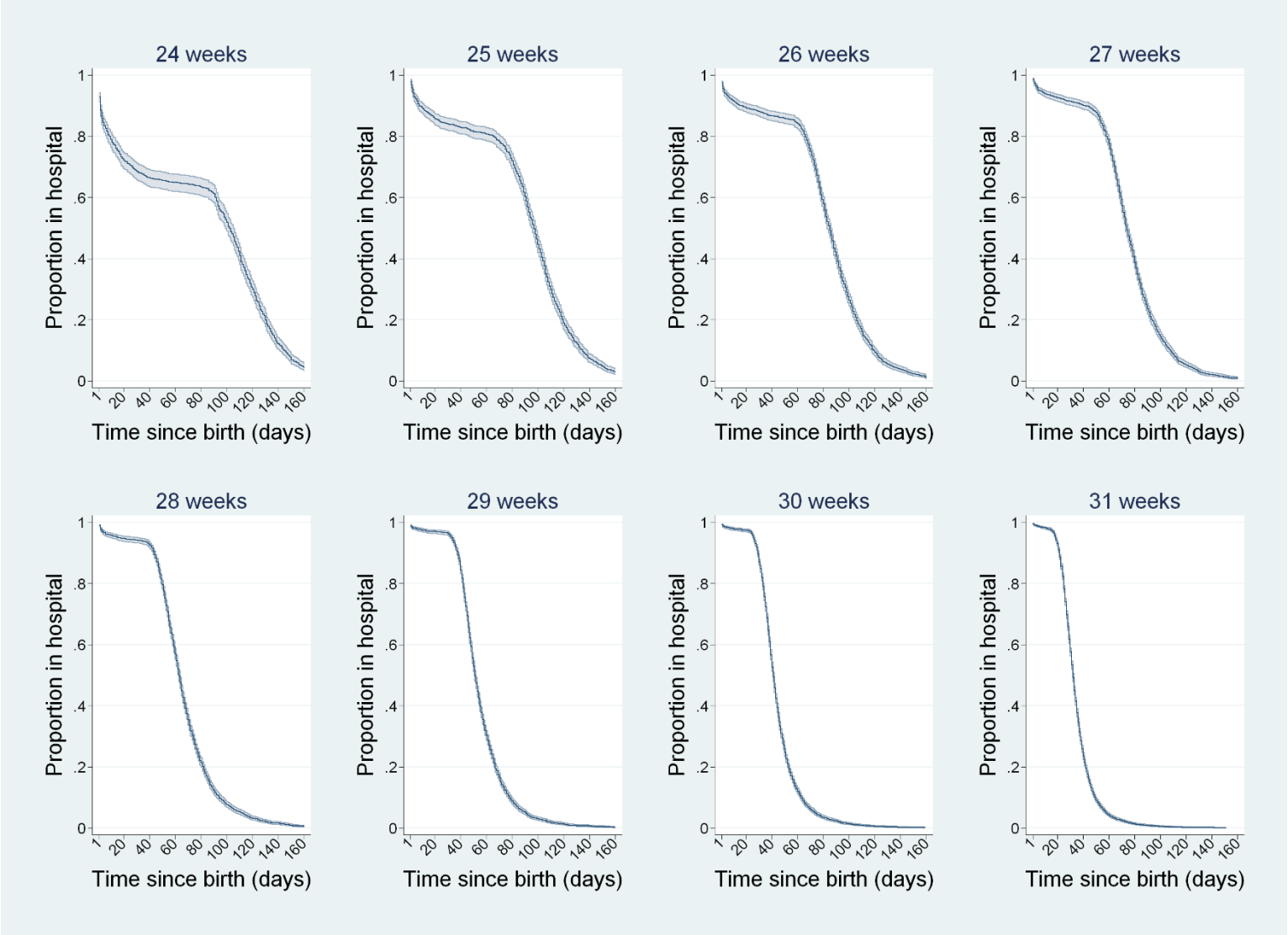
4.7.2 PRELIMINARY SURVIVAL ANALYSIS USING THE COX MODEL

I undertook a preliminary analysis using the Cox model, fitted via the *stcox* command in Stata v 14 to investigate the uses of survival analysis to investigate length of stay. Standard survival analyses such as that introduced here can only consider one outcome, and therefore all babies will be analysed together with a combined end point of having left hospital for any reason, i.e. the analysis measures the time spent in hospital. The 21,038 babies included in this analysis were described in Chapter 3. The distributions of length of stay for babies that survived and babies that died were seen to be very different in Figure 3-1, although in this section of the thesis they are combined.

Kaplan-Meier plots for the proportion of babies remaining in hospital over time are presented in Figure 4-3 by week of gestational age at birth. At the earlier weeks of gestational age at birth, the Kaplan-Meier curves initially decrease, then plateau and then decrease again. The initial decrease in the curve is created by the occurrence

deaths in the initial days after birth. The curve then plateaus at the approximate death proportion for that gestational age, although some deaths do occur later as seen in Figure 3-1. The curve then decreases again as the discharges from the neonatal unit start to occur. All curves eventually end at around zero when all babies will have either died or been discharged from neonatal care.

Figure 4-3: Kaplan-Meier plots to denote the proportion of babies remaining in neonatal care over time by week of gestational age.



A Cox proportional hazards model was fitted including gestational age at birth, modelled categorically with 27 weeks as the baseline group. Babies born at the extremes of gestational age in this cohort (24 weeks and 31 weeks) are very different from the other babies and therefore 27 weeks of gestational age was chosen to provide a more meaningful comparison. The hazard ratios, with their 95% confidence intervals and p-values, are presented in Table 4-1. These are the hazard ratios for leaving hospital for each week of gestational age at birth.

The combined endpoint of time to death or discharge results in hazard ratios which relate to the hazard of leaving hospital. At any point in time the babies born at 31 weeks leave hospital at 6.4 times the rate of those born at 27 weeks (95% CI: 6.07 to 6.76, p<0.001). Conversely, the babies born at 24 weeks are at a 41% reduced hazard of leaving hospital at all time points compared to those born at 27 weeks (95% CI: 0.54 to 0.63, p<0.001).

Table 4-1: Hazard ratios with 95% confidence interval for the hazard of leaving hospital.

Gestational age (weeks)	**Hazard ratio**	**95% Confidence Interval**	**p-value**
24	0.59	0.54, 0.63	<0.001
25	0.61	0.57, 0.66	<0.001
26	0.77	0.73, 0.83	<0.001
27	Baseline	Baseline	Baseline
28	1.34	1.27, 1.42	<0.001
29	2.07	1.96, 2.19	<0.001
30	3.48	3.30, 3.68	<0.001
31	6.40	6.07, 6.76	<0.001

The survival probability, here the probability of remaining in hospital, can be estimated using Equation 4-13. These probabilities are presented in Figure 4-4 by week of gestational age. These probabilities are a step function as the survival probability is only estimated at observed event times. However, as this is a large dataset, with many events, the curves appear reasonably smooth.

Figure 4-4: The predicted probability of remaining in hospital over time estimated from the Cox model by week of gestational age.

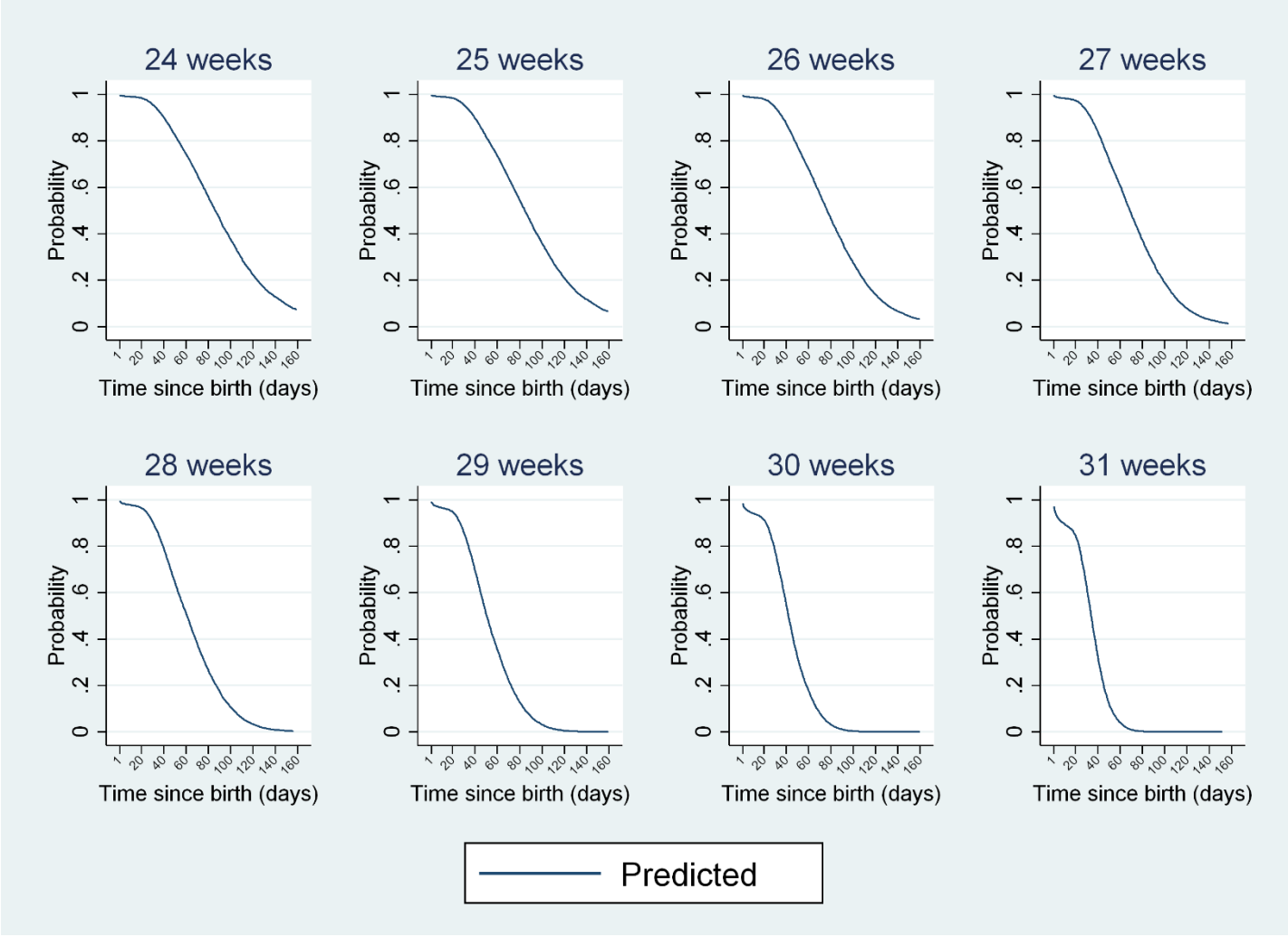
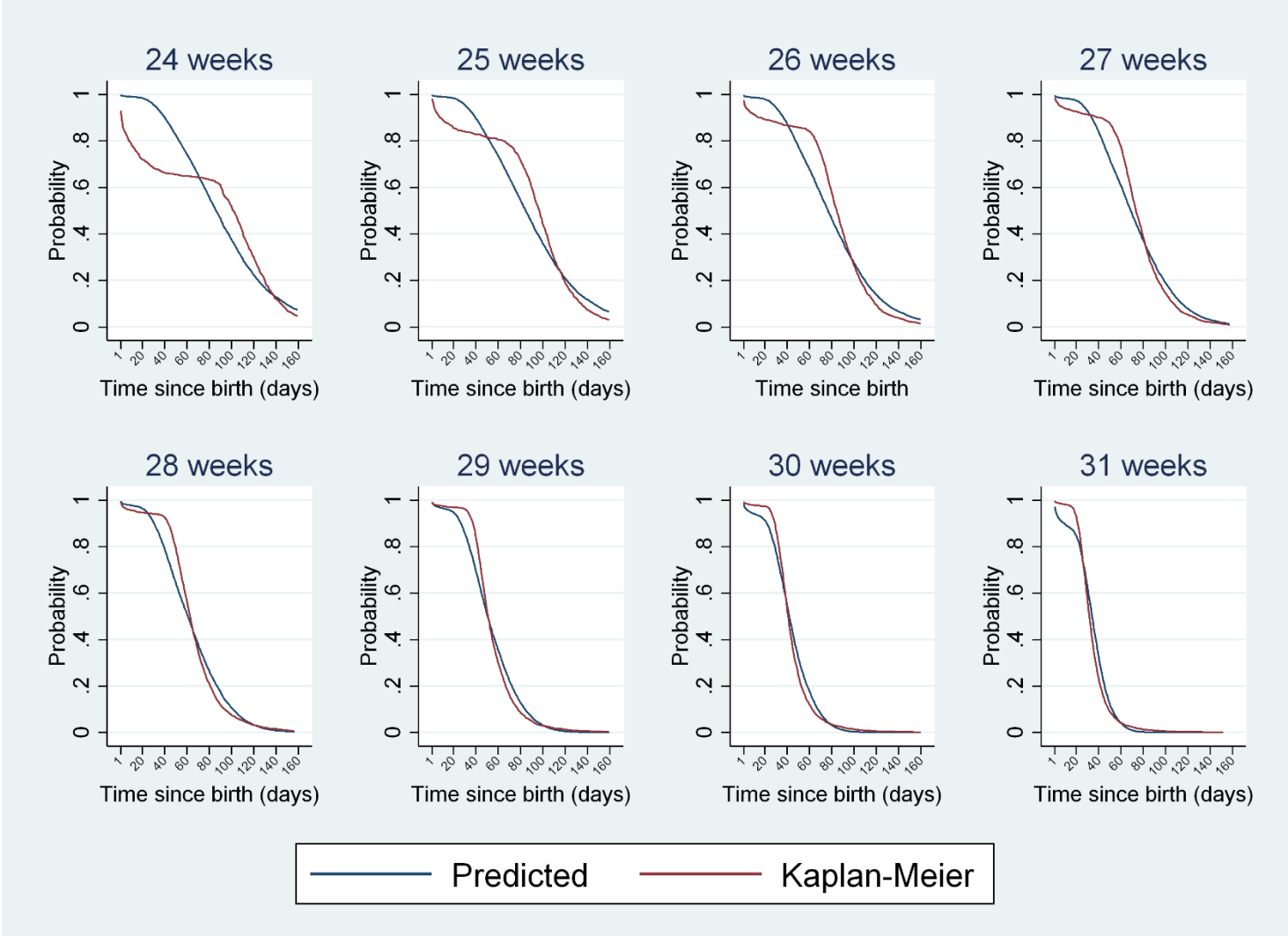


Figure 4-5: The predicted probability of remaining in hospital estimated from the Cox model overlaying the Kaplan-Meier estimates, by week of gestational age.



The model fit seen here was poor which can be seen when comparing the Kaplan-Meier estimates (Figure 4-3) with the probabilities predicted from the Cox model (Figure 4-4) which should broadly follow similar shapes. These have been overlaid together in Figure 4-5 to demonstrate the differences. This poor fit is due to two inter-related reasons: 1) the assumption of proportional hazards is likely to not hold and 2) the model fit is driven by the babies born at 30 and 31 weeks gestational age who have a different shaped '*survival*' as seen in Figure 4-3.

The predicted median length of stay in hospital for all babies can be identified from Figure 4-4 and these are also provided by week of gestational age with range in Table 4-2. For example, 50% of babies born at 24 weeks remain in hospital at least 87 days after birth (Table 4-2). For babies born at 31 weeks gestational age, 50% still remained in hospital 34 days after birth. Other percentiles could be identified and these could be used as estimates of length of stay.

The predictions of median length of stay are different to the observed median length of stay (Table 4-2), particularly for babies born at less than 28 weeks gestational age where mortality is high. However, the length of stay is estimated well when the mortality rate is lower.

Table 4-2: Predicted (from the Cox model) and observed median length of stay for all babies by week of gestational age, with 10th and 90th centile.

Gestational age (weeks)	**Predicted median length of stay (days)**	**Observed median length of stay (days)**
24	87 (40, 151)	102 (2, 145)
25	85 (39, 148)	97 (8, 135)
26	77 (36, 129)	85 (16, 119)
27	69 (33, 118)	74 (42, 108)
28	61 (30, 104)	63 (44, 94)
29	51 (25, 86)	51 (38, 78)
30	42 (16, 69)	41 (29, 63)
31	34 (11, 54)	32 (22, 50)

One approach to improve the model fit is the introduction of time-dependent covariates, thereby relaxing the assumption of proportional hazards, allowing the estimated hazards for the different weeks of gestational age to not be proportional

with each other. This will be discussed in more detail in Chapter 5. An alternative approach is by stratifying the analysis to babies who are similar to each other.

4.7.3 STRATIFICATION OF THE COX MODEL

The model fit seen before was poor, particularly when comparing the Kaplan-Meier estimates with the probabilities predicted from the Cox model (Figure 4-5). The Cox model is a flexible approach which makes no assumption about the shape of the hazard, and could therefore capture a shape with two or more turning points. However, as the babies born at 30 and 31 weeks contribute nearly 40% of all births (Table 3-1) the model fit is largely driven by this group, who have a low mortality rate and shorter length of stay compared to those born at less than 28 weeks gestational age.

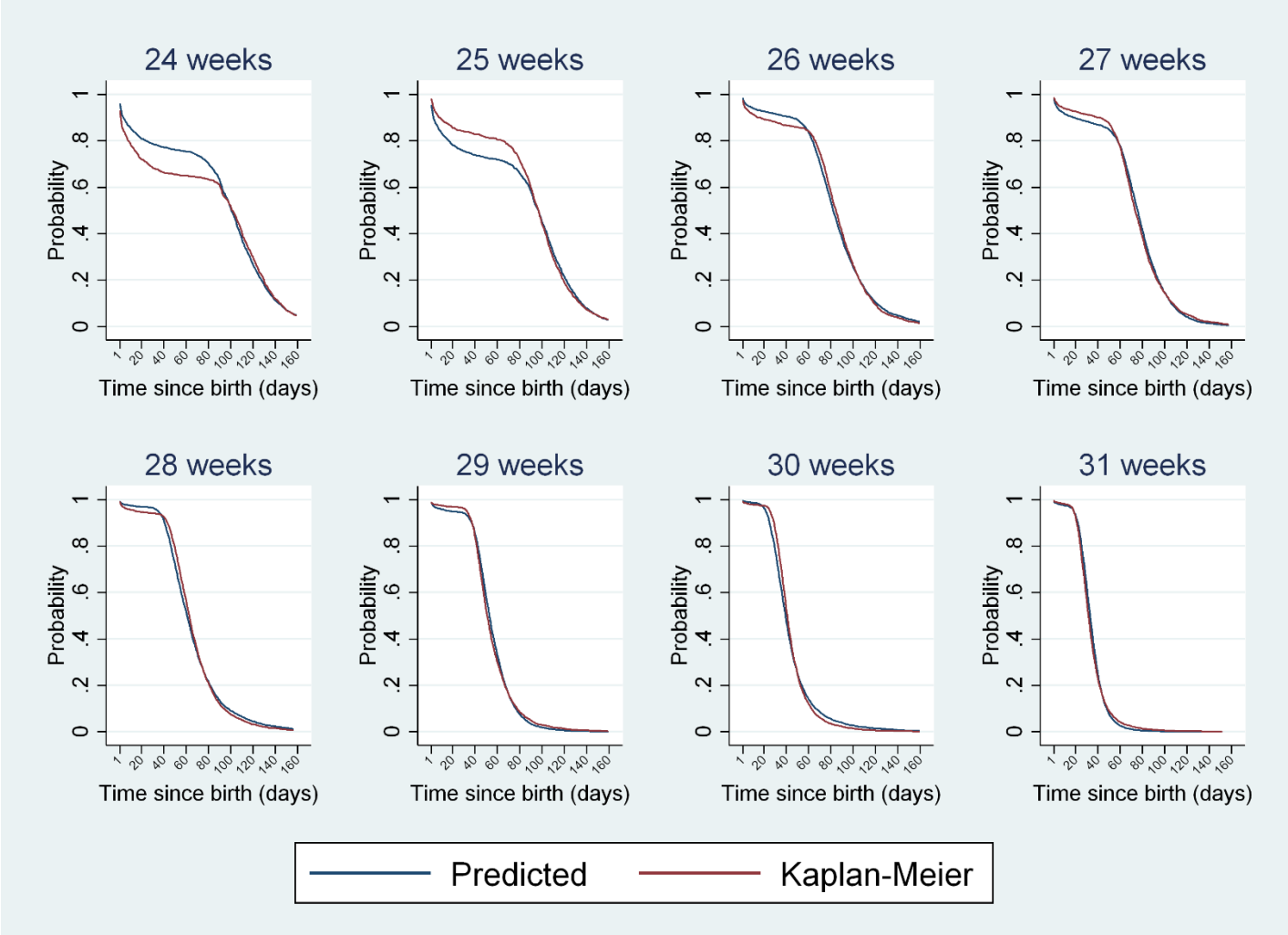
In Figure 4-6 the analysis is stratified by babies born at a gestational age of 24 to 25 weeks; 26 to 27 weeks; 28 to 29 weeks and 30 to 31 weeks. The predicted probabilities are again overlaid with the Kaplan-Meier curves. There is now greater agreement between the model predictions and observed data, indicating the model has been improved. However, some issues remain with the prediction of probabilities for babies born at 24 and 25 weeks gestational age.

Following stratification, the estimated median length of stay in hospital (Table 4-3) is now closer to the observed median time across all gestational ages. Even for babies born at 24 and 25 weeks the median length of stay is close to the observed data indicating the model performs well for estimating length of stay. Therefore, the Cox model can potentially be flexible enough to capture this complex shape which varies between weeks of gestational ages, if proportional hazards can be relaxed in some way such as via stratification.

Table 4-3: Predicted (from the stratified Cox model) and observed median length of stay for all babies by week of gestational age, with 10th and 90th centile.

Gestational age (weeks)	**Predicted length of stay (days) (10th, 90th)**	**Observed median length of stay (days) (10th, 90th)**
24	102 (7, 147)	102 (2, 145)
25	97 (5, 139)	97 (8, 135)
26	83 (33, 125)	85 (16, 119)
27	76 (27, 109)	74 (42, 108)
28	61 (40, 103)	63 (44, 94)
29	53 (30, 79)	51 (38, 78)
30	39 (25, 71)	41 (29, 63)
31	34 (22, 50)	32 (22, 50)

Figure 4-6: Predicted probabilities (from the stratified Cox model) of remaining in hospital by week of gestational age overlaid with the Kaplan-Meier plots.



4.8 EXTENDING THE COX MODEL

An alternative approach to the Cox proportional hazards model is use of a parametric model. This approach replaces the baseline hazard $\alpha_o(t)$ with a specified function. Various distributions can be used to specify the analytical form of the baseline hazard including the Exponential; Gompertz and Log-Normal (81). The Weibull distribution is commonly used. If the survival times are assumed to have a Weibull distribution $W(\delta, \gamma)$ then the survival and hazard functions are defined as:

$$S(t) = \exp(-\delta t^{\gamma})$$

$$\alpha(t) = \delta\gamma t^{\gamma-1}$$

Equation 4-16

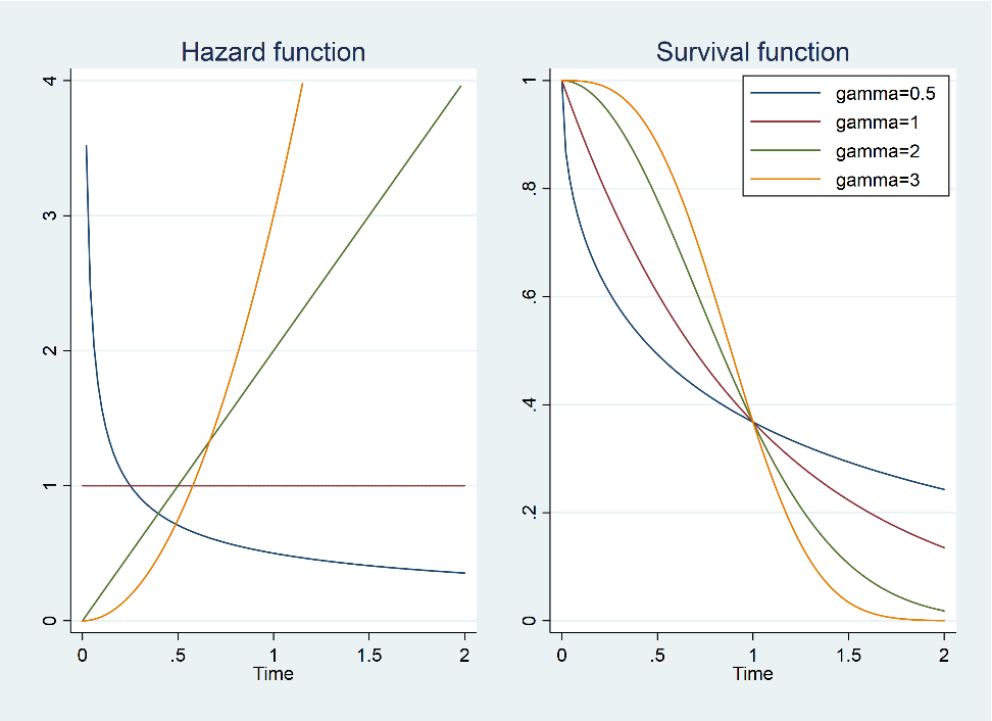
where the values of δ and γ relate to the scale and shape parameters of the Weibull distribution and can provide a range of different shapes for the hazard and survival function (Figure 4-7). If the value of $\gamma > 1$ then the hazard is increasing, constant if $\gamma = 1$ (this is also the exponential model) and if $\gamma < 1$ then the hazard is decreasing. Whilst a variety of hazard shapes can be obtained, these are restricted to be monotonic, i.e. have no turning points, whereas the Cox model does not impose this condition. Therefore parametric models may be unable to capture the shape of the hazard (89).

The cumulative hazard is defined as:

$$A(t) = \delta t^{\gamma}$$

Equation 4-17

Figure 4-7: Different shapes of hazard and survival functions obtained from the Weibull distributions with values of γ=0.5, 1, 2 or 3 and $\delta = 1$.



4.8.1 Extending the Weibull model to the Flexible parametric approach

The Weibull model can be extended further to increase the flexibility. One approach (89) considers the Weibull cumulative hazard (Equation 4-17) written in logarithmic form as a constant term and a linear function of log time:

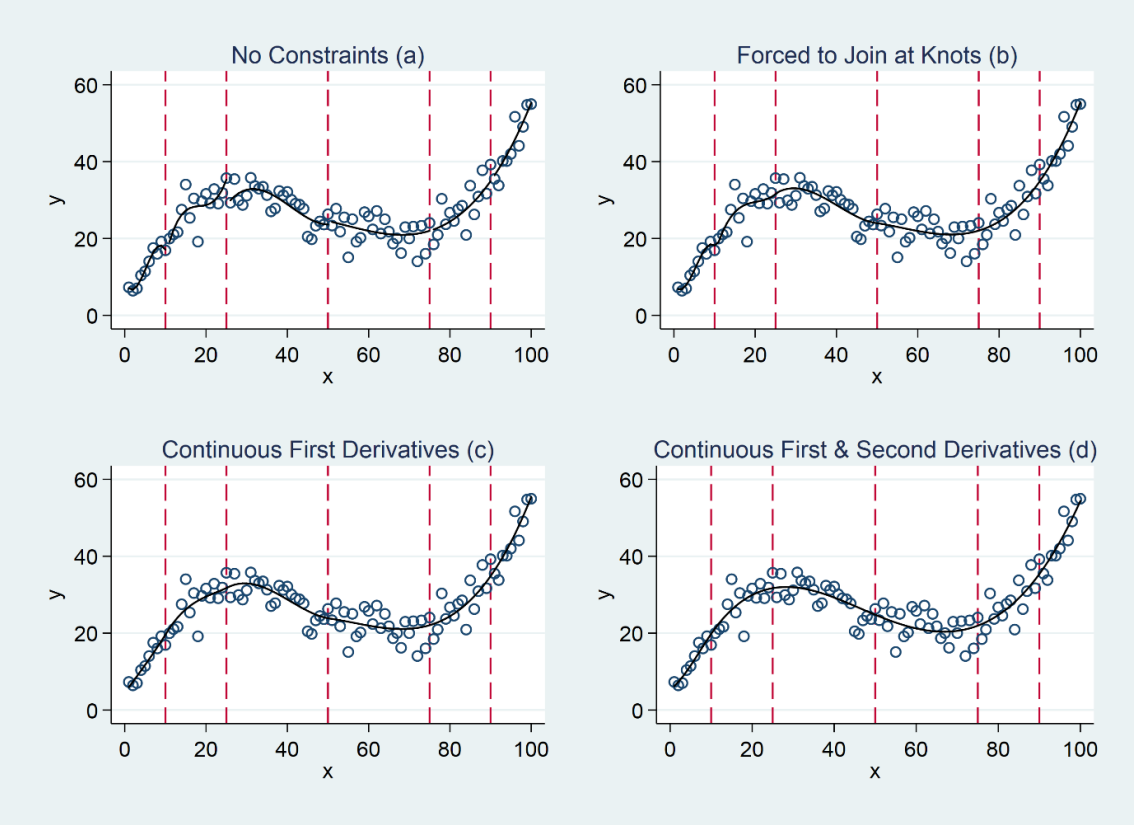
$$\ln A(t) = \ln(\delta) + \boldsymbol{\gamma}_1 \ln(t)$$

Equation 4-18

This function of time may not have adequate flexibility to capture the shape of the cumulative hazard and extensions were proposed by Royston and Parmar (80) and subsequently extended by Lambert and Royston (79). The revised approach uses restricted cubic splines (90), piecewise polynomial functions joined at pre-specified time-points known as knots, to model the baseline cumulative hazard.

A visual demonstration of splines is presented in Figure 4-8 where knots, denoted by the red dashed lines, are placed throughout a random function. To create the restricted cubic splines, firstly separate cubic polynomials are fitted between the knots which assume nothing about the polynomials between the other knots (a). The first constraint is that the separate curves must meet at the knots to form a continuous, but not necessarily smooth, function (b). The second constraint forces the function to have continuous first derivatives, which smooths the function at the knots (c). Finally, the second derivatives are forced to be continuous and, with restricted cubic splines, the function is forced to be linear before the first knot and after the final knot (d). The use of these constraints aids the creation of a smooth, non-linear function, known as restricted cubic splines which can be used to model complex shapes. The production of Figure 4-8 uses Stata code provided by Paul Lambert as part of the MSc Medical Statistics (University of Leicester).

Figure 4-8: Constraints placed on non-linear functions to create restricted cubic splines.



Use of restricted cubic splines extends this methodology to flexible parametric modelling. With this approach it is possible to assume proportional hazards providing

similar results to that from the Cox model. When assuming proportional hazards, similarly to Equation 4-14, the cumulative hazard including covariates can be written as:

$$\ln\{A(t|\boldsymbol{Z})\} = \ln\{A_0(t)\} + \boldsymbol{\beta}'\boldsymbol{Z}$$

Equation 4-19

A restricted cubic spline with $\boldsymbol{n_0}$ knots can be denoted as $s\{\ln(t)|\boldsymbol{\gamma}, \boldsymbol{n_0}\}$ and used to re-estimate the baseline log cumulative hazard as:

$$\ln\{A(t|\boldsymbol{Z})\} = \boldsymbol{\eta} = s\{\ln(t)\,|\boldsymbol{\gamma}, \boldsymbol{n_0}\} + \boldsymbol{\beta}'\boldsymbol{Z}$$

Equation 4-20

This can be transformed to the survival and hazard functions as:

$$S(t|\boldsymbol{Z}) = \exp\{-\exp(\boldsymbol{\eta})\} \text{ and } \alpha(t|\boldsymbol{Z}) = \frac{ds\{\ln(t)|\boldsymbol{\gamma}, \boldsymbol{k}_0\}}{dt}\exp(\boldsymbol{\eta})$$

Equation 4-21

In this section the flexible parametric modelling still assumes proportional hazards, although an advantage of this approach is that it is possible to model complex time-dependent effects with relative ease, and relax the proportional hazards assumption, and this will be explored in Chapter 5. Other advantages of flexible parametric modelling are that the survival and hazard function are smooth, there is interpretation of the hazard and survival even at time points when events have not been observed and the baseline hazard is directly estimated (79).

4.8.2 Preliminary survival analysis using the Flexible Parametric model

I fitted a flexible parametric model with four knots using *stpm2* in Stata v 14 which included a covariate for gestational age at birth, modelled categorically with 27 weeks as the baseline group. The hazard ratios, with their 95% confidence intervals and p-values, are presented in Table 4-4. As before with the Cox model (Table 4-1), these are the hazard ratios for leaving hospital for each week of gestational age at birth. This flexible parametric model assumes proportional hazards and therefore the results are

equivalent to that seen in Table 4-1. The probabilities are not presented but these would be equivalent to that seen in Chapter 4.7.2 (Figure 4-4). However, the predicted median length of stay is provided in

Table 4-5 and the estimates are seen to be similar to those in Table 4-2, with only small differences in the range.

Table 4-4: Hazard ratios from the flexible parametric model, with 95% confidence interval, for the hazard of leaving hospital.

Gestational age (weeks)	Hazard ratio	95% Confidence Interval	p-value
24	0.58	0.54, 0.63	<0.001
25	0.61	0.57, 0.65	<0.001
26	0.78	0.73, 0.83	<0.001
27	Baseline	Baseline	Baseline
28	1.34	1.27, 1.42	<0.001
29	2.10	1.98, 2.22	<0.001
30	3.54	3.35, 3.74	<0.001
31	6.40	6.03, 6.72	<0.001

Table 4-5: Predicted (from the flexible parametric model) and observed median length of stay for all babies by week of gestational age, with 10^{th} and 90^{th} centile.

Gestational age (weeks)	Predicted median length of stay (days) (10^{th}, 90^{th})	Observed length of stay (days) (10^{th}, 90^{th})
24	87 (41, 148)	102 (2, 145)
25	85 (40, 145)	97 (8, 135)
26	77 (37, 130)	85 (16, 119)
27	69 (33, 117)	74 (42, 108)
28	61 (30, 103)	63 (44, 94)
29	51 (24, 85)	51 (38, 78)
30	42 (19, 68)	41 (29, 63)
31	34 (13, 54)	32 (22, 50)

4.9 DISCUSSION

In this preliminary analysis, I combined the outcomes of death and discharge to calculate estimates of length of stay using Cox proportional hazards modelling and flexible parametric modelling. For the Cox model, the probability of remaining in the

neonatal unit was calculated by week of gestational age at birth, and the median length of stay was predicted. The purpose of this chapter was to introduce the concepts of survival analysis which will be extended in later chapters.

4.9.1 PROPORTIONAL HAZARDS

The journal article which introduced the Cox proportional hazards model is one of the most cited papers in medicine, having received over 43,000 citations as of February 2017 (78). The wide use of this methodology is due to the ease of implementation and interpretation, and because there is no need to investigate the shape of the baseline hazard. In some circumstances alternative methods, including those that impose parametric assumptions, may lead to inaccurate results (91) and therefore there are advantages to using the Cox model which makes no assumption about the baseline hazard shape (78). However, a criticism by many is that the Cox model is often overused particularly in situations where it may not be appropriate and indeed even Sir David Cox who devised the method has agreed this is true in certain circumstances (92).

The assumption of proportional hazards is often not tested, with a review of cancer studies from 1995 indicating that only 5% of studies investigated the assumption at that time (93). It is possible to assess this assumption after fitting the model using the Therneau-Grambsch test (85, 94) and by plotting Schoenfeld residuals (87). The proportional hazards assumption can be relaxed by incorporating time-dependent effects, that is, an interaction between covariates and time. However, this can be problematic with the Cox model, due to the computational power required (88). An alternative presented here was to stratify the model by groups of babies who share similar characteristics where proportionality within each stratum may be more likely to hold.

Proportional hazards are unlikely to hold in this analysis and the shape of the baseline hazard was driven by the babies born at 30 and 31 weeks gestational age. Most of these babies survive to discharge and therefore the shape of the hazard will be heavily influenced by this group leading to poorly estimated probabilities, particularly for babies born at 24 and 25 weeks gestational age who are very different to the rest of

the cohort. To overcome this, the analysis was stratified so that babies born at similar weeks of gestational age were grouped together.

Whilst stratification overcame some of the issues caused by proportional hazards, the combined endpoint was not informative for use by clinicians in parental counselling. The combining of these two groups makes the estimates of length of stay difficult to interpret as babies who die generally spend less time in hospital than those who survive to discharge. The assumption of proportional hazards was not tested here as this analysis is for illustration as it is fundamentally flawed from a clinical perspective.

The flexible parametric analysis which was undertaken in this chapter also assumed proportional hazards and therefore had the same limitations as the Cox model. However, time-dependent covariates could have been incorporated and this will be discussed further in Chapter 5.

4.9.2 PREDICTION OF LENGTH OF STAY

As the Cox model does not impose a distribution on the shape of the hazard, it is only possible to model events at observed time points, and prediction beyond the range of the data (extrapolation) or where events did not occur within the data (interpolation) is not possible. In these data there are generally events on every day and this potential limitation can be overcome further by using the flexible parametric model which provides a functional form for the hazard.

For babies born at the earliest weeks of gestational age the predicted median length of stay will be much lower than the actual length of stay for surviving babies. An alternative approach to allow consideration of the outcomes of death and discharge from neonatal care is to model these as two outcomes with different event time distributions. These would be considered 'competing' events (76, 95, 96), as the occurrence of one prevents the other from ever occurring. The most appropriate methodology to analyse these data is an extension of survival analysis known as competing risks analysis and this will be introduced and investigated in Chapter 5.

For commissioning purposes, it may not be important to separate the two groups of babies as irrespective of their outcome, care was required. However, estimates of

length of stay, as provided here are not informative for commissioning of neonatal specialist services as they provide no detail about the specific types of care required. To include more detailed information on the length of stay for commissioning purposes, such as the care required by a baby, a further extension of this analysis known as multistate modelling is required and this will be introduced in Chapter 6.

4.10 CHAPTER CONCLUSION

This chapter has introduced the concept of survival analysis and provided an introduction of the methods of the Cox proportional hazards model and flexible parametric modelling. A preliminary analysis was undertaken using the Cox model, the basis of much of the advanced work in this thesis. An equivalent analysis was provided using a flexible parametric model. The probability of remaining in the neonatal unit and the median length of stay were estimated.

The combination of outcomes here means that it is difficult to provide a useful interpretation of the hazard ratios. The methods here are also not appropriate for understanding two different outcomes simultaneously, such as being able to estimate length of stay for babies who survive and those who die. In an area of medicine with high levels of mortality it is important to consider deaths and discharges from the neonatal unit as separate outcomes, both of which are important in their own right.

5 COMPETING RISKS ANALYSIS FOR THE ESTIMATION OF MORTALITY AND LENGTH OF STAY

5.1 CHAPTER OVERVIEW

Survival analysis, introduced in the preceding chapter, is used to model the time until an event occurs. However, time to one of multiple events may be of interest. For example, there may be interest in the specific cause of death, where there are multiple causes from which a person can potentially die. These endpoints need not be mortality, but as they are considered simultaneously, the occurrence of one must prevent the others from occurring. These events *compete* with each other to occur first, and the statistical approach to model this is known as a competing risks analysis.

In this chapter, competing risks methodology is introduced as an alternative approach for investigating length of stay, which considers death or discharge from neonatal care as competing events. The probability of survival is considered along with estimates of length of stay for babies of different gestational ages. The results are then adjusted for babies of differing characteristics.

5.2 INTRODUCTION

The history of **competing risks** analyses extends back to the 18th century when Bernoulli studied the consequences of eradicating smallpox on mortality rates (97, 98). More recently, competing risks methods gained popularity in the 1970s (95, 99), and recent advances in statistical software have made implementation less computationally intensive (76, 100, 101).

Competing risks methods have often been used when there is interest in different causes of death, for example: death from respiratory distress, infection or prematurity. The occurrence of one of these events prevents the other events from occurring and can be considered in a single analysis known as a competing risks analysis.

In this thesis two competing events are considered: death or discharge from neonatal care (Figure 5-1). All babies experience an event in this study and when a baby experiences one outcome (e.g. death in the neonatal unit) they are censored from receiving the other outcome (e.g. discharge from the neonatal unit).

The final endpoints in a competing risks analysis are known as **absorbing states** because upon entry it is not possible to exit them again. The movement into an absorbing state is known as a transition. A standard survival analysis can be considered as a competing risks model with one initial state and one final absorbing state. Similarly, the approach to competing risks analysis is a special case of multistate modelling, to be introduced in Chapter 6. Although only two endpoints are considered here, a competing risks analysis can be used for more outcomes.

5.2.1 DATA SET-UP FOR COMPETING RISKS

Most medical research data are collected and stored in wide format, i.e. each study participant has one row of data as seen in Table 5-1. In this example, the first baby dies after five days whilst the second survives to discharge from the neonatal unit after 37 days. To use competing risks analysis the data need to be manipulated into long format, with a row of data for each possible event as in Table 5-2. The status variable takes the value of one if the event occurred and zero otherwise. Use of data in this format allows implementation of a competing risk analysis to consider multiple events simultaneously.

Table 5-1: Details for two hypothetical babies in wide format.

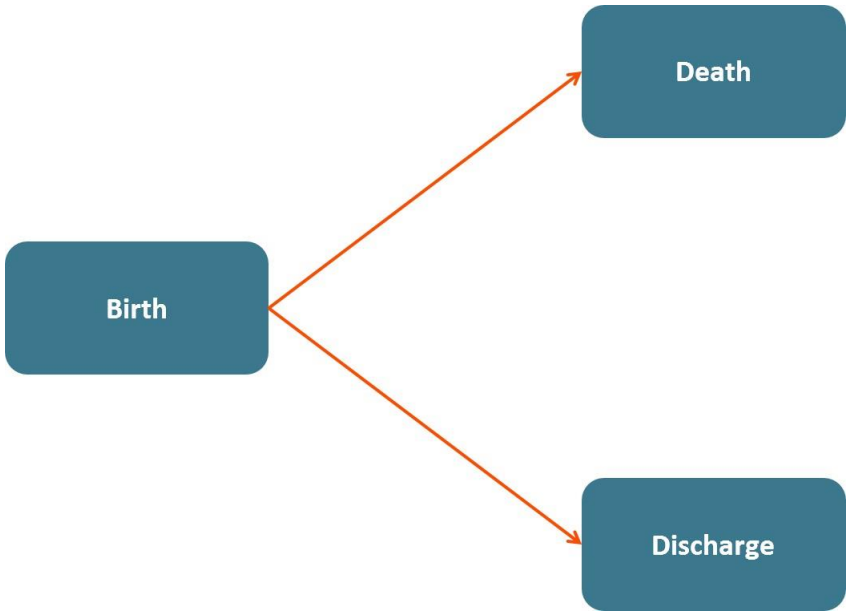
Baby ID	Gestational age	Time	Event
1	24	5	Died
2	30	37	Discharged home

Table 5-2: Details for the same two hypothetical babies presented in long format.

Baby ID	Gestational age	Time	Event	Status
1	24	5	Died	1
1	24	5	Discharged home	0
2	30	37	Died	0
2	30	37	Discharged home	1

A statistical model, such as the Cox proportional hazards model, can be fitted to these two events simultaneously by stratifying the analysis on the event. Different covariates can be considered for the different events (**event-specific covariates**) or the covariate effect can be shared across the events (**shared covariates**).

Figure 5-1: An example of a competing risks model with two outcomes: death or discharge.



5.3 CAUSE-SPECIFIC HAZARD AND CUMULATIVE INCIDENCE FUNCTION

There are two main statistics of interest in a competing risks analysis: the cause-specific hazard and the cumulative incidence function. Whilst in the literature the term used is cause-specific hazard here this will be renamed as the **event-specific hazard** as the term cause-specific can create confusion when the outcomes are not causes of

death. The **cumulative incidence function** provides the proportion of individuals that have experienced a specific event over the follow up-time.

5.3.1 EVENT-SPECIFIC HAZARD

If an individual is at risk of experiencing D potential events, then the event-specific hazard $\alpha_k(t)$ is the rate of failure from k at time t, given that the patient has not already experienced an event, where T is the time of failure from any event. The event-specific hazard function (95) is denoted:

$$\alpha_k(t) = \lim_{\delta \to 0} \frac{P(t \leq T \leq t + \delta, D = k | T \geq t)}{\delta}$$

Equation 5-1

The event-specific hazard is calculated conditionally, i.e. for the patient to experience one event they cannot have experienced any other event and are censored from experiencing them. There is an event-specific hazard for every event and from all D of these it is possible to estimate the cumulative incidence function.

The event-specific cumulative hazard is obtained by integrating the overall hazard as:

$$A_k(t) = \int_0^t \alpha_k(s) ds$$

Equation 5-2

5.3.2 SURVIVAL FUNCTION AND CUMULATIVE INCIDENCE FUNCTION

From the event-specific hazard functions it is possible to obtain a survival function:

$$S_k(t) = \exp(-A_k(t))$$

Equation 5-3

which is the survival from event k, the marginal survival function, and the complement can be interpreted as the probability of experiencing event k in a hypothetical world where it is not possible to experience any other event. However, this assumption is unlikely to hold and therefore a more appropriate overall survival function is:

$$S(t) = \exp\left(-\sum_{k=1}^{D} A_k(t)\right)$$

Equation 5-4

which is the probability of not having experienced any event at time t. This is used to define the cumulative incidence function of event k, defined as $\mathrm{I}_k(\mathrm{t})$ which is the probability of experiencing event k before time t:

$$\mathrm{I}_k(\mathrm{t}) = \int_0^{\mathrm{t}} \alpha_k(\mathrm{s})\mathrm{S}(\mathrm{s})\mathrm{ds}$$

Equation 5-5

or with covariates:

$$\mathrm{I}_k(\mathrm{t}|\mathbf{Z}_k) = \int_0^{\mathrm{t}} \alpha_k(\mathrm{s}|\mathbf{Z}_k)\mathrm{S}(\mathrm{s}|\mathbf{Z})\mathrm{ds}$$

Equation 5-6

There are two main approaches to modelling competing risks: 1) model the event-specific hazards and transform to the cumulative incidence function and 2) model the cumulative incidence function directly (102). The first approach is often advocated as both the event-specific hazard and the cumulative incidence function are useful for communication of risk and provide an absolute measure that prognosis and clinical

decision making can use (103, 104). Therefore, this is the approach used throughout the competing risks analysis in this thesis.

5.4 COX PROPORTIONAL HAZARDS MODELLING

In a similar approach to that of a standard survival analysis, the concepts of the hazard and survival functions can be extended to a competing risks analysis framework using Cox proportional hazards modelling. To fit the model, the data are arranged in the format seen in Table 5-2.

5.4.1 EVENT-SPECIFIC HAZARD RATE

As was seen for the Cox model in Chapter 4 it is possible to estimate the event-specific hazard rate, here written on the log scale, for event k with covariates $\boldsymbol{Z}$ and regression coefficients $\boldsymbol{\beta}$. The model can be fitted by stratifying for all events simultaneously:

$$\ln\big(\alpha_k(t|\boldsymbol{Z}_k)\big) = \ln\Big(\alpha_{k,0}(t)\Big) + \boldsymbol{\beta}_k'\boldsymbol{Z}_k + \boldsymbol{\beta}'\boldsymbol{Z}$$

Equation 5-7

where $\boldsymbol{\beta}_k'\boldsymbol{Z}_k$ are interaction terms between event k and the covariates, therefore allowing the covariates to differ between events. These are known as **event-specific covariates**, and different covariates can be considered for different events if appropriate. The term $\boldsymbol{\beta}'\boldsymbol{Z}$ refers to covariates where the values are shared across all the events and these are referred to as **shared covariates**. The baseline event-specific hazard for event k is defined as: $\alpha_{k,0}(t)$. When someone experiences an event they are censored from experiencing the other events at the time the transition occurs (76).

5.4.2 CUMULATIVE INCIDENCE FUNCTION

In Equation 5-5, the event-specific cumulative incidence function were derived from the event-specific hazards via the cumulative hazard (101). The cumulative incidence function can be written in terms of the cause-specific cumulative hazard:

$$I_k(t|\mathbf{Z}) = A_k(t|\boldsymbol{Z})\exp\left(\sum_{k=1}^{D} A_k(t|\boldsymbol{Z})\right)$$

Equation 5-8

where

$$A_k(t|\boldsymbol{Z}) = A_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{Z})$$

Equation 5-9

and the event-specific cumulative baseline hazard is estimated using the Breslow estimator (101):

$$\hat{A}_0(t) = \sum_{j:t_j \le t} \frac{1}{\sum_{l \in R_j} \exp(\widehat{\boldsymbol{\beta}}'\boldsymbol{Z}_l)}$$

Equation 5-10

where $R_j$is the risk set at time t_j and l represents each individual at risk (82).

5.4.3 ASSUMPTIONS OF THE COX MODEL

As with the standard survival analysis, one assumption of the Cox competing risks model is that the hazards are proportional over time. This assumption is extended so that for each event, the effects of the event-specific covariates are modelled to be proportional over time. Stratification of the variables which are non-proportional is one approach to overcome some of the limitations, as seen in Chapter 4. An alternative method such as the use of flexible parametric modelling may provide advantages if relaxation of the proportional hazards assumption is required and this will be discussed in Chapter 5.7.

5.5 COX PROPORTIONAL HAZARDS COMPETING RISKS ANALYSIS

I undertook an analysis using the Cox competing risks model to estimate the length of stay and survival for babies born very preterm. The 21,038 babies introduced and summarised in Chapter 3 and modelled in Chapter 4 were used. All data manipulation was undertaken in Stata v 14 and I fitted a Cox proportional hazards competing risks model using the *mstate* command in R 3.0.2 which was stratified for the events of interest: death or discharge from the neonatal unit. Gestational age was modelled as an event-specific covariate. Event-specific hazard ratios were estimated and are compared to the baseline group selected as 27 weeks of gestational age at birth (Table 5-3). Equivalent results to those estimated by *mstate* were obtained using the *stpm2* command in Stata v 14 (results not shown).

The probability of death at any time, i.e. the hazard of death, for babies born at 24 weeks was 3.44 times higher than in the babies born at 27 weeks gestational age (95% CI: 2.93 to 4.04, p<0.001, Table 5-3). For babies born at 31 weeks, the hazard of being discharged home alive was 8.79 times higher than babies born at 27 weeks gestational age at all time points (95% CI: 8.31 to 9.29, p<0.001, Table 5-3).

Table 5-3: Event-specific hazard ratios and 95% confidence intervals for death or discharge, by week of gestational age (Cox model).

Gestational age (weeks)	Hazard ratio	95% Confidence Interval	p-value
Died			
24	3.44	2.93, 4.04	<0.001
25	1.80	1.51, 2.14	<0.001
26	1.34	1.13, 1.59	<0.001
27	Baseline	Baseline	Baseline
28	0.65	0.54, 0.78	<0.001
29	0.34	0.28, 0.43	<0.001
30	0.27	0.22, 0.34	<0.001
31	0.23	0.22, 0.34	<0.001
Discharged			
24	0.39	0.36, 0.42	<0.001
25	0.52	0.48, 0.56	<0.001
26	0.73	0.68, 0.78	<0.001
27	Baseline	Baseline	Baseline
28	1.44	1.36, 1.53	<0.001
29	2.42	2.28, 2.57	<0.001
30	4.35	4.10, 4.60	<0.001
31	8.79	8.31, 9.29	<0.001

From the event-specific hazards the cumulative incidence functions were estimated for each week of gestational age using Equation 5-8. These are interpreted as the predicted proportion of babies that have died or been discharged, or the probability of an event occurring, from the neonatal unit over time. These estimates are often considered more clinically useful as it is possible to visually inspect the proportion of babies having experienced an event over time.

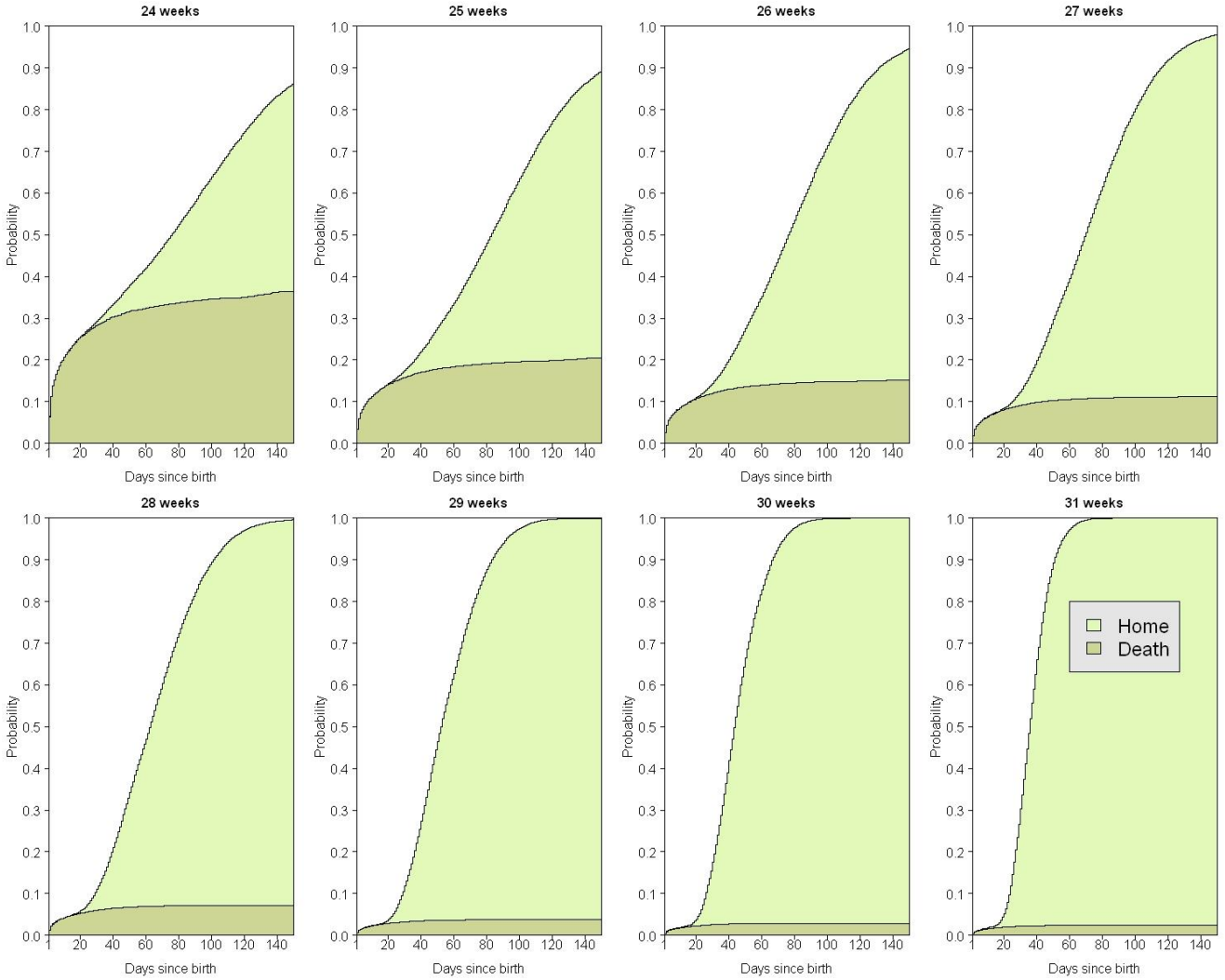
The stacked cumulative incidence functions are presented graphically for each week of gestational age in Figure 5-2. The different shaded regions represent the proportions of babies who have died or been discharged from neonatal care over time. The unshaded area represents the babies who remain in the neonatal unit, having not yet experienced either event. For most weeks of gestational age no unshaded area remains at 150 days, i.e. most babies have been discharged or died.

The proportion of babies who were predicted to die was highest for those born most preterm. The area representing death plateaued at the total proportion of deaths for each week of gestational age and the proportion of babies that died corresponded with other published evidence for this population (14). Discharge from the neonatal unit occurred earlier for babies born at the later weeks of gestational age. From Figure 5-2 the proportion of babies who died or were discharged home by certain points in time can be identified. Selected time points are provided in Table 5-4: one day, ten days and 30 days after birth.

Table 5-4: Proportion of babies who have died or been discharged home one day, ten days and 30 days after birth (Cox model).

Gestational age (weeks)	**1 day**		**10 days**		**30 days**	
	Died	**Discharged**	**Died**	**Discharged**	**Died**	**Discharged**
24	0.064	0.000	0.206	0.000	0.283	0.010
25	0.034	0.000	0.112	0.000	0.158	0.017
26	0.025	0.000	0.085	0.000	0.120	0.024
27	0.019	0.000	0.064	0.000	0.091	0.034
28	0.012	0.000	0.042	0.000	0.060	0.050
29	0.006	0.000	0.022	0.000	0.032	0.085
30	0.005	0.000	0.018	0.001	0.025	0.149
31	0.004	0.000	0.015	0.003	0.022	0.281

Figure 5-2: Stacked cumulative incidence functions over time to estimate the proportion of deaths or discharges by week of gestational age (Cox model).



5.5.1 TESTING MODEL ASSUMPTION: PROPORTIONAL HAZARDS

The Therneau-Grambsch test was undertaken to investigate the assumption of proportional hazards. Significant results, indicating the assumption of proportional hazards does not hold, were seen for all weeks of gestational age who were discharged except 28 weeks (Table 5-5). This issue was seen in Figure 5-2 where discharge from neonatal care begins to occur at the same point in time for all weeks of gestational age, although at different rates. For babies born at the earliest gestational ages, discharge from neonatal care is unlikely to occur at this early time point. There are no issues of proportional hazards for the hazard of death except for babies born at 24 weeks. Further investigation of this will be undertaken in Chapter 5.7.

Table 5-5: Therneau-Grambsch test for proportional hazards (Cox model).

Gestational age (weeks)	**p-value**
Discharged	
24	<0.001
25	<0.001
26	<0.001
27	Baseline
28	0.275
29	<0.001
30	<0.001
31	<0.001
Died	
24	<0.001
25	0.851
26	0.035
27	Baseline
28	0.287
29	0.196
30	0.845
31	0.057

5.6 FLEXIBLE PARAMETRIC MODELLING

The Cox model can be extended to a flexible parametric setting and this aids the relaxation of the proportional hazards assumption. The flexible parametric approach was extended to include competing risks methods in 2013 by Hinchliffe and Lambert (101, 103), and is used here. An example of the potential use of this approach for

modelling length of stay in neonatal care was published in Paediatric and Perinatal Epidemiology by Hinchliffe, Seaton et al (6).

The use of a flexible parametric approach to competing risks analysis provides equivalent results to those from the Cox model when proportional hazards are assumed. The advantage of this approach is that time-dependent effects can be incorporated and this will be explored in Chapter 5.6.2.

5.6.1 EXTENDING TO A FLEXIBLE PARAMETRIC MODEL

The Cox model approach required the dataset to be in long format and then the model was stratified by the event. A flexible parametric proportional hazards model can be fitted in a similar way using the expanded dataset:

$$\ln\big(A_k(t|\boldsymbol{Z})\big) = s\big\{\ln(t)\,\big|\boldsymbol{\gamma}_{0,k}, \boldsymbol{n}_{0,k}\big\} + \boldsymbol{\beta}_k'\boldsymbol{Z}_k + \boldsymbol{\beta}'\boldsymbol{Z}$$

Equation 5-11

where $s\big\{\ln(t)\,\big|\boldsymbol{\gamma}_{0,k}, \boldsymbol{n}_{0,k}\big\}$ is the log cumulative hazard function for the event k modelled using restricted cubic splines with knots $\boldsymbol{n}_{0,k}$. Using the event-specific hazards, the cumulative incidence function can be obtained using Equation 5-5 via numerical integration (103).

It is possible to obtain equivalent results from the Cox proportional hazards model and the flexible parametric proportional hazards model within the competing risks setting, as seen in Chapter 4 for the standard survival analysis.

5.6.2 INCORPORATION OF TIME-DEPENDENT EFFECTS

All analyses undertaken so far have relied on the proportional hazards assumption (see Chapter 4) and as seen in Table 5-5 this is not always appropriate. To relax the proportional hazards assumption, time-dependent effects are created by introducing interactions between the spline terms for the hazard and the parameters of interest (79). Equation 5-11 can be extended to incorporate time-dependent covariates:

$$\ln\big(A_k(t|\boldsymbol{Z})\big) = s\big\{\ln(t)\,\big|\boldsymbol{\gamma}_{0,k}, \boldsymbol{n}_{0,k}\big\} + \sum_{j=1}^{TD_k} s\big\{\ln(t)\,\big|\boldsymbol{\gamma}_{0,k}, \boldsymbol{n}_{0,k}\big\}\, z_j + \boldsymbol{\beta}_k'\boldsymbol{Z}_k + \boldsymbol{\beta}'\boldsymbol{Z}$$

Equation 5-12

where TD_k is the number of time-dependent effects and $s\big\{\ln(t)\,\big|\boldsymbol{\gamma}_{0,k}, \boldsymbol{n}_{0,k}\big\}z_j$ is the spline function for the j^{th} time-dependent effect for event k (101). This allows for the introduction of non-proportional effects and the difference between covariates does not have to be constant over time. Although the model estimates can no longer be expressed as a single hazard ratio estimates of the cumulative incidence function are still informative and may be more biologically plausible than those obtained from assuming proportional hazards.

5.7 FLEXIBLE PARAMETRIC MODEL WITH TIME-DEPENDENT COVARIATE

I fitted a flexible parametric competing risks model using the *stpm2* command in Stata v 14, including a categorical term for gestational age. The *mstate* command does not support the inclusion of time-dependent covariates. The incorporation of a time-dependent effect for gestational age lowered the Akaike Information Criterion (AIC) from 29974 to 23929 indicating that use of time-dependent effects improved the model fit (Table 5-6) (105). The AIC measures goodness of fit (assessed by the likelihood function) which is penalised for an increasing number of parameters. A lower AIC indicates a better model fit (105).

To select the number of knots required to provide the best model fit, I fitted models with different numbers of knots in the baseline hazard and selected the model with the lowest AIC. The final model had four knots giving an AIC of 23811 (Table 5-6).

Table 5-6: Comparison of the number of knots to model gestational age as a time-dependent effect and corresponding AIC.

Approach for modelling gestational age	**AIC**
Gestational age (not time-dependent). Three knots.	29974
Gestational age (time-dependent). Three knots.	23929
Gestational age (time-dependent). Four knots.	**23811**
Gestational age (time-dependent). Five knots.	23826
Gestational age (time-dependent). Six knots.	24049 (model failed to converge)

5.7.1 EVENT-SPECIFIC HAZARDS

The inclusion of a time-dependent effect for gestational age prevents interpretation of hazard ratios that can describe whether one group is consistently at an increased (or decreased) hazard than another group at every time point. However, it is possible to plot the event-specific hazard rates to see how these vary over time and these are provided in Figure 5-3 and Figure 5-4.

Between the different gestational ages, the relaxation of proportional hazards allows the difference between the hazards to vary over time. The hazard of death for babies born at 24 weeks gestational age decreases over time (Figure 5-3). For babies born at 25 to 28 weeks gestational age the hazard of death reduces until around 30 days when it increases again briefly, then reduces before increasing after day 60 (Figure 5-3). Although in this analysis babies born at 24 weeks may have a different shaped hazard, this is potentially due to the small number of babies in this group and so should not be over-interpreted. These babies may be more similar to the group of babies born at 25 weeks gestational age than seen here. For babies born at 29 to 31 weeks the hazard of

death increases after 30 days. This is likely to reflect the level of sickness of those babies in this group who remain in hospital at this point in time.

The differences between babies born at 24 to 28 weeks gestational age and those born at 29 to 31 weeks were apparent when investigating the hazard of discharge (Figure 5-4). For babies born at 24 to 28 weeks the hazard of discharge increases over time. However, for babies born at 30 and 31 weeks gestational age the hazard of discharge increases and then decreases after approximately 50 days. This corresponds with the increase in the hazard of mortality for this group of babies. This may reflect a higher level of sickness in this group of babies compared to their peers who have been discharged. However, towards the end of follow up time, data are limited and so these results may also be related to lack of data.

Figure 5-3: Predicted event-specific hazards of death over time by week of gestational age after allowing for time-dependent effects.

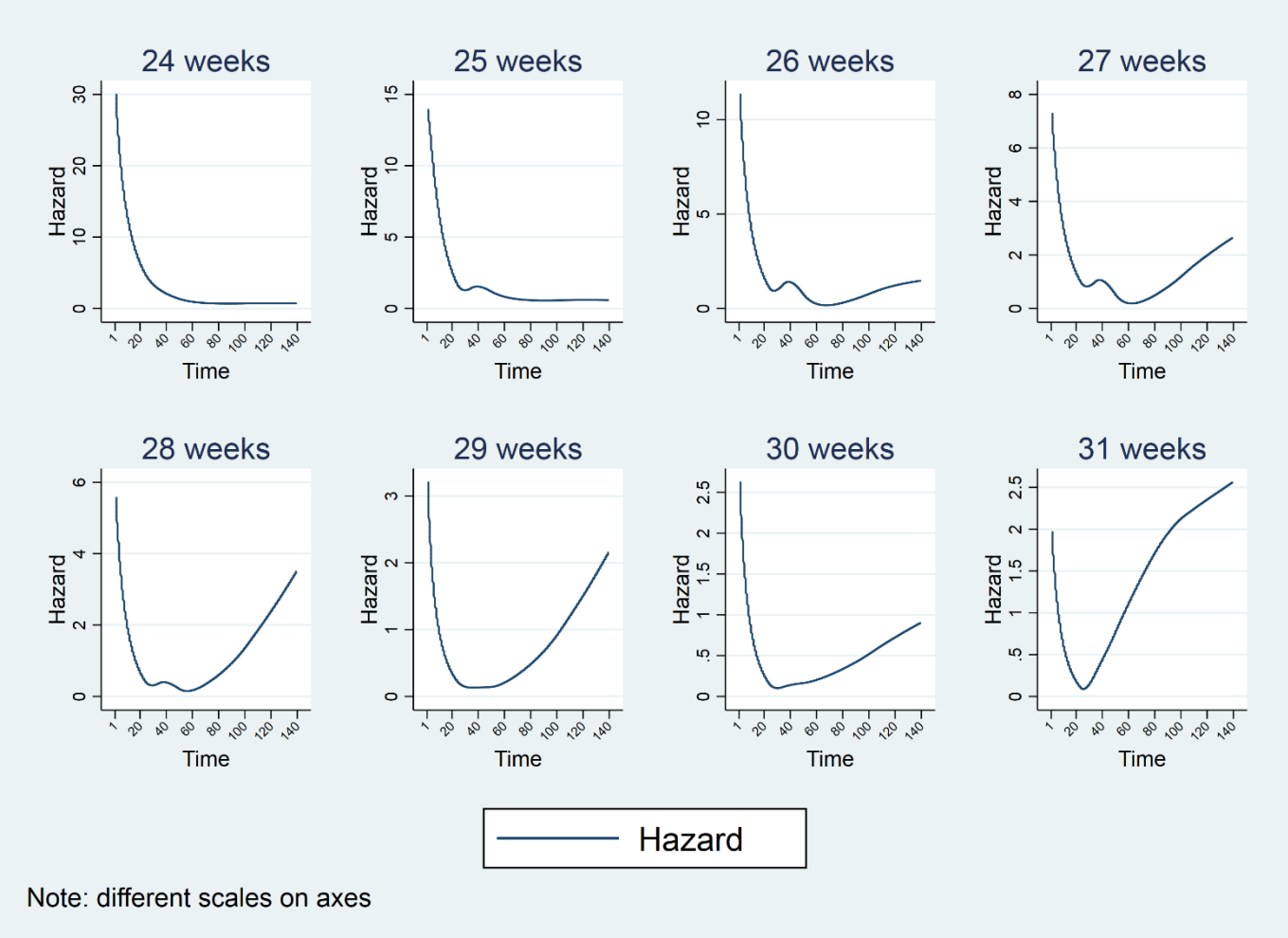
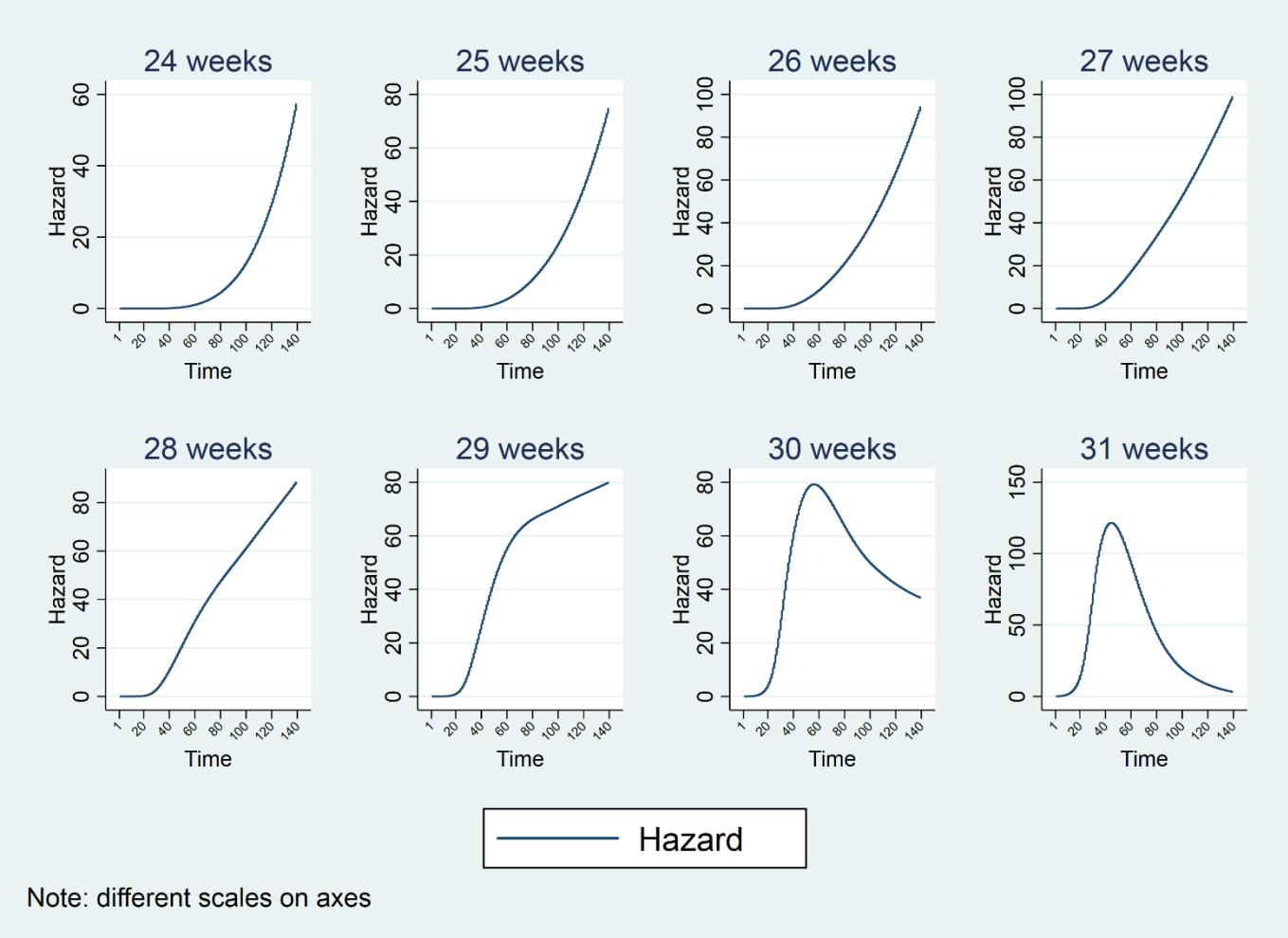


Figure 5-4: Predicted event-specific hazards of discharge by week of gestational age after allowing for time-dependent effects.



5.7.2 CUMULATIVE INCIDENCE FUNCTIONS

The cumulative incidence functions are estimated using the *stpm2cif* post-estimation command in Stata v 14 and presented as stacked plots in Figure 5-5. The proportion of babies estimated to have died or been discharged are provided for one, ten and 30 days after birth (Table 5-7). As expected, as gestational age increased, the proportion of babies estimated to die reduced, and the proportion discharged from neonatal care increased. No babies born at 24 and 25 weeks gestational age were discharged in the first 30 days of life.

Table 5-7: Proportion of babies who have died or been discharged home one, ten and 30 days after birth estimated from the flexible parametric model with time-dependent covariate.

Gestational age (weeks)	**1 day**		**10 days**		**30 days**	
	Died	**Discharged**	**Died**	**Discharged**	**Died**	**Discharged**
24	0.059	0.000	0.213	0.000	0.313	0.000
25	0.024	0.000	0.108	0.000	0.159	0.000
26	0.023	0.000	0.084	0.000	0.117	0.001
27	0.014	0.000	0.056	0.000	0.084	0.004
28	0.011	0.000	0.041	0.000	0.057	0.013
29	0.008	0.000	0.023	0.000	0.031	0.042
30	0.006	0.000	0.019	0.001	0.025	0.141
31	0.004	0.000	0.014	0.005	0.019	0.357

5.7.3 UNCERTAINTY ESTIMATES FOR THE CUMULATIVE INCIDENCE FUNCTION

Confidence intervals for the estimate of the cumulative incidence functions are estimated using the delta method (103). Estimates of the cumulative incidence functions for death and discharge, with 95% confidence intervals are provided for babies born at 24 weeks (Figure 5-6) and those born at 31 weeks (Figure 5-7). Confidence interval estimates for other weeks of gestational age can be found in Appendix 5. The proportion of babies who have died or been discharged at selected time points is reproduced in Table 5-8 with the 95% confidence intervals.

Figure 5-5: Predicted cumulative incidence function stacked plots from the flexible parametric model with a time-dependent effect for gestational age.

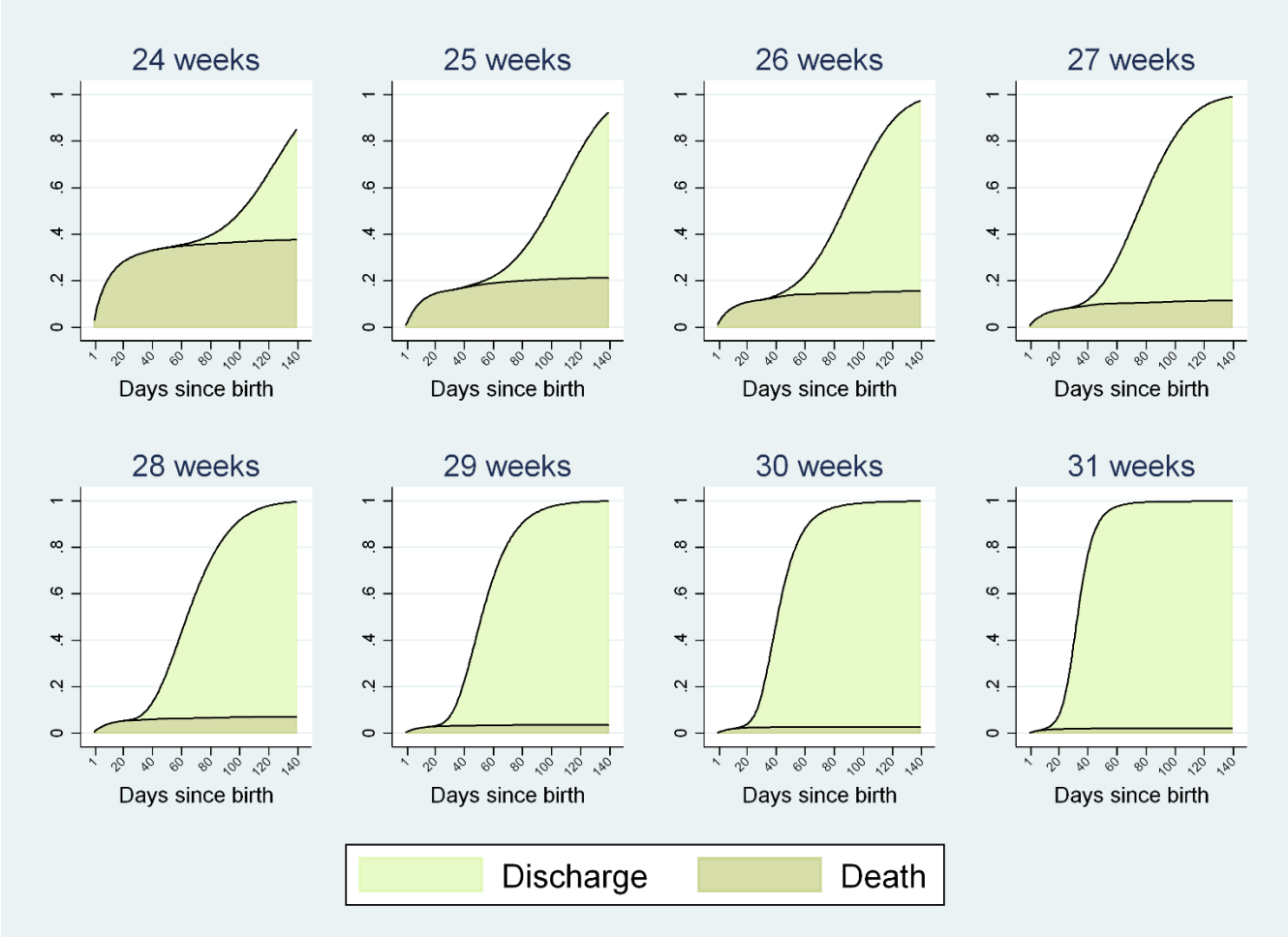


Table 5-8: Estimated proportion of babies, with 95% confidence interval, who have died or been discharged from neonatal care at one, ten and 30 days. Estimated from the flexible parametric model with time-dependent effect for gestational age.

Gestational age (weeks)	**1 day**		**10 days**		**30 days**	
	Died	**Discharged**	**Died**	**Discharged**	**Died**	**Discharged**
24	0.059 (0.049, 0.069)	0.0000 (0.000, 0.000)	0.213 (0.191, 0.236)	0.000 (0.000, 0.000)	0.313 (0.286, 0.339)	0.000 (0.000, 0.000)
25	0.024 (0.017, 0.030)	0.000 (0.000, 0.000)	0.108 (0.092, 0.123)	0.000 (0.000, 0.000)	0.159 (0.139, 0.178)	0.000 (0.000, 0.000)
26	0.023 (0.018, 0.028)	0.000 (0.000, 0.000)	0.084 (0.072, 0.095)	0.000 (0.000, 0.000)	0.117 (0.102, 0.132)	0.001 (0.000, 0.001)
27	0.014 (0.010, 0.018)	0.000 (0.000, 0.000)	0.056 (0.047, 0.065)	0.000 (0.000, 0.000)	0.084 (0.072, 0.096)	0.004 (0.003, 0.004)
28	0.011 (0.008, 0.014)	0.000 (0.000, 0.000)	0.041 (0.035, 0.048)	0.000 (0.000, 0.000)	0.057 (0.048, 0.065)	0.013 (0.011, 0.014)
29	0.008 (0.005, 0.010)	0.000 (0.000, 0.000)	0.023 (0.019, 0.028)	0.000 (0.000, 0.000)	0.031 (0.025, 0.037)	0.042 (0.039, 0.045)
30	0.006 (0.004, 0.008)	0.000 (0.000, 0.000)	0.019 (0.015, 0.023)	0.001 (0.001, 0.001)	0.025 (0.020, 0.030)	0.141 (0.133, 0.148)
31	0.004 (0.003, 0.005)	0.000 (0.000, 0.000)	0.014 (0.011, 0.017)	0.005 (0.004, 0.006)	0.019 (0.015, 0.022)	0.357 (0.346, 0.368)

Figure 5-6: Cumulative incidence functions with 95% confidence intervals for death or discharge for babies born at 24 weeks gestational age (note: different scales).

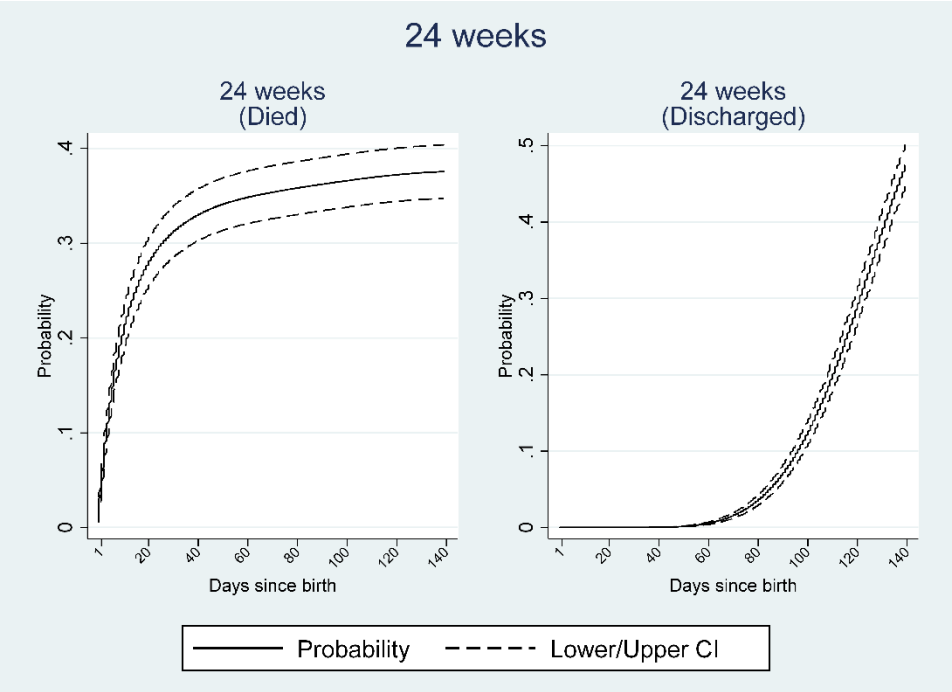
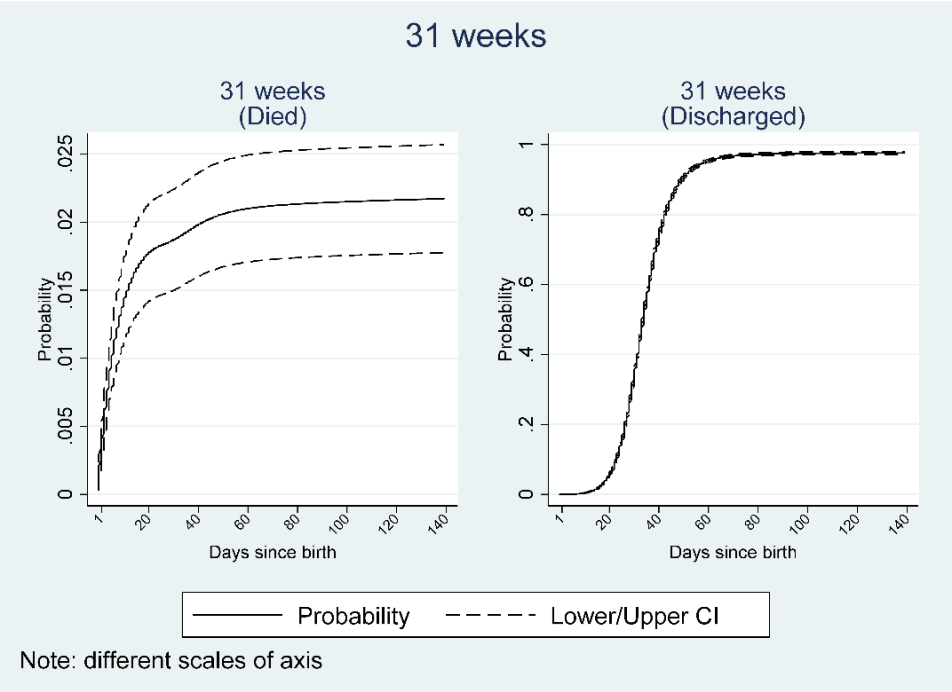


Figure 5-7: Cumulative incidence functions with 95% confidence intervals of death or discharge for babies born at 31 weeks gestational age (note: different scales).



5.7.4 ESTIMATING LENGTH OF STAY

Median length of stay was estimated by outcome (Table 5-9). Results should be interpreted with care when the events are rare, for example the 10th and 90th centiles for length of stay for babies who died who were born at 30 and 31 weeks could be heavily influenced by one event. The time to the estimated date of delivery (commonly known as the EDD) is provided in Table 5-9 assuming that all births are measured in completed weeks and that term delivery is taken as 40 weeks gestational age. Babies born at 24 to 26 weeks gestational age are discharged from neonatal care around their due date. Babies born at 27 to 29 weeks are discharged a few weeks before their due date. Babies born at 30 to 31 weeks seem to be discharged around a month before their due date, i.e. at the point they are around 35 to 36 weeks corrected age.

As around half of deaths appear to occur in the first seven to ten days, it would appear clinically appropriate to use this as an estimate of the length of stay for babies who die. At around ten days after birth it may be appropriate to consider if a baby is clinically stable and likely to survive to discharge, what their potential length of stay may be.

Table 5-9: Estimated median length of stay (10th, 90th centile) for babies who survive or who die from the flexible parametric model with time-dependent effects.

Gestational age (weeks)	Length of stay (days) of discharges	Length of stay (days) of deaths	Time to estimated due date
24	122 (86, 152)	8 (1, 49)	112
25	107 (73, 139)	9 (1, 62)	105
26	91 (60, 124)	9 (1, 66)	98
27	78 (51, 110)	8 (1, 61)	91
28	66 (42, 99)	10 (1, 49)	84
29	53 (35, 79)	8 (1, 131)	77
30	42 (28, 62)	3 (1, 106)	70
31	33 (23, 47)	7 (1, 97)	63

5.7.5 SENSITIVITY ANALYSIS

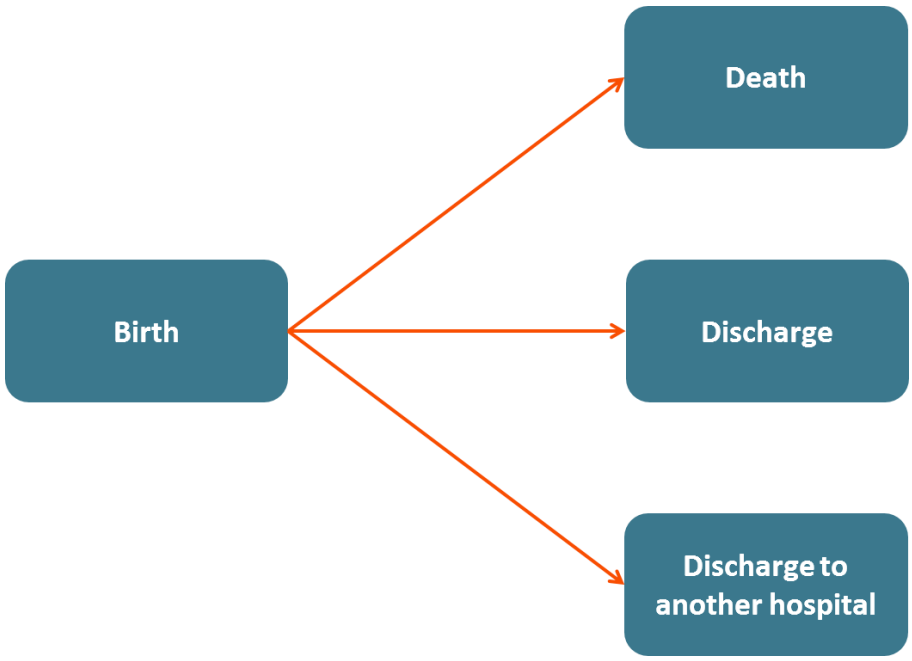
The outcome of discharge from neonatal care was an amalgamation of several outcomes including: discharge home; discharge to another hospital and discharge for

surgery (Table 3-6). For the babies discharged to another hospital it is possible that they may have received ongoing neonatal care in a hospital outside the area covered by the NNRD and therefore their true stay in neonatal care may be longer than that observed. This could have been considered as an additional outcome in the competing risks analysis (Figure 5-8) although the number of babies experiencing this was small. Therefore, as a sensitivity analysis, the observations for the babies discharged to another hospital were censored (81) (n=219). The hazard ratios were not reported previously in this analysis as they have no immediate interpretation as use of a time-dependent covariate for gestational age means they will vary over time. However, the hazard ratios are provided to aid comparison of the standard analysis and the sensitivity analysis to investigate model fit (Table 5-10). The sensitivity analysis resulted in small changes to the hazard ratios for babies discharged at 30 and 31 weeks gestational age (Table 5-10) but other results remained consistent.

Table 5-10: Hazard ratios (95% confidence intervals) from the non-censored analysis and the sensitivity analysis investigating babies discharged to another hospital.

Gestational age (weeks)	**Hazard ratio (95% confidence interval)**	**Hazard ratio (95% confidence interval) censored analysis**
Died	0.10 (0.09, 0.11)	0.10 (0.09, 0.11)
24	3.92 (3.33, 4.62)	3.92 (3.33, 4.62)
25	1.86 (1.55, 2.23)	1.86 (1.55, 2.23)
26	1.39 (1.16, 1.66)	1.39 (1.16, 1.66)
27	Baseline	Baseline
28	0.67 (0.55, 0.81)	0.67 (0.55, 0.81)
29	0.38 (0.30, 0.48)	0.38 (0.30, 0.48)
30	0.30 (0.24, 0.38)	0.30 (0.24, 0.38)
31	0.35 (0.27, 0.45)	0.35 (0.27, 0.45)
Discharged	0.03 (0.02, 0.04)	0.03 (0.02, 0.03)
24	0.02 (0.01, 0.03)	0.02 (0.01, 0.03)
25	0.10 (0.08, 0.13)	0.10 (0.08, 0.13)
26	0.35 (0.30, 0.42)	0.37 (0.31, 0.44)
27	Baseline	Baseline
28	2.65 (2.33, 3.01)	2.73 (2.41, 3.11)
29	6.79 (6.04, 7.64)	6.93 (6.16, 7.81)
30	17.02 (15.21, 19.06)	17.56 (15.67, 19.68)
31	37.04 (33.15, 41.39)	38.15 (34.11, 42.68)

Figure 5-8: An alternative competing risks model with three outcomes: death, discharge or discharge to another hospital.



An additional sensitivity analysis was undertaken for the 141 babies who were discharged to receive surgery (Table 3-4) as there is no information about their later care. These babies may have been discharged home; discharged to paediatric services or died following surgery. The robustness of the analysis was investigated in a sensitivity analysis by assuming these babies died and re-estimating the hazard ratios (Table 5-11). Again, this did not substantially alter the results, with only small changes seen in the estimated hazard ratios for discharge for babies born at 30 and 31 weeks.

This sensitivity analysis indicates that the results presented here are robust to the assumption that all these babies can be considered to have been discharged from neonatal care without this having an impact on the overall estimates.

Table 5-11: Hazard ratios (95% confidence intervals) from the non-censored and the sensitivity analysis assuming all babies discharged to surgery died.

Gestational age (weeks)	**Hazard ratio (95% confidence interval)**	**Hazard ratio (95% confidence interval) sensitivity analysis**
Died	0.10 (0.09, 0.11)	0.10 (0.09, 0.11)
24	3.92 (3.33, 4.62)	3.82 (3.25, 4.49)
25	1.86 (1.55, 2.23)	1.83 (1.53, 2.20)
26	1.39 (1.16, 1.66)	1.37 (1.15, 1.64)
27	Baseline	Baseline
28	0.67 (0.55, 0.81)	0.69 (0.57, 0.83)
29	0.38 (0.30, 0.48)	0.42 (0.34, 0.52)
30	0.30 (0.24, 0.38)	0.41 (0.33, 0.51)
31	0.35 (0.27, 0.45)	0.43 (0.34, 0.53)
Discharged	0.03 (0.02, 0.04)	0.03 (0.02, 0.03)
24	0.02 (0.01, 0.03)	0.02 (0.01, 0.03)
25	0.10 (0.08, 0.13)	0.10 (0.08, 0.13)
26	0.35 (0.30, 0.42)	0.35 (0.30, 0.42)
27	Baseline	Baseline
28	2.65 (2.33, 3.01)	2.64 (2.33, 3.00)
29	6.79 (6.04, 7.64)	6.80 (6.05, 7.66)
30	17.02 (15.21, 19.06)	17.07 (15.25, 19.12)
31	37.04 (33.15, 41.39)	37.21 (33.28, 41.58)

5.7.6 MODEL FIT AND PERFORMANCE

An acknowledged limitation of competing risks methodology is the difficulty in assessing model fit and future work is needed to develop approaches to improve this (6). However, research indicates that these methods are robust to selection of the number and location of knots (80, 89).

To assess the model fit when using the time-dependent covariate for gestational age, comparisons were made between the observed median length of stay and that estimated from the model for babies surviving to discharge (Table 5-12) and those who died in neonatal care (Table 5-13). Median length of stay of babies who survived was estimated well, with the maximum difference between estimated and observed being three days for the most preterm babies. For babies born at 27 to 31 weeks the

estimated median length of stay was within one day of the observed median length of stay.

The estimated length of stay for babies who died varied more than for those who survived, with the largest difference being an under-estimation of 5 days for babies born at 29 weeks. This result was less precise as there was less power to estimate it.

Table 5-12: Predicted and observed median length of stay for babies who survived (competing risks model with time-dependent covariate for gestational age).

Gestational age (weeks)	Observed median length of stay (days)	Predicted median length of stay (days)	Difference (Estimated - Observed)
24	119	122	+3 days
25	104	107	+3 days
26	89	91	+2 days
27	77	78	+1 day
28	65	66	+1 day
29	52	53	+1 day
30	42	42	-1 day
31	32	33	+1 day

Table 5-13: Predicted and observed median length of stay for babies who died (competing risks model with time-dependent covariate for gestational age).

Gestational age (weeks)	Observed median survival time (days)	Estimated median survival time (days)	Difference (Estimated - Observed)
24	8	8	0 days
25	9	11	+2 days
26	9	8	-1 day
27	8	10	+2 days
28	10	6	-4 days
29	8	3	-5 days
30	3	3	0 days
31	7	5	-2 days

The cumulative incidence functions shown in Figure 5-5 provided the proportion of babies who have died or been discharged over time. The final observed proportion of babies who died is compared with that estimated from the model and provided in

Table 5-14. The observed versus estimated proportion of deaths matched each other well at the end of the time period.

Table 5-14: Estimated proportion of babies who have died versus the observed proportion.

Gestational age (weeks)	**Estimated proportion died**	**Observed proportion died**	**Difference**
24	0.377	0.380	-0.003
25	0.212	0.213	-0.001
26	0.155	0.156	-0.001
27	0.114	0.115	-0.001
28	0.071	0.073	-0.002
29	0.036	0.038	-0.002
30	0.027	0.028	-0.001
31	0.022	0.023	-0.001

Plots comparing the observed and estimated proportions of discharge and death over the time since birth are found in Figure 5-9 and Figure 5-10 respectively. Whilst there is a small amount of over- and under-estimation in the babies born at the earliest gestational ages, particularly for the probability of discharge, on the whole the model is robust.

Figure 5-9: Predicted and observed proportions of discharge (competing risks model with time-dependent term for gestational age) over time.

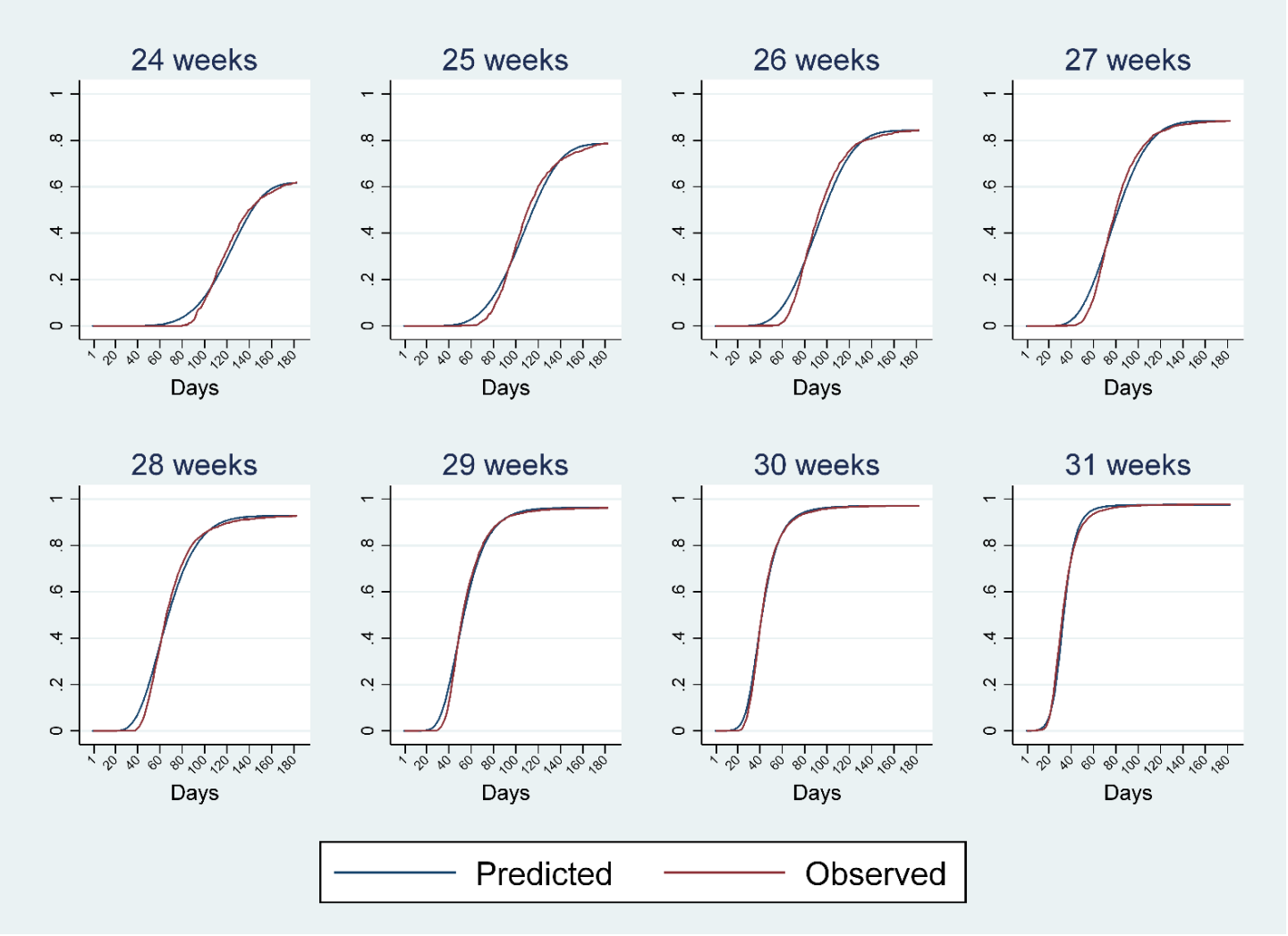
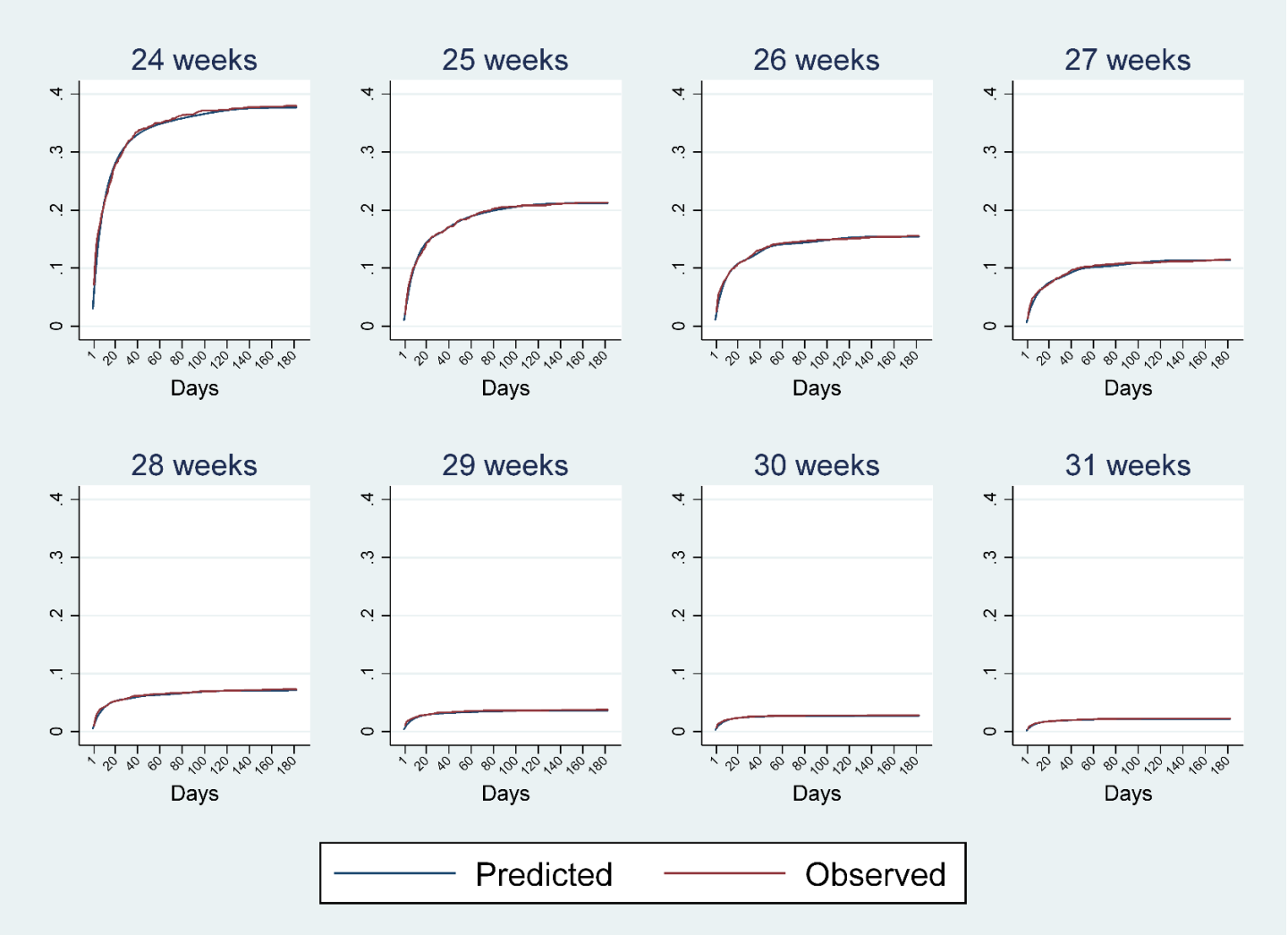


Figure 5-10: Predicted and observed proportions of death (competing risks model with time-dependent term for gestational age) over time.



5.7.7 COMPARISON OF FLEXIBLE PARAMETRIC MODEL WITH THE COX MODEL

This analysis has provided clinically informative estimates of length of stay for babies who survive and those who die, by week of gestational age. A time-dependent covariate was used to model gestational age.

Although time-dependent effects were introduced for gestational age for both events, the proportion of babies who were estimated to die in the proportional hazards model and the non-proportional hazards model was similar. For example, for babies born at 26 weeks the proportion who were estimated to die by one, ten and 30 days from the Cox proportional hazards model was 0.025, 0.085 and 0.120 (Table 5-4) compared to 0.023, 0.084 and 0.117 (Table 5-7) from the flexible parametric model with non-proportional hazards. Similarly, for babies born at 31 weeks gestational age the proportion who were estimated to have died by one, ten and 30 days was 0.004, 0.015 and 0.022 (Table 5-4) respectively from the proportional hazards model. This was similar to the flexible parametric model, with non-proportional hazards which estimated 0.004, 0.014 and 0.019 (Table 5-7). This corroborates with the Therneau-Grambsch test (Table 5-5) which indicated that the hazard of death was modelled adequately by the proportional hazards model.

However, this was not the case with the predicted proportion of babies who were discharged over time. At 30 days after birth the Cox proportional hazards model estimated that 0.022 and 0.034 of babies born at 26 and 27 weeks had been discharged from the neonatal unit (Table 5-4) whilst the non-proportional hazards model estimated this as 0.001 and 0.004 (Table 5-7). The proportional hazards competing risks model poorly estimated the proportion of babies discharged from the neonatal unit, especially in early days after birth, and this may relate to the two different shapes of hazards of discharge seen in Figure 5-3 for babies born at 24 to 28 weeks and for babies born at 29 to 31 weeks gestational age, or that most of the data related to babies born at the later gestational ages. The flexible parametric model with a time-dependent term for gestational age estimated the time at which discharge occurred well.

5.8 FLEXIBLE PARAMETRIC COMPETING RISKS MODEL INCLUDING: GESTATIONAL AGE, SEX AND BIRTHWEIGHT Z-SCORE

The systematic review undertaken in Chapter 2 identified that birthweight, sex and gestational age were the variables most often included in analyses which aimed to predict length of stay and mortality in neonatal care. Therefore, the analysis undertaken earlier in Chapter 5.7 was extended to also include birthweight and sex. Birthweight was modelled as a z-score which provides a measurement of the distance an individual baby's birthweight is from the average birthweight for their gestational age and sex (106). Results are presented as birthweight centiles, i.e. a z-score of 0 is the 50th centile.

Babies with indeterminate sex (n=20) or missing or implausible birthweight (n=118) were excluded from the analysis leaving 20,900 babies included. Only babies with complete data for all these variables were considered in this analysis to allow comparison between the different models.

Variables were considered for inclusion in the model as both fixed effects and time-dependent effects. Table 5-15 describes the model building process and the model with the lowest AIC, highlighted in bold was ultimately selected. This model contained: gestational age (modelled as a time-dependent covariate), sex and birthweight z-score (modelled using splines).

5.8.1 EVENT-SPECIFIC HAZARDS

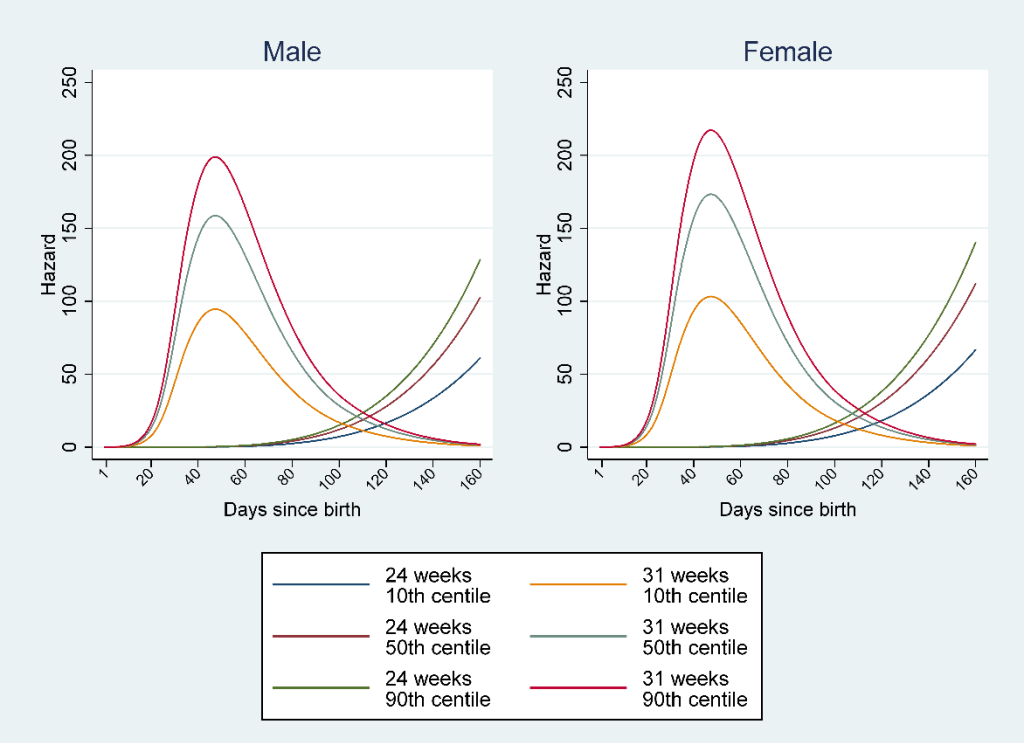
As before, a single value for the event-specific hazard ratio cannot be estimated. Event-specific hazards of discharge are presented in Figure 5-11 for male and female babies born at 24 and 31 weeks with a birthweight at the 10th, 50th or 90th centile.

Table 5-15: Model building process with the corresponding AIC values. A minus/plus sign indicates a decrease/increase in the AIC from the previous lowest value.

Covariates (how they were modelled)	**AIC**
Gestational age (categorical: time-dependent)	23053[4]
Gestational age (categorical: time-dependent) Sex (categorical)	23013 (-)
Gestational age (categorical: time-dependent) Sex (categorical: time-dependent)	23018 (+)
Gestational age (categorical: time-dependent) Sex (categorical) Birthweight z-score (linear)	20429 (-)
Gestational age (categorical: time-dependent) Sex (categorical) Birthweight z-score (linear, time-dependent)	22787 (+)
Gestational age (categorical: time-dependent) **Sex (categorical)** **Birthweight z-score (splines: 3 degrees of freedom)**	**20031 (-)**
Gestational age (categorical: time-dependent) Sex (categorical) Birthweight z-score (splines: 3 degrees of freedom: time-dependent)	23323 (+)

[4] This is different to the previous AIC due to the exclusion of observations with missing data.

Figure 5-11: Event-specific hazards of discharge for selected covariate patterns.

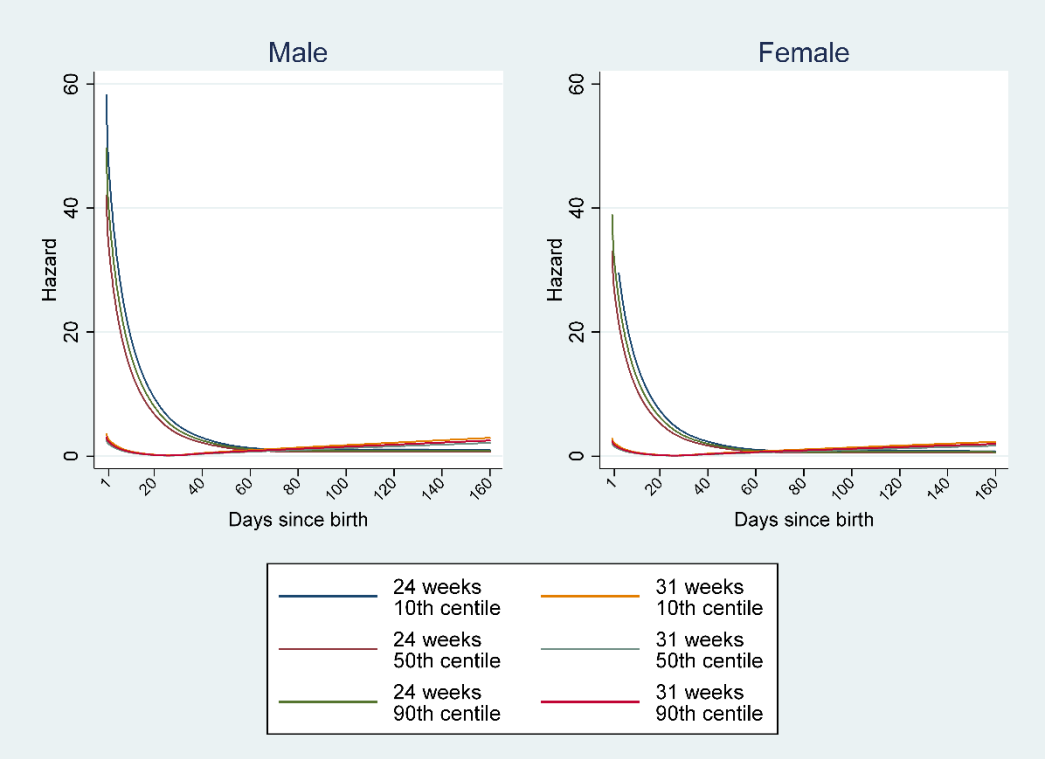


For both sexes, the hazards of discharge for babies born at 24 weeks gestational age compared to babies born at 31 weeks are very different shapes (Figure 5-11). For babies born at 31 weeks gestational age, the hazard of discharge initially increases and peaks at around 50 days before decreasing again. This mirrors that seen in Figure 5-4 where the model was only adjusted for gestational age. Babies born at 24 weeks have a low hazard of discharge which begins to increase after 50 days and continues to increase over time.

Event-specific hazards of death are provided in Figure 5-12 for male and female babies born at 24 and 31 weeks gestational age and a birthweight at the 10th, 50th and 90th centile. The hazard rate for babies born at 24 weeks is initially high and reduces over time (Figure 5-12). Conversely, for babies born at 31 weeks the hazard is initially very low, and it increases steadily after 50 days. This indicates that babies born at this gestational age, who remain in hospital for an extended period, are at an increased hazard of death. This is clinically valid as these babies are likely to be those who are sicker than other 31 week babies, whereas for the very preterm babies most deaths occur earlier. However, the hazard always remains much lower than the initial hazard

for babies born at 24 weeks gestational age. Again, this is similar to that seen from the model with only gestational age (Figure 5-3) although here the results are on the same scale so the shape of the hazard for babies born at 31 weeks gestational age is not as pronounced.

Figure 5-12: Event-specific hazards of death for selected covariate patterns.



5.8.2 CUMULATIVE INCIDENCE FUNCTIONS

Cumulative incidence plots are estimated for combinations of clinically meaningful covariate patterns. Birthweight was estimated at the 10th, 50th and 90th centile for males by week of gestational age to demonstrate the cumulative incidence functions in Figure 5-13 and Figure 5-14. Results for female babies can be found in Appendix 5.

As gestational age increased the proportion of babies that died decreased and the point at which discharge began to occur was earlier. For the most preterm babies born at 24 weeks, there was still a small proportion remaining in hospital even after 160 days, when this axis was cut off. The lowest risk of mortality was in the babies born at 31 weeks with a birthweight at the 50th or 90th centile, where a proportion of 0.018 and 0.020 died respectively. These plots should be interpreted with caution,

particularly where numbers are small. For example, for babies born at 24 weeks gestational age the proportion of deaths is higher in those with a birthweight at the 90th centile than those with a birthweight at the 50th centile. This is potentially due to the small number of babies in this group, causing even one death to contribute a large difference to the total. Figure 5-15 presents the cumulative incidence functions with 95% confidence intervals to demonstrate uncertainty in the estimates for babies born at 24 weeks.

Figure 5-13: Cumulative incidence function plots for males by week of gestational age with a birthweight at the 10th, 50th and 90th centile.

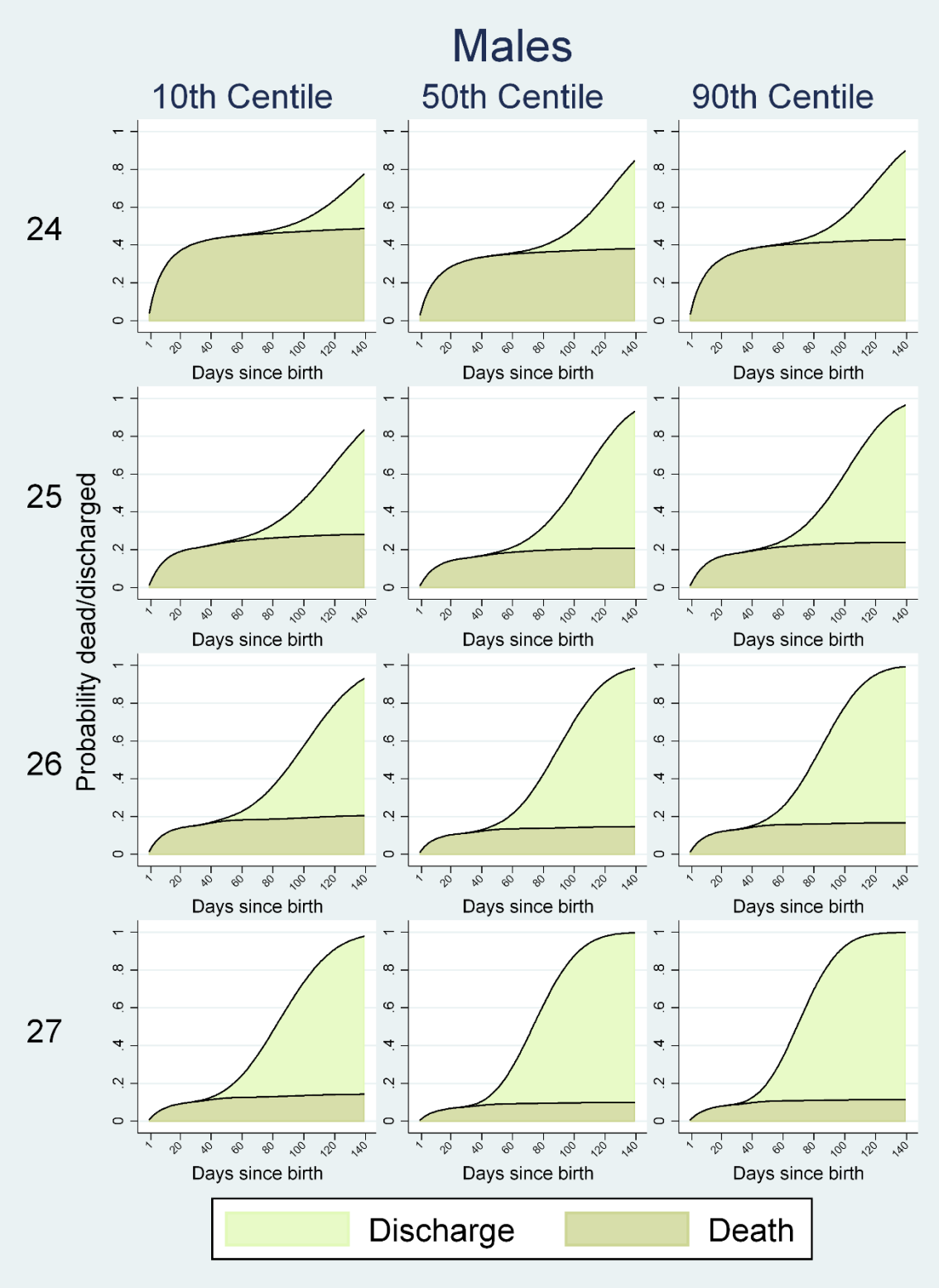


Figure 5-14: Cumulative incidence function plots for males by week of gestational age with a birthweight at the 10th, 50th and 90th centile.

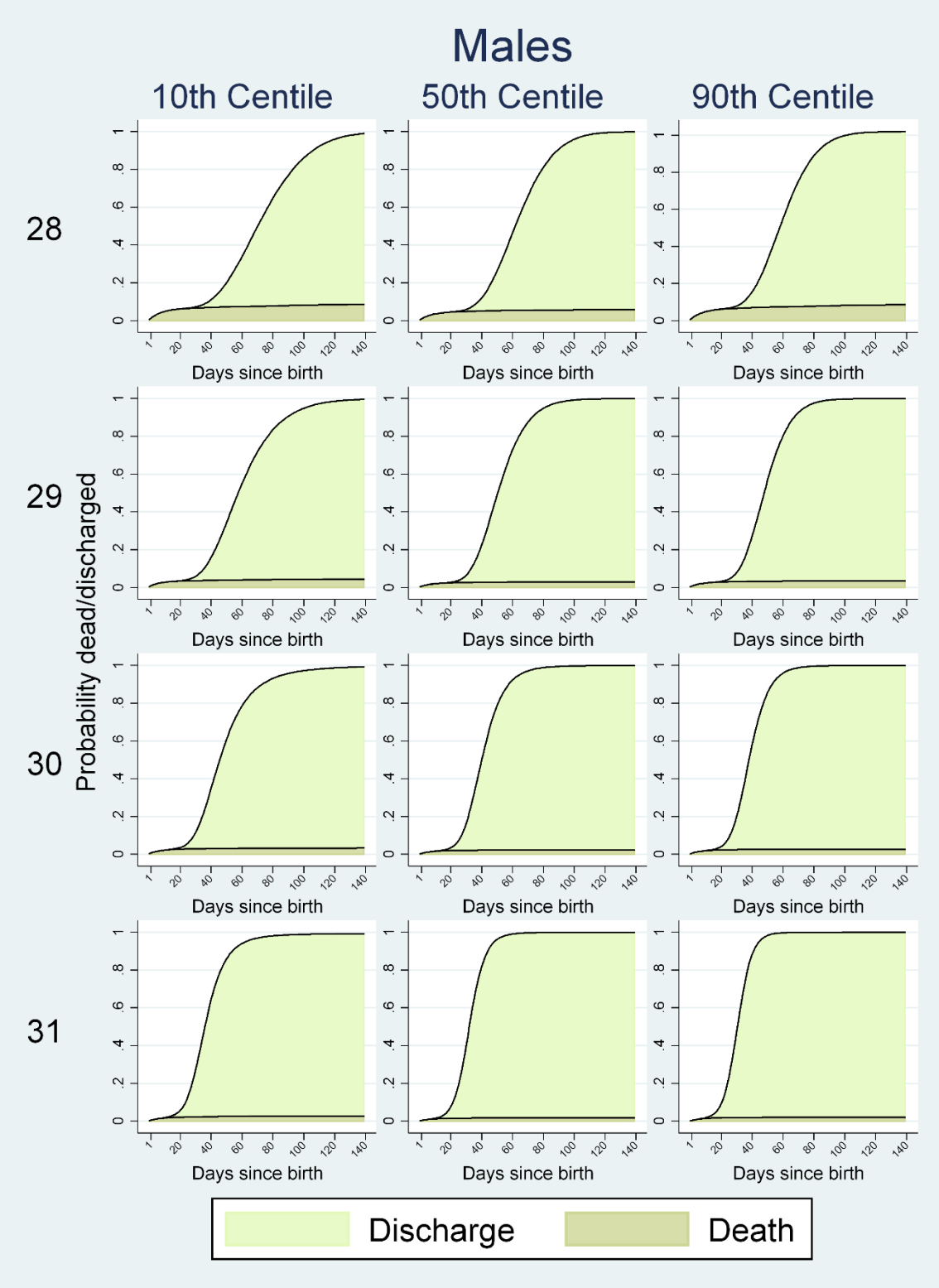
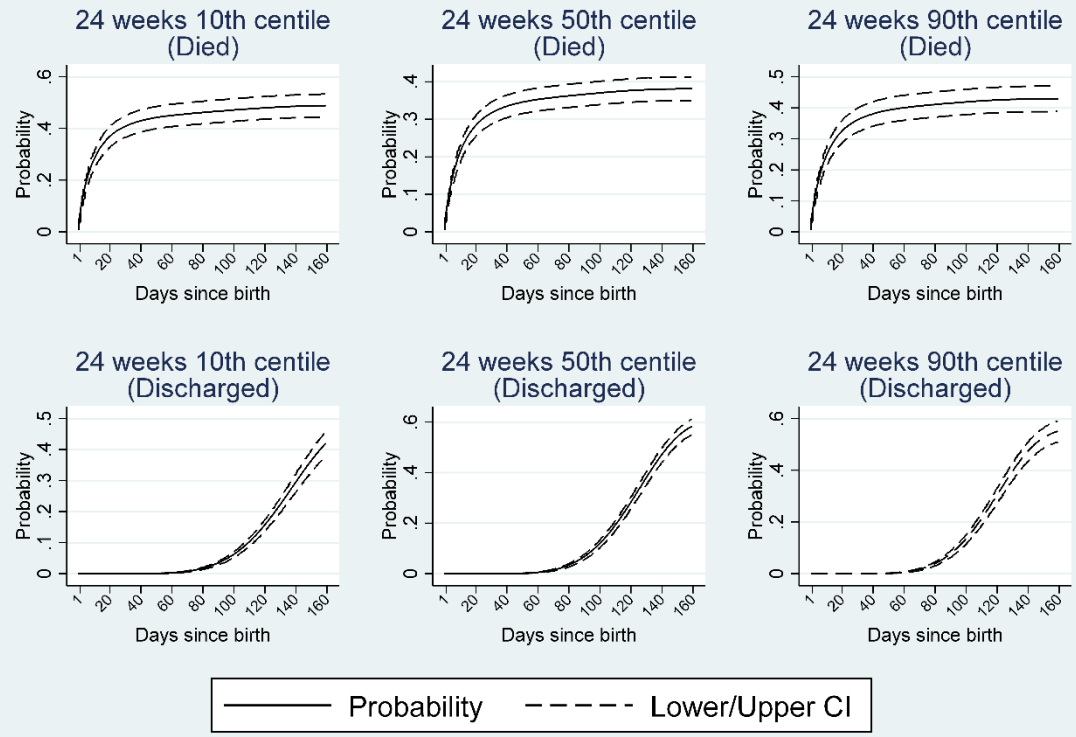


Figure 5-15: Cumulative incidence function for death or discharge with 95% confidence intervals for male babies born at 24 weeks gestational age.



Note: different scales on axes

The proportions of babies who have died or been discharged at certain time points for a given set of characteristics are provided in Table 5-16 to facilitate reading of the cumulative incidence function plots. These results are similar to those found in Table 5-7, where birthweight and sex was not adjusted for, because Table 5-16 represents an 'average' baby.

Table 5-16: Predicted proportion of male babies with a birthweight at the 50th centile who have died or been discharged home one, ten and 30 days after birth.

Gestational age (weeks)	1 day		10 days		30 days	
	Died	**Discharged**	**Died**	**Discharged**	**Died**	**Discharged**
24	0.059	0.000	0.216	0.000	0.317	0.000
25	0.023	0.000	0.106	0.000	0.156	0.0001
26	0.022	0.000	0.080	0.000	0.112	0.001
27	0.013	0.000	0.051	0.000	0.076	0.003
28	0.009	0.000	0.035	0.000	0.048	0.011
29	0.007	0.000	0.020	0.000	0.026	0.039
30	0.005	0.000	0.016	0.000	0.021	0.145
31	0.004	0.000	0.012	0.003	0.016	0.389

5.8.3 ESTIMATING LENGTH OF STAY

To inform clinical counselling of parents, estimates of length of stay can be calculated (Table 5-17 and Table 5-18). Table 5-17 presents median length of stay for babies who die by specific characteristics. Approximately half of all deaths occur in around the first ten days of life. Therefore, ten days of life may be an appropriate time point to prompt clinicians to consider discussing the possibility of a long length of stay, particularly for the early gestational ages, as the risk of mortality reduces. These estimates can be used alongside information from the cumulative incidence function and clinical judgement to counsel parents. For example, for a male baby born at 27 weeks gestational age with a birthweight at the 50th centile around half of deaths (total proportion of deaths around 0.10 from Figure 5-13) have occurred in the first ten days (Table 5-7). At around ten days of life, and using their clinical judgement, a clinician could explain to a parent that the risk of mortality has reduced, but that their baby could be in hospital for a long time. The estimate of median length of stay for a baby of these characteristics is 76 days (66 days by day ten) but I would suggest that clinicians use a more general description, e.g. “around two months” to reflect the uncertainty in this estimate. Future qualitative research could focus on the issues of how to communicate the risk of mortality and length of stay to parents.

Table 5-17: Estimated median length of stay of babies who die in neonatal care, by sex, birthweight and gestational age.

Gestational age (weeks)	**Males**			**Females**		
	Birthweight			**Birthweight**		
	10th centile	**50th centile**	**90th centile**	**10th centile**	**50th centile**	**90th centile**
24	8	8	8	8	9	8
25	10	9	10	11	12	9
26	9	9	10	9	9	7
27	10	10	8	9	10	12
28	11	8	11	12	6	12
29	5	3	2	3	2	4
30	3	5	3	4	9	6
31	5	10	6	8	1	19

The median length of stay for babies who survive is provided in Table 5-18. For example, half of male babies born at 24 weeks with a birthweight at the 50th centile, will have been discharged by 123 days after birth (Table 5-18). For male babies born at 31 weeks with a birthweight at the 50th centile, the median length of stay is 32 days.

Babies born at 24 to 26 weeks gestational age were discharged home at approximately their due date. Babies born at 27 and 28 weeks seem to be generally discharged a few days before their due date, whilst those born at 29 to 31 weeks gestational age, are discharged home much earlier than their due date, sometimes at around 35 weeks PMA.

The differences between the lengths of stay for males and females are very small. Babies born at the 90th centile of birthweight have a similar length of stay to babies born a week later but with a birthweight at the 10th centile. Birthweight for gestational age does impact on length of stay, with the babies born at the 10th centile staying around two weeks longer than those at the 90th centile, particularly at the early gestational ages. These differences reduce for the babies born at 30 and 31 weeks with the length of stay between the smallest and largest babies in those gestational week varying by around one week. These birthweight differences may potentially reflect

clinical practice in not discharging a baby until they reach a certain weight, and those smaller babies have more to gain to achieve this, or it may reflect that birthweight is a proxy for the sickness of the baby. Alternatively these may be errors in the estimation of gestational age and a larger baby may actually be a week older in terms of their gestational age than estimated.

It is not suggested that estimates of these results should be produced for every potential birthweight centile. These estimates of the 10^{th}, 50^{th} and 90^{th} centile could represent a small, average and large baby by gestational age.

Table 5-18: Estimated median length of stay (10th, 90th centile) of discharges for males and females with birthweights at the 10th, 50th and 90th centile.

Gestational age (weeks)	**Males**			**Females**			**Days to EDD**
	Birthweight			**Birthweight**			
	10th centile	**50th centile**	**90th centile**	**10th centile**	**50th centile**	**90th centile**	
24	134 (96, 164)	123 (88, 152)	117 (86, 145)	131 (95, 161)	121 (88, 149)	116 (83, 145)	112
25	117 (80, 153)	106 (73, 137)	101 (71, 130)	115 (80, 150)	104 (71.5, 135)	100 (69, 129)	105
26	102 (67, 136)	91 (62, 121)	85 (58, 114)	99 (66, 134)	88 (60, 117)	84 (58, 112)	98
27	86 (57, 121)	76 (51, 104)	72 (49, 98)	71 (55, 118)	62 (50, 102)	58 (48, 96)	91
28	73 (46, 107)	63 (42, 90)	60 (40, 83)	84 (45, 104)	74 (41, 87)	71 (40, 82)	84
29	58 (38, 90)	51 (35, 73)	48.5 (34, 67)	57.5 (38, 87)	50 (35, 71)	48 (33, 65)	77
30	46 (31, 71.5)	41 (28, 58)	39 (27, 53)	45 (30, 69)	40 (28, 56)	38 (27, 52)	70
31	36 (25, 52)	32 (22, 43.5)	30.5 (22, 41)	36 (24, 51)	32 (22, 43)	30 (21, 40)	63

5.8.4 MODEL FIT AND PERFORMANCE

It is not possible to investigate model fit of this analysis as birthweight z-score has been modelled continuously using splines alongside further adjustments. The cumulative incidence function was estimated for babies born at the 10th, 50th and 90th centile, but in reality there are likely to be few or no babies in this dataset that have a birthweight at exactly this point. As such it is not possible to plot the observed and the estimated cumulative incidence function as done previously (Figure 5-9 and Figure 5-10). However, the median length of stay for babies born at the 50th centile by gestational age (Table 5-17) approximates that estimated from the previous model containing only gestational age (Table 5-8). Since that model performed well, it seems logical that this model with further adjustments also performs well.

The poor availability of model fit measures is a limitation of these methods and a potential compromise would be to categorise all included variables to aid comparison with the observed data. For example, birthweight could be grouped into less than the 10th centile, 10th to 90th centile and greater than the 90th centile. However, this is likely to create other issues as this would lead to loss of data and statistical power and therefore this approach was not applied. Future methodological work is required to identify appropriate methods of assessing model fit.

5.9 DISCUSSION

The second aim of this thesis was to consider methods for estimating length of stay for babies irrespective of whether they survive to discharge or die during their time in neonatal care. Whilst the use of competing risks analysis to model death and discharge from neonatal care is relatively novel, it has been used in other clinical areas including investigating the risk factors for pneumonia and the impact this has on length of stay in adult intensive care units (107). Researchers within the area of pneumonia have advocated the use of competing risks methods to appropriately model death and discharge as competing events when investigating length of stay (108). One study undertaken before this thesis by Hinchliffe, Seaton et al considered use of competing risks methods within the context of neonatal care (6).

This chapter has investigated two different approaches to model competing risks: Cox proportional hazards modelling and flexible parametric modelling. Competing risks methods have been well developed in statistical software (96, 103, 109) with Stata (examples include: *stcrreg* and *stpm2cif*) and R commands (examples include: *mstate* and *cmprsk*) which allow for the use of the Cox model and the flexible parametric model. Software in Stata allows the inclusion of time-dependent covariates (110) within competing risks methods (Stata commands: *stpm2* and *stpm2cif*). In this chapter I used the *stpm2cif* post-estimation command in Stata v 14 and *mstate* in R 3.0.2.

In its simplest form, the flexible parametric model provides the same results as the Cox model, which was demonstrated in Chapter 4 (Table 4-1 and Table 4-4). Whilst it is possible to relax proportional hazards in the Cox model, the main advantage of flexible parametric modelling is that time-dependent effects can incorporated easily in routine software. This allows differences in the hazard between groups to vary over time.

A flexible parametric model with a time-dependent effect for gestational age reduced the AIC substantially, indicating that model fit was improved by modelling gestational age in this way. When comparing the cumulative incidence functions with that estimated from the proportional hazards model, results were most different when comparing the proportion of babies who had been discharged over time. This introduction of a time-dependent effect allowed the time to discharge to vary between the different gestational ages over time.

A flexible parametric model was then built using variables identified as being clinically informative for estimating length of stay and the risk of mortality from the systematic review (Chapter 2). This model included gestational age; birthweight z-score and sex of the baby. Cumulative incidence plots and estimates of length of stay were calculated for this model. From this it was possible to provide estimates of median length of stay for babies who survived and those who died. Broadly, babies born extremely preterm who survived to discharge remained in the neonatal unit until around their due date. However, babies born later were discharged home earlier, and those born at 30 and 31 weeks gestational age were often discharged within four to six weeks of birth. Clinicians can use these length of stay estimates when counselling families. But they should also use their judgement and not provide specific estimates of estimated

discharge dates without explanation that these are not definite dates. For example, if a baby's predicted length of stay is 60 days the clinicians may wish to say "around two months" rather than providing a specific estimated discharge date. Alternatively, if a 'target' discharge date is wanted, then the clinician may wish to provide a date calculated from the 75th centile for length of stay for a baby of given characteristics, and this can be revised down to the median length of stay if a baby appears to be doing well. These estimates are useful to prepare parents for the likely length of stay of their baby.

Plots of the cumulative incidence function can be produced for combinations of covariates as appropriate. These provide useful information about the risk of mortality and alongside the median length of stay these can aid communication of the results in this work. These results should not be produced for every birthweight centile, as there is likely to be little difference to results between specific centiles. Instead cut points such as those provided for the 10th, 50th and 90th centile could be used to describe small, average and large babies.

5.9.1 USE OF THIS RESEARCH FOR INFORMING CONVERSATIONS WITH PARENTS OF PRETERM BABIES

The cumulative incidence functions produced provide an estimated proportion of deaths for babies of given characteristics who are admitted to the neonatal unit. The plots can also be used to identify a time point when the risk of mortality has plateaued, providing a potentially appropriate time to discuss length of stay with parents. Generally, this point was at around ten days of life.

A parent panel meeting to discuss how to communicate these results to parents of very preterm babies suggested that parents should not be shown these plots or results. Similarly, this group did not want parent-friendly versions of these results produced in leaflets. This group recommended that the results should be used by clinicians to form the basis of a conversation which was also informed by clinician experience and expertise. For example, for the parents of a male singleton baby born at 24 weeks these results could be used to say the following:

- Around the time of admission, using clinician judgement: *"around 70% of babies like your baby will survive"*
- After approximately ten days, and with use of clinician judgement: *"it is quite likely now that your baby will survive, but he is likely to stay in hospital until around the day he should have been born"*

For a baby boy born at 31 weeks the conversations are likely to use the results here to frame the conversation differently. For example:

- Around the time of admission, using clinician judgement: *"most babies like your baby will survive"*
- After approximately ten days, and with use of clinician judgement: *"it is likely that your baby will need to be in hospital for a few more weeks"*

Parent panel meetings have been held to discuss communication of these results. Ongoing work with Bliss (charity to support families of preterm and sick babies) and other stakeholders is identifying appropriate methods of communication of these results with clinicians.

5.9.2 COMPARISON WITH OTHER PUBLISHED RESEARCH

There has been little published research that has investigated length of stay in the neonatal unit, and the work that has been undertaken generally excludes babies that die. The published systematic review from this thesis identified nine papers published which met the eligibility criteria within this area (18) although other studies were also published in areas of the world not covered by this review (65, 66).

One paper investigated how birthweight impacted on length of stay and found similar results to those presented here. The analysis by Lee et al (5) suggested birthweight for gestational age had the largest impact on length of stay from all the variables included in the analysis. Whilst gestational age was not included in the analysis by Lee et al, they found that birthweight impacted on length of stay by around two weeks for the smallest babies (<1000g) but only a few days for the bigger babies, which is similar to the results from this analysis. Lee et al also noted that sex of the baby had a minimal impact on length of stay, with a difference of one to four days on overall length of stay

between the sexes, and this result has been replicated in this thesis (5). Research by Altman et al (58) also found that the sex of the baby did not have an impact on the PMA at which they were discharged.

The paper by Hinchliffe, Seaton et al undertaken prior to this thesis (6) focused on all births at 24 to 28 weeks whilst in this analysis only singletons were considered. Therefore, the results presented in this chapter have a higher survival rate as multiple births represent a large proportion of preterm births and have a higher risk of mortality in the preterm population, particularly at 24 to 27 weeks gestational age (111).

5.9.3 ALTERNATIVE STATISTICAL APPROACHES

An event-specific hazard approach (commonly referred to as a cause-specific approach elsewhere) was used for this analysis. An alternative would have been to use an approach such as the Fine and Gray method (102), which models the sub-distribution hazard. The difference between the event-specific hazard and the sub-distribution hazard is the risk-set. For the event-specific approach the risk-set decreases each time there is an event due to any reason, i.e. there is censoring. However, when the sub-distribution approach is used the individuals remain in the risk-set. This alternative approach of using sub-distribution hazards can also provide a cumulative incidence function.

The disadvantage of the event-specific hazard approach is that the cumulative incidence is a function of all the event-specific hazards and the probability of each event occurring. Therefore, there is no one-to-one relationship between the event-specific hazard and the probability of that specific outcome (101). Whilst the Fine and Gray approach could overcome this, it has other disadvantages. As participants who fail from one event remain in the risk-set for the other causes they theoretically remain in the risk-set forever. Therefore, the hazard function from this approach does not represent an epidemiological rate (112) and potentially the probability of all the competing events can sum to more than one (113).

Event-specific hazard approaches have been advocated when there is interest in all of the competing events as the event-specific hazard rates and cumulative incidence

functions can provide useful information for clinical discussion (101, 112). The total probability can be broken down into the different competing events and, as discussed, this can provide a useful measure to aid clinical decision making (104).

5.9.4 STRENGTHS AND LIMITATIONS

A strength of this work is that it is the first time estimates of length of stay have been produced for all singleton very preterm babies which also account for the outcome of the baby. A large, national database has been used to provide these estimates which can be used by clinicians in the counselling of parents throughout the neonatal period.

A limitation of this work is that a small number of babies are discharged to facilities not captured by the NNRD, and because they do not return to an NNRD unit it is unclear what their final length of stay was. However, a sensitivity analysis was undertaken which indicated that the impact this assumption had on the results was minimal, and therefore including them rather than excluding them and losing statistical power appears appropriate. Also, inspection of these data indicated that most of these babies were discharged to postnatal wards based in other English hospitals. Some of these babies could have been considered as an additional outcome, for example discharge to other specialist services. However, the small number of babies experiencing this outcome would have resulted in unreliable estimates.

A final limitation of this analysis is the difficulty in assessing model fit. The observed and estimated probabilities and lengths of stay were compared where possible. The results appeared robust, although as additional covariates were included it was not possible to assess this further.

5.10 CHAPTER CONCLUSION

This chapter used competing risks analyses to estimate the length of stay for both babies who survive to leave neonatal care and those who die in neonatal care. A Cox proportional hazards model was used before the approach was extended to the flexible parametric model approach, which allowed the introduction of time-dependent covariates. A model adjusted for gestational age, birthweight z-score and

sex was used to estimate the cumulative incidence functions, the proportion of babies experiencing each of the events, over time.

The median length of stay for babies who died was generally around ten days. Therefore, using clinical judgement, this time point may be appropriate to discuss length of stay, if it has not already been discussed, as the risk of mortality begins to reduce. Median length of stay was estimated for babies of specified characteristics who survived to discharge, for example for babies born at 24 weeks gestational age the median length of stay was 122 days. These estimates can be used by clinicians in discussion with parents about length of stay.

6 MULTISTATE MODELLING FOR THE ESTIMATION OF MORTALITY AND LENGTH OF STAY AT EACH LEVEL OF CARE

6.1 OVERVIEW OF CHAPTER

In this chapter the theory of survival analysis and competing risks is extended to a multistate analysis framework. This allows events which occur before the final endpoint to be considered as transient or intermediate states. The levels of neonatal care will be considered as transient states within the analysis, and an unadjusted and adjusted multistate model will be undertaken to describe the neonatal care pathway and provide estimates of mean length of stay at each level of care.

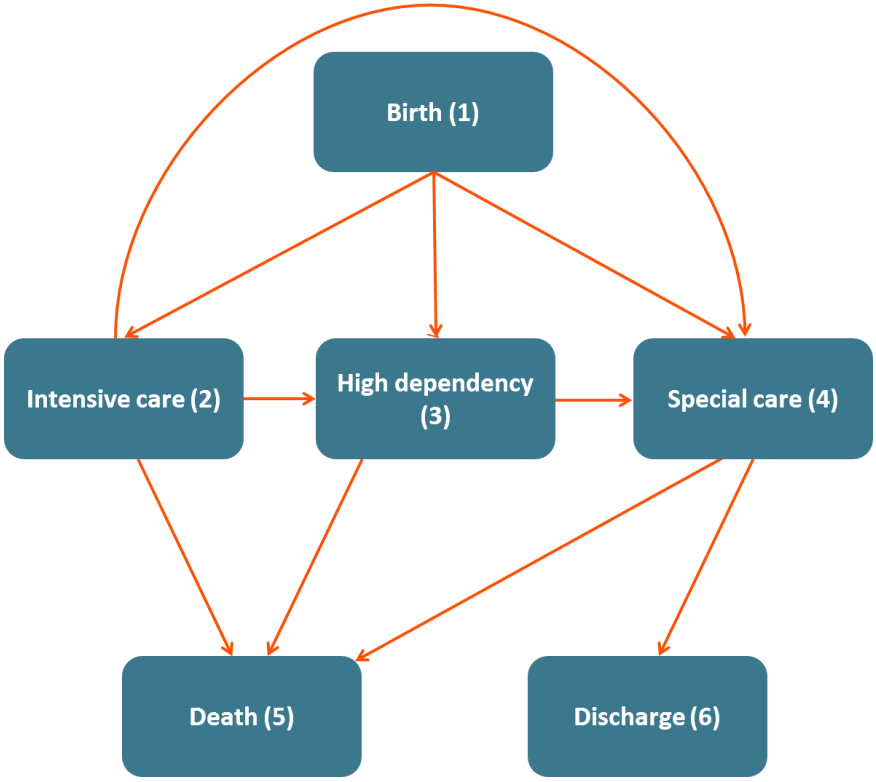
6.2 INTRODUCTION

In Chapter 5, competing risks methodology was introduced where one initial state and several mutually exclusive competing endpoints (absorbing states) were considered, for example, death and discharge. To extend this further, intermediate states known as **transient states** can be introduced and this then becomes a multistate model. These transient states can represent any event that occurs before the final absorbing state. For example, a state could represent relapse from a disease. In this analysis, the transient states are the different levels of care (Chapter 1.3) required by a baby before they either die or are discharged from the neonatal unit (Figure 6-1).

A standard survival analysis can be considered as being a multistate model that has one initial state and one final absorbing state. A competing risks model, such as that presented in Chapter 5, is a special case of a multistate model where there is one initial state and multiple absorbing states. A multistate model can be considered as a series of nested competing risks models. For example, for a baby in the intensive care state the competing events are: having their care stepped down to high dependency care; having their care stepped down to special care or dying (Figure 6-1). If the baby's care was stepped down to high dependency care then they would be censored from being at risk of death (along the transition from intensive care) and of stepping down

from intensive care to special care. A new competing risks model would be formed with the competing events of: death (following high dependency care) or stepping down to special care. A multistate model contains a finite number of states $S = \{1,2,3,4 \dots J\}$. In Figure 6-1 there are six states. One is an initial starting state: birth and two are absorbing states: death and discharge. The other three states are intermediate or transient states, meaning that upon entering these states it is possible to exit them again.

Figure 6-1: An example of a multistate model with birth as the initial state and absorbing states of death or discharge. All other states are transient.



6.2.1 DATA PREPARATION FOR MULTISTATE MODELLING

The approach for data preparation for a multistate model is similar to that used for the competing risks methods in Chapter 5. This is extended to provide one row of data for each potential transition that an individual is at risk of experiencing, up to a maximum number of rows equal to the maximum number of transitions in the model. If an absorbing state is reached before becoming at risk of another transition, then they will

not have a row of data for that transition. In this chapter I undertook data manipulation in Stata v 14 and the analysis in R 3.0.2 using the *mstate* command.

An example of two hypothetical patients is provided in Table 6-1 in wide format. The status variables indicate whether the transition occurred (0: no, 1: yes) whilst the entry time variables indicate when the event was experienced, or when the transition was censored. This data can be manipulated into long format as in Table 6-2 where the start and stop times indicate when the patient started and stopped being at risk of a transition.

Baby 1 spends time receiving intensive care, high dependency care, and special care before being discharged home alive after 163 days. Therefore, in the long dataset they have ten rows of data (Table 6-2) as they have been at risk of experiencing all ten potential transitions (Figure 6-1). Baby 2 receives 52 days of intensive care before dying. Therefore, they only have six rows of data (Table 6-2) as they are only ever at risk of experiencing six transitions. These six transitions are the three from the birth state and the three from the intensive care state (two to lower levels of care, one to death).

Throughout this chapter the *neonatal care pathway* refers to the hierarchically collapsed total days at each level of care which a baby receives, rather than the individual movements between levels of care received on a day-by-day basis for an individual baby. This approach to considering care means that the results are informative for commissioning, and provide information about the neonatal care service as a whole, but cannot be used to inform the day-to-day clinical management of individual babies.

Table 6-1: The standard format of datasets in wide format with one row per baby.[5]

ID	IC entry time	IC status	HD entry time	HD status	SC entry time	SC status	Died entry time	Died status	Home entry time	Home status
1	1	1	56	1	88	1	163	0	163	1
2	1	1	52	0	52	0	52	1	52	0

Table 6-2: The data following modification into long format, with one row for each potential transition in the model.[5]

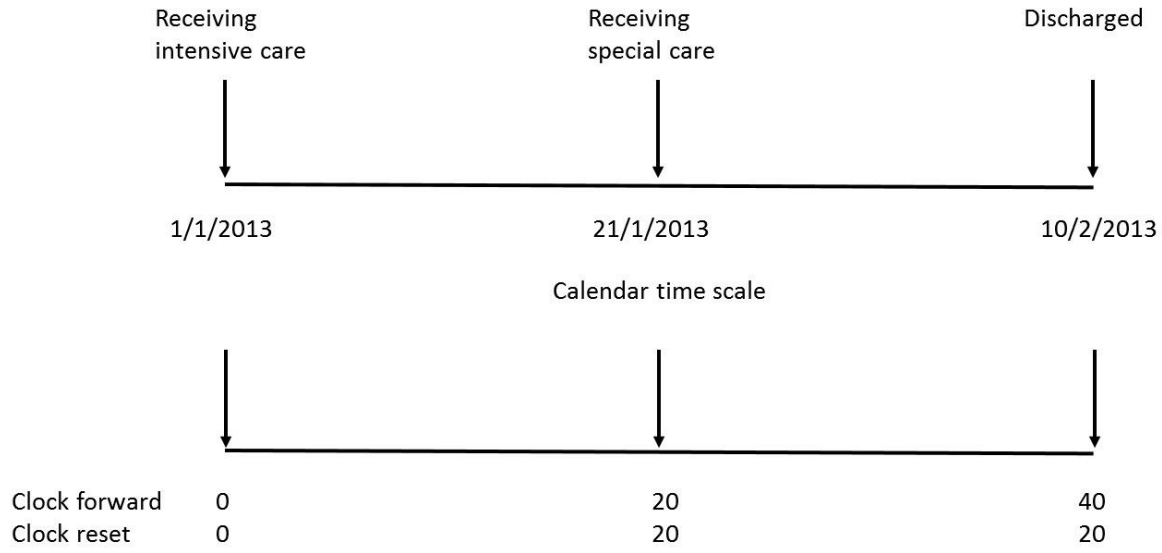
ID	Start	Stop	Status	From	To
1	0	1	1	Birth	IC
1	0	1	0	Birth	HD
1	0	1	0	Birth	SC
1	1	56	1	IC	HD
1	1	56	0	IC	SC
1	1	56	0	IC	Death
1	56	88	1	HD	SC
1	56	88	0	HD	Death
1	88	163	1	SC	Home
1	88	163	0	SC	Death
2	0	1	1	Birth	IC
2	0	1	0	Birth	HD
2	0	1	0	Birth	SC
2	1	52	0	IC	HD
2	1	52	0	IC	SC
2	1	52	1	IC	Death

[5] In this table the following acronyms apply: IC intensive care; HD high dependency; SC special care.

6.2.2 TIMESCALES AND THE MARKOV ASSUMPTION

In the context of multistate modelling there are two different time scales which the time t can refer to known as: ***clock forward*** or ***clock reset***.

Figure 6-2: Example of the two time scales in multistate modelling: clock forward and clock reset.



The time scale in a clock forward analysis refers to the time since the participant entered the initial starting state. The clock continues moving forward for the participant as intermediate events occur, until the final event occurs. In a clock reset analysis, the time t of the hazard rate depends on the time since entry into a given state j and the clock is reset to 0 each time the patient enters a new state.

A natural choice of time scale when considering neonatal care is birth denoting time zero and time since birth being measured, i.e. a clock forward approach. However, in reality there is thought to be very little difference in the parameter estimates from use of either clock forward or clock reset (76). A preliminary investigation of the clock reset approach in this thesis found similar results for most transitions (results not presented). Throughout this thesis the clock forward approach is used.

Many multistate models rely on the Markov assumption. This assumption is often described as being 'memoryless' as the state that a participant occupies at time t only

depends on the current time $s, s \leq t,$ and on the current state which they occupy. It does not depend on the history of the participant until that point.

If the Markov assumption is strictly applied then only clock forward models can be defined as Markov models and a clock reset model, which forces a slight relaxation of the Markov assumption, is often referred to as semi-Markov (76). This is because a clock reset approach does not depend on the current time, but on the duration spent in the current state (114). Although the Markov assumption can appear restrictive, it is convenient as the calculation required to estimate the probabilities would otherwise become very complex. In this thesis, all care is assumed to occur hierarchically and therefore the Markov assumption is likely to be reasonable as stepping up or a subsequent step down between the same levels of care is not considered in the analysis. This allows for interpretation and analysis which is meaningful for commissioning, but not for clinical decision making which will be interested in the movements between levels of care which a baby may take. If this was considered then the Markov assumption would need to be relaxed as information about the movement between levels of care would need to be considered, thus requiring a non-Markov model approach.

It is possible to assess the validity of the Markov assumption by including the entry time into a state in the model as a covariate to assess whether the hazard ratio is different from one, indicating a potential violation of the Markov assumption (115). If the hazard ratio is greater than one this indicates that as entry time into a given state increases the hazard of making the next transition increases. Similarly, if the hazard ratio is less than one, this indicates that as entry time increases, the hazard of making the next transition decreases. Both of these situations indicate a potential violation of the Markov assumption. Competing risks methods, introduced in Chapter 5, are always Markov models as there is no event history (76).

Relaxing the Markov assumption via a change of timescale is a possibility, although some researchers do not recommend it as changing the timescale in this way introduces confusion (100). Generally, the choice of the timescale is a question of how to best answer the question of interest (76, 116) although more formal procedures have been suggested to identify the appropriate timescale (117). Limited research has

investigated non-Markov models (118) and future work is needed in this area. Throughout this chapter, Markov multistate models will be used, to allow for use of the clock forward timescale, but in keeping with convention, the word 'Markov' will be dropped.

6.3 CALCULATION OF THE HAZARD RATE AND TRANSITION PROBABILITIES

There are two statistics of interest within multistate modelling: the **transition hazards** and the **transition probabilities**. The transition hazards can be interpreted in a similar way to those from a standard survival analysis (the hazard) or a competing risks approach (referred to in this thesis as the event-specific hazard). The hazard ratio for a specific transition can provide a comparison between two or more groups and indicate whether a group is at an increased hazard of that specific transition. The probabilities of being in any state over time can be presented and used in a similar way to the cumulative incidence functions from a competing risks analysis.

6.3.1 EXTENDING THE HAZARD RATE TO MULTISTATE MODELLING

In Figure 6-1 each potential state is represented by a box with transitions denoted by arrows between the states. The transition between state l and state j can be represented by $l \rightarrow j$. If T denotes the time that j is reached from l then the **transition hazard rate** (also known as the transition intensity) is:

$$\alpha_{lj}(t) = \lim_{\delta \to 0} \frac{P(t \leq T < t + \delta | T \geq t)}{\delta}$$

Equation 6-1

Comparison of the transition hazard rate with the hazard for a standard survival analysis (Equation 4-5) and the event-specific hazard for the competing risks approach (Equation 5-1) demonstrates the similarity between all these approaches. The cumulative hazard for the $l \rightarrow j$ transition is:

$$A_{lj}(t) = \int_0^t \alpha_{lj}(u)du$$

Equation 6-2

As in Equation 4-9, the Nelson-Aalen estimator of the cumulative hazard can be used and is extended to multistate modelling by being considered for each separate transition. Further details can be found in the references: (100, 119).

The hazards can be described in a matrix, and from Figure 6-1 this can be expressed as:

TO / **FROM**	BIRTH	INTENSIVE	HDU	SPECIAL	DEATH	DISCHARGE
BIRTH	–	$\alpha_{12}(t)$	$\alpha_{13}(t)$	$\alpha_{14}(t)$		
INTENSIVE		–	$\alpha_{23}(t)$	$\alpha_{24}(t)$	$\alpha_{25}(t)$	–
HDU		–	–	$\alpha_{34}(t)$	$\alpha_{35}(t)$	–
SPECIAL		–	–	–	$\alpha_{45}(t)$	$\alpha_{46}(t)$
DEATH		–	–	–	–	–
DISCHARGE		–	–	–	–	–

The time between birth and the next state is artificially produced as a random number between zero and one to allow prediction of all babies from the same initial state. Therefore this hazard should not be interpreted.

6.3.2 CALCULATION OF TRANSITION PROBABILITIES

Transition probabilities are calculated as the probability that a random participant will be in state j at time t conditional on them having previously been in l. The probabilities are estimated using the Aalen-Johansen estimator, which is the matrix version of the Kaplan-Meier estimator. The Aalen-Johansen estimator is also known as the empirical transition matrix and full details can be found in references: (100, 120). In Markov models, given the cumulative hazard, the transition probability matrix can be calculated via use of a product integral (119, 121). When the transition probabilities are stacked in a plot the distance between them represents the probability of being in the corresponding state (96).

Multistate analyses can be undertaken using the *mstate* command (assuming Cox proportional hazards) or *flexsurv* (parametric models) in R 3.0.2. In this thesis I use the *mstate* command throughout.

6.3.3 CALCULATION OF EXPECTED LENGTH OF STAY

The expected length of stay (122) in any given state k is calculated by integrating the probability of being in that state:

$$E_k^{\tau} = \int_0^{\tau} P(X_u = k)\, du$$

Equation 6-3

where τ is a fixed value, usually representing the end of follow up time. For $\tau < \infty$ this is known as the restricted expected length of stay. In this work, expected length of stay is reported in completed days as for commissioning purposes only whole days are measured. It is currently not possible to calculate asymptotic confidence intervals for the expected length of stay (personal communication with developers of *mstate* command). Therefore, where possible measures of uncertainty for expected length of stay are provided from 1,000 bootstrap samples to create percentile confidence intervals (123). In this thesis, I formed a bootstrap sample from random data which were selected from the original sample with replacement to create a new sample the same size as the original dataset. The analysis is replicated on this newly created dataset, and

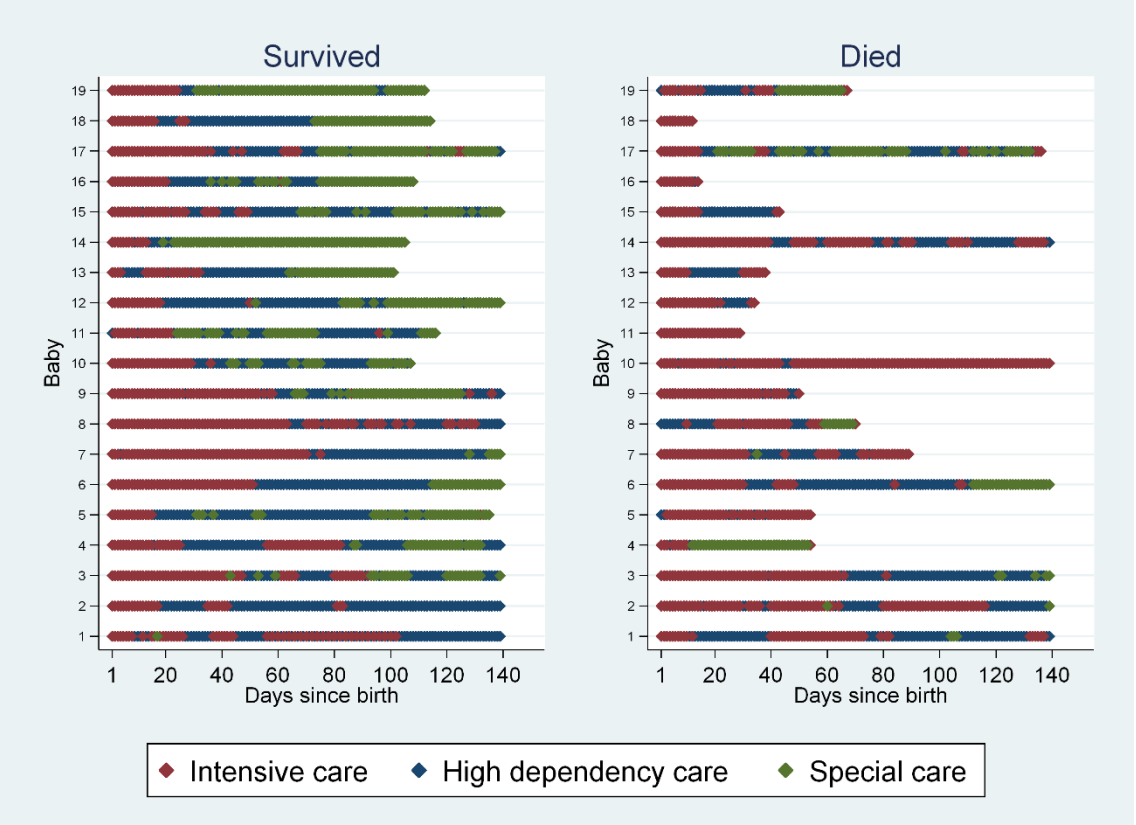
this is repeated multiple times, with the results stored each time, and the 2.5th and 97.5th centiles used to provide measures of uncertainty.

6.4 OBSERVED DATA

The 21,038 singleton very preterm babies required 1,205,422 days of care. The majority of this care (approximately 52%) was at the lowest level: special care, with a total of 629,413 days provided. High dependency care accounted for 329,568 days (27%) and intensive care accounted for 246,411 days (21%) (Table 3-1).

Care was assumed to occur hierarchically, that is, all intensive care was received before high dependency care, and all high dependency care before special care to form the model presented in Figure 6-1. Figure 6-3 presents examples of the observed care and from this all intensive care was collected together to form the time spent in state two of the multistate model, all high dependency care was amalgamated to form the time spent in state three, and all special care was amalgamated to form the time spent in state four. Whilst this assumption, that care occurs hierarchically, is an over-simplification, for commissioning purposes it is irrelevant what order the care is received in as the number of days and costs remain the same. However, this is a limitation if interest lies in the clinical care pathways of an individual baby.

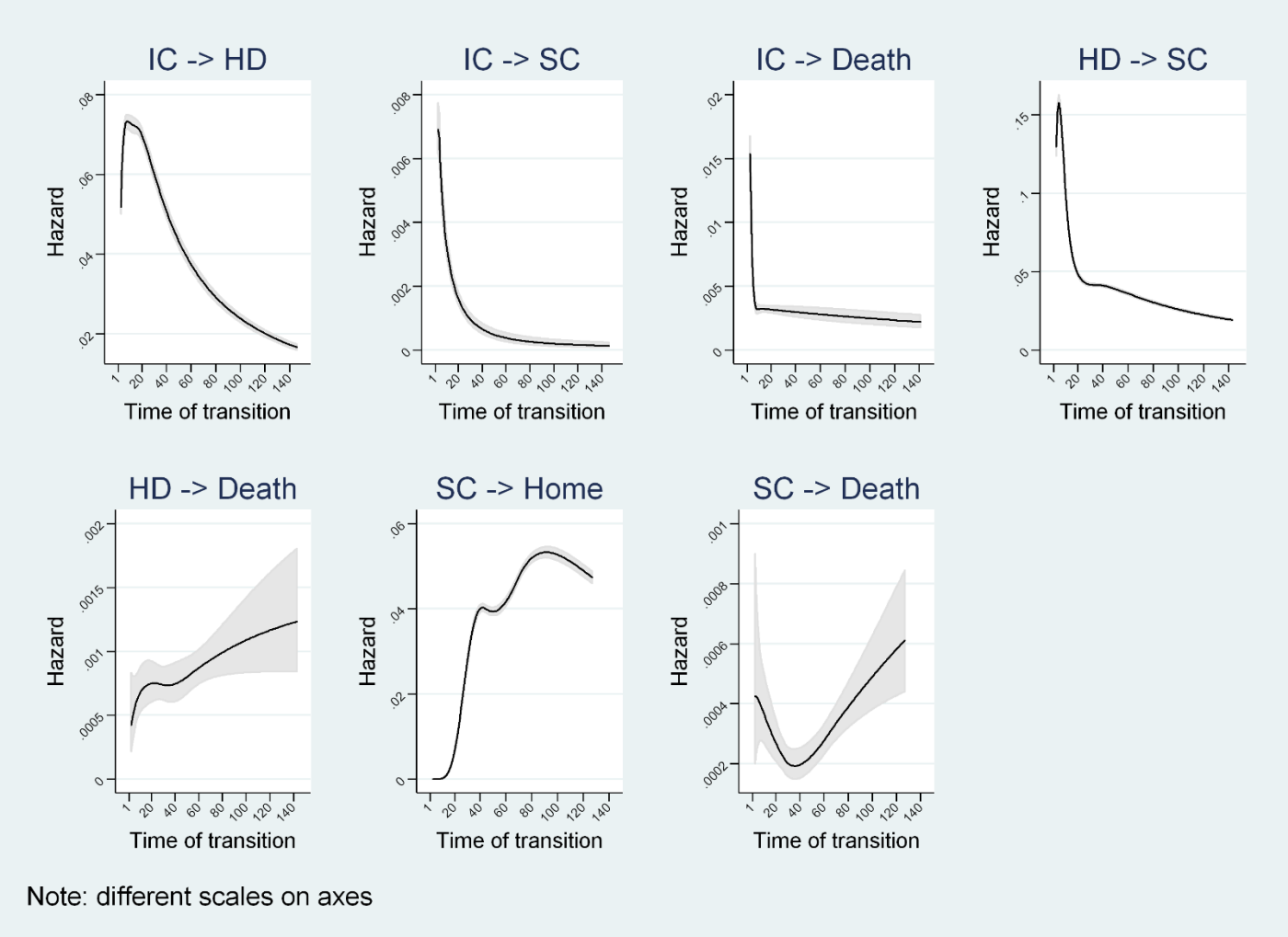
Figure 6-3: Patterns of observed levels of care for a random selection of babies who survive to discharge from neonatal care and who die in neonatal care.



The observed hazard of each transition was estimated (Figure 6-1) (89) with 95% confidence interval (Figure 6-4). The hazard can be interpreted as a rate of the transition although the shape of the hazard is of more interest to investigate. For example, the hazard of dying following only intensive care sharply decreases over time until around day seven and thereafter it continues to decrease more slowly over time. This indicates that if a baby survives to day seven then the hazard of experiencing death immediately after only intensive care is low.

Several of the observed hazards (Figure 6-4) have turning points and use of a parametric approach would be unlikely to capture the shape adequately and a Cox model may help as no assumptions are imposed on the shape of the hazard.

Figure 6-4: The observed hazard with 95% confidence intervals for each transition in the multistate model.[6]



[6] In this figure the following acronyms are used: IC intensive care; HD high dependency and SC special care

A matrix of transitions with the number of babies making the identified transition is:

TO / **FROM**	BIRTH	INTENSIVE	HDU	SPECIAL	DEATH	DISCHARGE	TOTAL
BIRTH	–	17,269	2,796	973			21,038
INTENSIVE		–	15,129	824	1,316	–	17,269
HDU		–	–	17,665	260	–	17,925
SPECIAL		–	–	–	186	19,276	19,462
DEATH		–	–	–	–	–	–
DISCHARGE		–	–	–	–	–	–
TOTAL		17,269	17,925	19,462	1,762	19,276	75,694

As babies can experience multiple transitions some babies contribute to the total for several transitions. The sum of the deaths and discharges will total 21,038 as these are final events (absorbing states).

A stacked plot providing the observed proportion of babies in that category over time is provided in Figure 6-5. It is possible to read the proportion in any category at any point in time. For example, seven days after birth the proportion of babies receiving intensive care was 0.525; receiving high dependency was 0.227; receiving special care was 0.207 and 0.04 had died. No babies had been discharged home at seven days after birth. The proportion of babies in each of the different states is also provided in Table 6-3 on selected days following birth.

Table 6-3 provides an estimate for all babies together, and to investigate this further, important factors such as gestational age need to be considered. Figure 6-6 and Figure 6-7 provides the observed data again for babies grouped into two gestational age brackets: 24 to 28 weeks (extremely preterm) and 29 to 31 weeks (very preterm). These plots demonstrate that the mortality is higher in the earlier gestational ages, and that, as expected, the need for higher levels of neonatal care, including intensive care, is increased in the babies born at 24 to 28 weeks gestational age. This will be explored further later in the chapter when the analysis will be adjusted for gestational age.

A plot of the observed care and of having died or been discharged for every week of gestational age is provided in Figure 6-8.

Figure 6-5: Observed proportion of babies in each of the different categories of levels of care or who have been discharged or died over time.

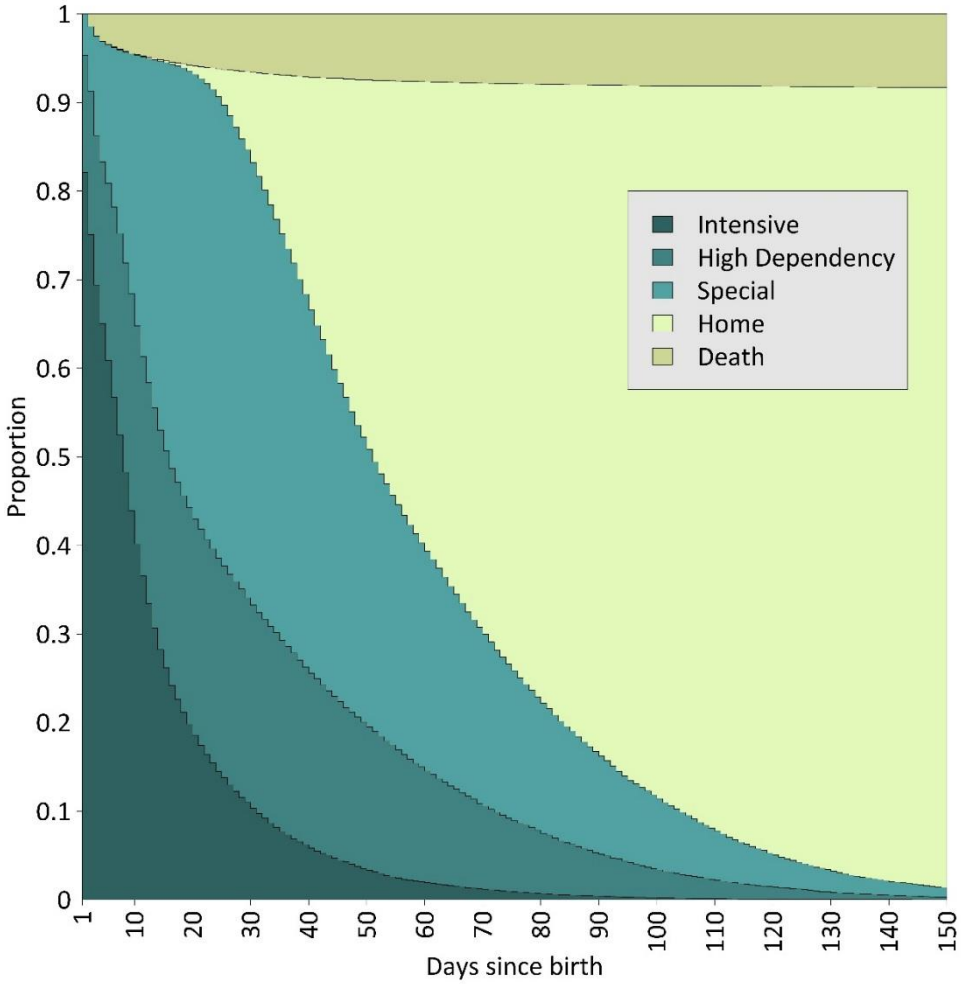


Table 6-3: Observed proportion (95% confidence interval) of babies in each of the different states on specific days following birth.

Day after birth	Intensive care	High dependency	Special care	Discharged	Died
1	0.821 (0.815, 0.827)	0.133 (0.129, 0.137)	0.046 (0.044, 0.048)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)
2	0.751 (0.745, 0.757)	0.162 (0.156, 0.168)	0.073 (0.069, 0.077)	0.000 (0.000, 0.000)	0.014 (0.012, 0.016)
3	0.694 (0.688, 0.700)	0.169 (0.163, 0.175)	0.112 (0.108, 0.116)	0.000 (0.000, 0.000)	0.025 (0.023, 0.027)
4	0.620 (0.614, 0.626)	0.183 (0.177, 0.189)	0.136 (0.132, 0.140)	0.000 (0.000, 0.000)	0.031 (0.029, 0.033)
5	0.609 (0.603, 0.615)	0.200 (0.194, 0.206)	0.157 (0.151, 0.163)	0.000 (0.000, 0.000)	0.034 (0.032, 0.036)
6	0.568 (0.562, 0.574)	0.214 (0.208, 0.220)	0.180 (0.174, 0.186)	0.000 (0.000, 0.000)	0.037 (0.035, 0.039)
7	0.525 (0.519, 0.531)	0.227 (0.221, 0.233)	0.207 (0.201, 0.213)	0.000 (0.000, 0.000)	0.040 (0.038, 0.042)
10	0.402 (0.396, 0.408)	0.247 (0.241, 0.253)	0.306 (0.300, 0.312)	0.001 (0.001, 0.001)	0.045 (0.043, 0.047)
14	0.282 (0.276, 0.288)	0.248 (0.242, 0.254)	0.416 (0.410, 0.422)	0.002 (0.002, 0.002)	0.051 (0.049, 0.053)
30	0.104 (0.100, 0.108)	0.229 (0.223, 0.235)	0.499 (0.493, 0.505)	0.102 (0.098, 0.106)	0.066 (0.062, 0.070)
50	0.034 (0.032, 0.036)	0.162 (0.156, 0.168)	0.313 (0.307, 0.319)	0.417 (0.411, 0.423)	0.074 (0.070, 0.078)
150	0.000 (0.000, 0.000)	0.003 (0.000, 0.000)	0.010 (0.010, 0.010)	0.904 (0.900, 0.908)	0.083 (0.079, 0.087)

Figure 6-6: Observed proportion of babies born at 24 to 28 weeks gestational age receiving each of the levels of care or who have been discharged or died.

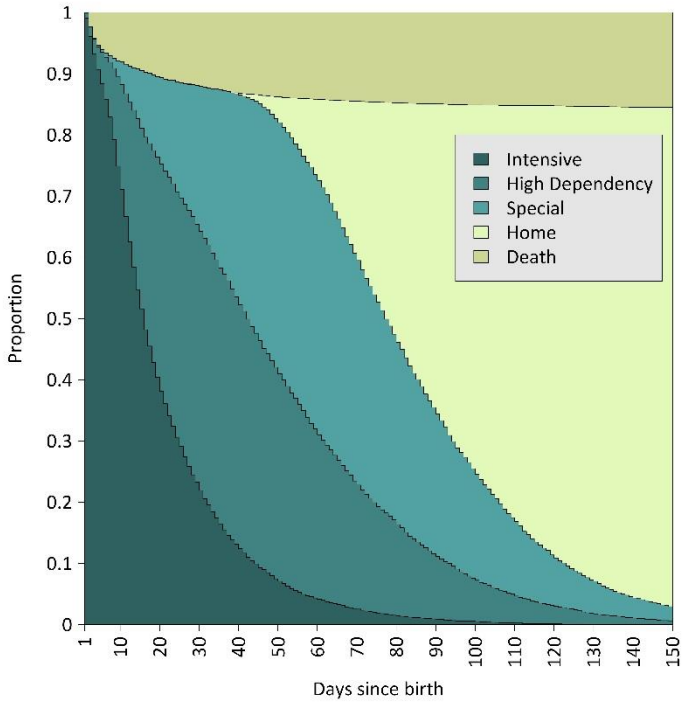


Figure 6-7: Observed proportion of babies born at 29 to 31 weeks gestational age receiving each of the levels of care or who have been discharged or died.

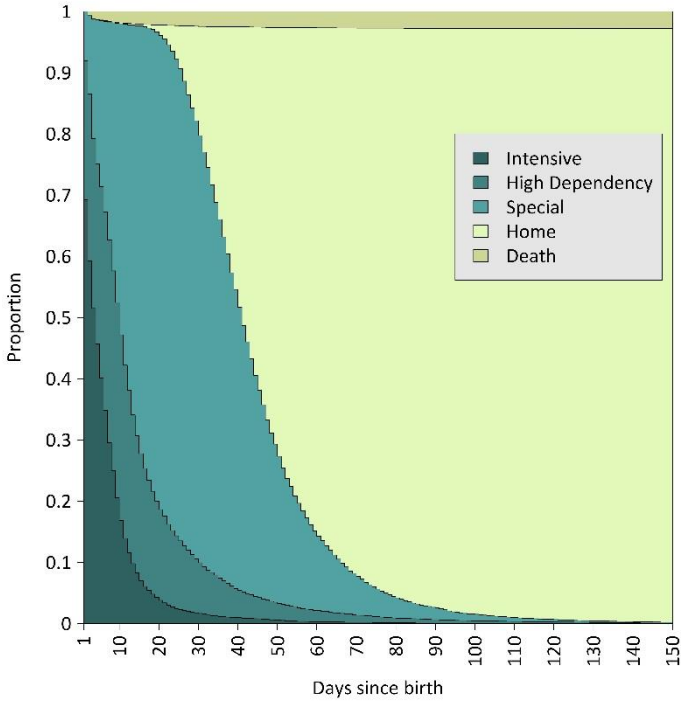
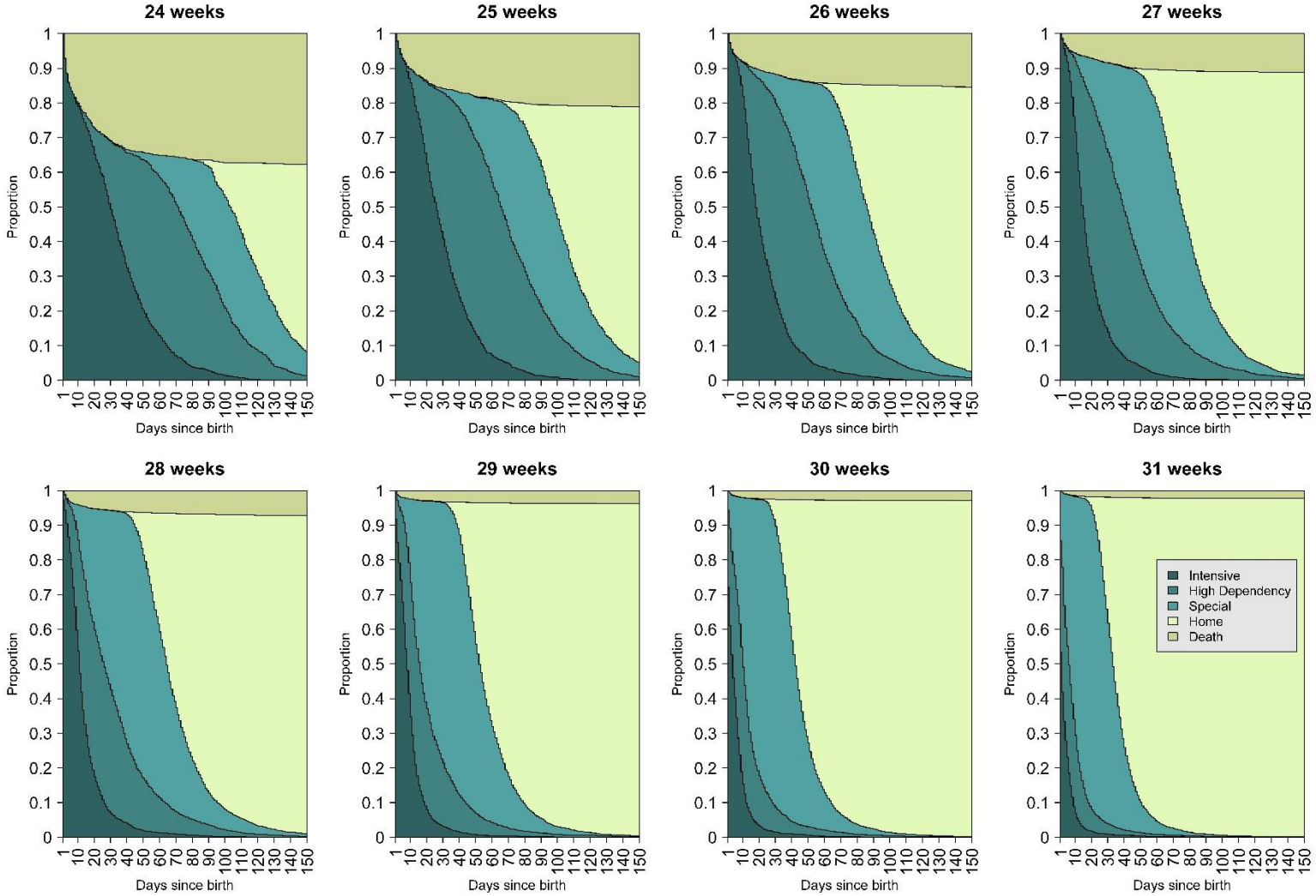


Figure 6-8: Observed proportion of babies requiring each level of care and those who have died or been discharged over time by gestational age.



6.5 COX PROPORTIONAL HAZARDS MULTISTATE MODELLING

Several approaches can be used to estimate multistate models and one approach is an extension of the Cox model (76, 96, 121). This approach estimates the probability of being in any state over time via use of a Aalen-Johansen-type estimator (101) as outlined in Chapter 6.3.2. As described previously in Chapter 4 and Chapter 5, in its simplest form the Cox model assumes proportional hazards. In this multistate model each transition is a separate stratum and so proportional hazards are assumed to hold within each transition. Therefore when proportional hazards are assessed, this should be investigated within each transition (85).

The transition-specific Cox model is defined as:

$$\alpha_{lj}(t|\mathbf{Z}) = \alpha_{l,j,0}(t)\exp(\boldsymbol{\beta}'_{lj}\mathbf{Z}_{\boldsymbol{lj}})$$

Equation 6-4

where $\boldsymbol{\beta}_{lj}$is the vector of regression coefficients and $\mathbf{Z}_{lj}$ are the transition-specific covariates that impact on the transition $l \rightarrow j$ (76, 78).

If all the regression coefficients are allowed to differ between transitions then this model is equivalent to fitting a separate model to each stratum (in this analysis each transition) with respect to the estimation of the hazards. However, there are advantages in the use of one model, for example post-estimation of transition probabilities as multiple models would give probabilities totalling more than one (96). It is possible to allow the covariates to be different for each transition, or only adjust certain transitions for certain covariates; these are known as **transition-specific covariates** (76). If covariates are assumed to have the same effect on all transitions then they are known as **shared covariates**.

6.6 COX MULTISTATE MODEL ADJUSTED FOR GESTATIONAL AGE

In order to estimate the levels of care required over time, and the outcomes of death or discharge from neonatal care, I fitted a Cox multistate model as described in Figure

6-1 with an adjustment for gestational age. All analysis in this chapter was conducted using the *mstate* command in R 3.0.2.

Hazard ratios are presented comparing each week of gestational age for each transition to the baseline group of 27 weeks gestational age (Table 6-4). The hazard of dying following intensive care is 3.03 times higher in the 24 week group compared to the 27 week group ($p<0.001$,Table 6-4). However, the hazard of dying after intensive care for babies born at 28 weeks is 30% lower than for those born at 27 weeks ($p<0.001$,Table 6-4).

Certain hazard ratios are poorly estimated due to a lack of data, for example the transition directly from intensive care to special care as seen by the large hazard ratios for 30 and 31 weeks, which are 26.7 and 53.6 respectively. This is because very few babies experience this transition and even fewer in those gestational age groups, mainly because these babies are likely to have received little or no intensive care. Therefore, these hazard ratios should not be over-interpreted. This can be seen in all gestational age groups for this transition, indicated by the wide confidence intervals, and thus most uncertainty. Similar levels of uncertainty can be seen in other transitions which are also uncommon, for example special care to death.

Table 6-4: Hazard ratios of each transition compared to 27 weeks gestational age with 95% confidence interval and p-value.[7]

Gestational age (weeks) **Transition**	**Hazard ratio**	**95% Confidence Interval**	**p-value**
24 weeks			
IC -> HD	0.38	0.35, 0.41	<0.001
IC -> SC	0.65	0.26, 1.62	0.36
IC -> Death	3.03	2.51, 3.65	<0.001
HD -> SC	0.51	0.47, 0.56	<0.001
HD -> Death	1.85	1.21, 2.82	0.005
SC -> Home	0.58	0.53, 0.64	<0.001
SC -> Death	4.51	2.15, 9.45	<0.001
25 weeks			
IC -> HD	0.52	0.48, 0.56	<0.001
IC -> SC	0.25	0.07, 0.88	0.03
IC -> Death	1.63	1.33, 1.99	<0.001
HD -> SC	0.61	0.57, 0.66	<0.001
HD -> Death	1.05	0.68, 1.61	0.84
SC -> Home	0.69	0.64, 0.75	<0.001
SC -> Death	3.40	1.73, 6.71	<0.001
26 weeks			
IC -> HD	0.75	0.70, 0.80	<0.001
IC -> SC	0.52	0.23, 1.32	0.18
IC -> Death	1.25	1.02, 1.53	0.033
HD -> SC	0.80	0.75, 0.86	<0.001
HD -> Death	0.88	0.58, 1.34	0.55
SC -> Home	0.80	0.75, 0.86	<0.001
SC -> Death	2.96	1.68, 5.21	<0.001
27 weeks			
IC -> HD	Reference	Reference	Reference
IC -> SC	Reference	Reference	Reference
IC -> Death	Reference	Reference	Reference
HD -> SC	Reference	Reference	Reference
HD -> Death	Reference	Reference	Reference
SC -> Home	Reference	Reference	Reference
SC -> Death	Reference	Reference	Reference
28 weeks			
IC -> HD	1.38	1.30, 1.47	<0.001
IC -> SC	3.95	2.21, 7.07	<0.001
IC -> Death	0.71	0.57, 0.87	0.003
HD -> SC	1.40	1.32, 1.48	<0.001

[7] In this table the following acronyms apply: IC intensive care; HD high dependency; SC special care.

Gestational age (weeks) Transition	Hazard ratio	95% Confidence Interval	p-value
HD -> Death	0.79	0.51, 1.22	0.29
SC -> Home	1.25	1.18, 1.33	<0.001
SC -> Death	0.59	0.31, 1.04	0.07
29 weeks			
IC -> HD	2.04	1.92, 2.17	<0.001
IC -> SC	9.93	5.71, 17.3	<0.001
IC -> Death	0.44	0.34, 0.57	<0.001
HD -> SC	1.97	1.85, 2.09	<0.001
HD -> Death	0.49	0.28, 0.86	0.012
SC -> Home	1.87	1.76, 1.98	<0.001
SC -> Death	0.27	0.15, 0.52	<0.001
30 weeks			
IC -> HD	2.86	2.69, 3.04	<0.001
IC -> SC	26.7	15.6, 45.9	<0.001
IC -> Death	0.55	0.42, 0.72	<0.001
HD -> SC	2.44	2.30, 2.60	<0.001
HD -> Death	0.26	0.12, 0.55	<0.001
SC -> Home	3.17	2.99, 3.36	<0.001
SC -> Death	0.16	0.08, 0.30	<0.001
31 weeks			
IC -> HD	3.48	3.27, 3.70	<0.001
IC -> SC	53.6	31.4, 91.6	<0.001
IC -> Death	0.64	0.49, 0.84	0.001
HD -> SC	3.10	2.91, 3.30	<0.001
HD -> Death	0.48	0.25, 0.95	0.034
SC -> Home	6.26	5.90, 6.63	<0.001
SC -> Death	0.10	0.05, 0.20	<0.001

6.6.1 PREDICTION OF PROBABILITIES

The estimates of the probability of being in each state over time are provided by week of gestational age in Figure 6-9. This may be more informative for the commissioning of care than the hazard ratios produced in Table 6-4. However, these probabilities are an over-simplification of the clinical care provided, where the ordering of care and past clinical experience of the baby will also be relevant. This will be discussed in Chapter 8.3.6. The stacked probability plots can be interpreted by considering the distance between the lines which represents the probability of being in that given state at that time. For example, for babies born at 24 weeks, the probability of needing intensive

care on the first day of life is one, and this reduces to 0.939 and then 0.897 on the second and third days of life. For these same babies, over the same period of time the probability of death increases from 0.061 (day two of life) to 0.103 (day three of life).

For the babies born at 24 weeks the risk of mortality is high with the probability of having died by 150 days of 0.38 (95% confidence interval: 0.345 to 0.415, Table 6-5). This compares to a probability of 0.024 (95% confidence interval: 0.020 to 0.028, Table 6-6) at the same time point for babies born at 31 weeks. The babies born at 24 weeks gestational age required higher levels of care, with the estimated probability needing intensive care at ten days being 0.706 (95% confidence interval: 0.688 to 0.724, Table 6-5) compared to 0.089 (95% confidence interval: 0.083 to 0.095, Table 6-6) for babies born at 31 weeks gestational age.

The predicted and observed proportions are similar for babies born at later gestational ages, indicating good model fit. However, the estimates show less agreement for babies born at 24 to 26 weeks (comparing Figure 6-8 with Figure 6-9). This is particularly apparent in the estimation of when babies are discharged from the neonatal unit, which is predicted to be earlier than would be expected for the extremely preterm babies. This is because the model is overwhelmed with data related to babies born at 30 and 31 weeks gestational age, and proportional hazards are assumed to hold across the weeks of gestational age. Alternative approaches will be discussed in Chapter 6.8 and Chapter 6.9 to overcome this issue.

When the transition probabilities are stacked, as in Figure 6-9, it is difficult to present the uncertainty around the estimates. Figure 6-10 and Figure 6-11 provide the probabilities again, with 95% confidence intervals to represent the uncertainty. The uncertainty in the babies born at 31 weeks is less than the uncertainty for those born at 24 weeks because there is more data available for babies born at 31 weeks of gestational age.

Figure 6-9: Stacked estimated proportions of babies receiving each level of care, or who have died or been discharged, over time.

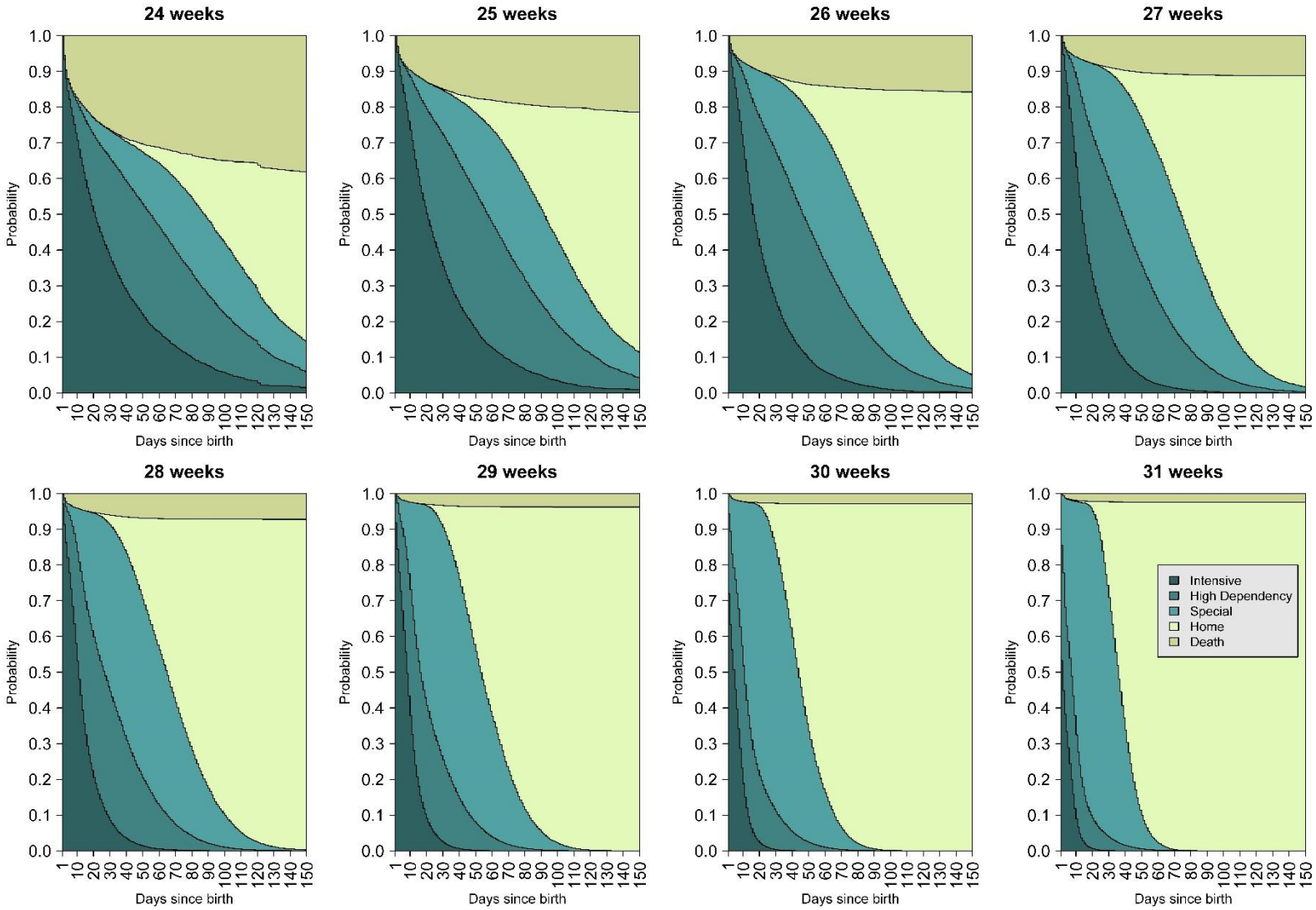


Table 6-5: Estimated proportions (95% confidence interval) of babies born at 24 week gestational age receiving each level of care or who have died or been discharged.

Day after birth	Intensive care	High dependency	Special care	Discharged	Died
1	1.000 (1.00, 1.00)	0.000 (0.000, 0.00)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)
2	0.932 (0.924, 0.940)	0.012 (0.010, 0.014)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.055 (0.047, 0.063)
3	0.879 (0.867, 0.891)	0.023 (0.021, 0.025)	0.001 (0.001, 0.001)	0.000 (0.000, 0.000)	0.096 (0.084, 0.108)
4	0.845 (0.831, 0.859)	0.032 (0.030, 0.034)	0.002 (0.002, 0.002)	0.000 (0.000, 0.000)	0.120 (0.106, 0.134)
5	0.821 (0.805, 0.837)	0.043 (0.039, 0.047)	0.003 (0.003, 0.003)	0.000 (0.000, 0.000)	0.133 (0.117, 0.149)
6	0.797 (0.781, 0.813)	0.054 (0.050, 0.058)	0.004 (0.002, 0.006)	0.000 (0.000, 0.000)	0.145 (0.129, 0.161)
7	0.775 (0.759, 0.791)	0.066 (0.060, 0.072)	0.005 (0.005, 0.005)	0.000 (0.000, 0.000)	0.154 (0.138, 0.170)
10	0.706 (0.688, 0.724)	0.105 (0.097, 0.113)	0.013 (0.011, 0.015)	0.000 (0.000, 0.000)	0.176 (0.158, 0.194)
30	0.380 (0.360, 0.400)	0.276 (0.256, 0.296)	0.077 (0.063, 0.091)	0.002 (0.002, 0.002)	0.265 (0.241, 0.289)
50	0.218 (0.198, 0.238)	0.310 (0.286, 0.334)	0.146 (0.126, 0.166)	0.023 (0.019, 0.027)	0.303 (0.276, 0.330)
150	0.015 (0.005, 0.025)	0.044 (0.032, 0.056)	0.082 (0.066, 0.098)	0.477 (0.444, 0.510)	0.380 (0.345, 0.415)

Table 6-6: Estimated proportions (95% confidence interval) of babies born at 31 weeks gestational age receiving each level of care or who have died or been discharged.

Day after birth	Intensive care	High dependency	Special care	Discharged	Died
1	0.533 (0.521, 0.545)	0.324 (0.314, 0.334)	0.144 (0.136, 0.152)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)
2	0.449 (0.439, 0.459)	0.330 (0.320, 0.400)	0.216 (0.206, 0.226)	0.000 (0.000, 0.000)	0.006 (0.004, 0.008)
3	0.383 (0.373, 0.393)	0.300 (0.290, 0.310)	0.306 (0.296, 0.316)	0.001 (0.000, 0.003)	0.011 (0.009, 0.013)
4	0.334 (0.324, 0.344)	0.296 (0.286, 0.306)	0.356 (0.344, 0.368)	0.001 (0.000, 0.003)	0.014 (0.012, 0.016)
5	0.288 (0.278, 0.298)	0.298 (0.288, 0.308)	0.398 (0.386, 0.410)	0.002 (0.000, 0.004)	0.015 (0.013, 0.017)
6	0.244 (0.234, 0.254)	0.296 (0.286, 0.306)	0.443 (0.431, 0.455)	0.002 (0.000, 0.004)	0.016 (0.014, 0.018)
7	0.200 (0.192, 0.208)	0.290 (0.280, 0.300)	0.493 (0.481, 0.505)	0.002 (0.000, 0.004)	0.017 (0.015, 0.019)
10	0.089 (0.083, 0.095)	0.236 (0.226, 0.246)	0.654 (0.642, 0.666)	0.003 (0.001, 0.005)	0.019 (0.015, 0.023)
30	0.000 (0.000, 0.000)	0.038 (0.034, 0.042)	0.671 (0.659, 0.683)	0.269 (0.259, 0.279)	0.023 (0.019, 0.027)
50	0.000 (0.000, 0.000)	0.005 (0.003, 0.007)	0.094 (0.088, 0.100)	0.878 (0.870, 0.886)	0.024 (0.020, 0.028)
150	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.977 (0.973, 0.980)	0.024 (0.020, 0.028)

6.6.2 EXPECTED LENGTH OF STAY IN EACH LEVEL OF CARE

The expected length of stay in each of the transient states, i.e. the time spent receiving each level of care is calculated as described in Chapter 6.3.3 and provided for all babies in Table 6-7. The percentile confidence interval is estimated from 1000 bootstrap samples. It is not possible to estimate expected length of stay by outcome as the probability of being in a state is an amalgamation of all babies irrespective of their future transition.

The observed mean length of stay receiving each level of care is provided in Table 6-8 to allow a comparison between the observed length of stay and that estimated from the model (Table 6-7). The difference in the observed and predicted length of stay is a maximum of two days, and overall the predictions from the model reflect the observed data. Therefore, whilst there was concern over the estimation of the probabilities in Chapter 6.6.1, the length of stay estimates appear unaffected.

An alternative approach for considering the care needs of this population of babies is presented in Table 6-9 where an estimation is provided of the total number of days of care which may be required by these babies using a total number of babies similar to that seen in this population (see Chapter 3).

Table 6-7: Expected time (days) spent receiving each level of care with 95% confidence interval.

Gestational age (weeks)	Intensive care	High dependency	Special care	Total
24	33 (30, 34)	29 (27, 31)	22 (19, 23)	84 (80, 87)
25	30 (28, 31)	33 (31, 34)	27 (25, 28)	90 (87, 92)
26	23 (21, 24)	29 (27, 30)	30 (28, 31)	82 (81, 84)
27	18 (17, 19)	25 (23, 26)	31 (29, 32)	74 (73, 76)
28	14 (12, 14)	19 (17, 19)	33 (33, 34)	66 (65, 67)
29	10 (9, 10)	13 (11, 13)	33 (32, 34)	56 (55, 57)
30	6 (5, 6)	9 (7, 9)	31 (30, 32)	46 (45, 47)
31	4 (3, 4)	6 (5, 6)	28 (27, 28)	38 (37, 38)

Table 6-8: Observed mean length of stay (days) at each level of care and overall.

Gestational age (weeks)	Intensive care	High dependency care	Special care	Total
24	32	31	20	83
25	29	35	26	90
26	22	31	30	83
27	18	26	32	76
28	14	18	35	67
29	9	12	35	56
30	6	8	32	46
31	4	6	27	37

Table 6-9: Estimated days of care in one year for each level of care using a hypothetical number of babies similar to the observed data.

Gestational age (weeks)	Hypothetical number of babies	Intensive care	High dependency	Special care	Total
24	280	9,240	8,120	6,160	23,520
25	320	9,600	10,560	8,640	28,800
26	440	10,120	12,760	13,200	36,080
27	540	9,720	13,500	16,740	39,960
28	700	9,800	13,300	23,100	46,200
29	800	8,000	10,400	26,400	44,800
30	1,000	6,000	9,000	31,000	46,000
31	1,300	5,200	7,800	36,400	49,400
Total					315,640

Figure 6-10: Proportion of babies in each state over time with 95% confidence interval for a baby born at 24 weeks gestational age.

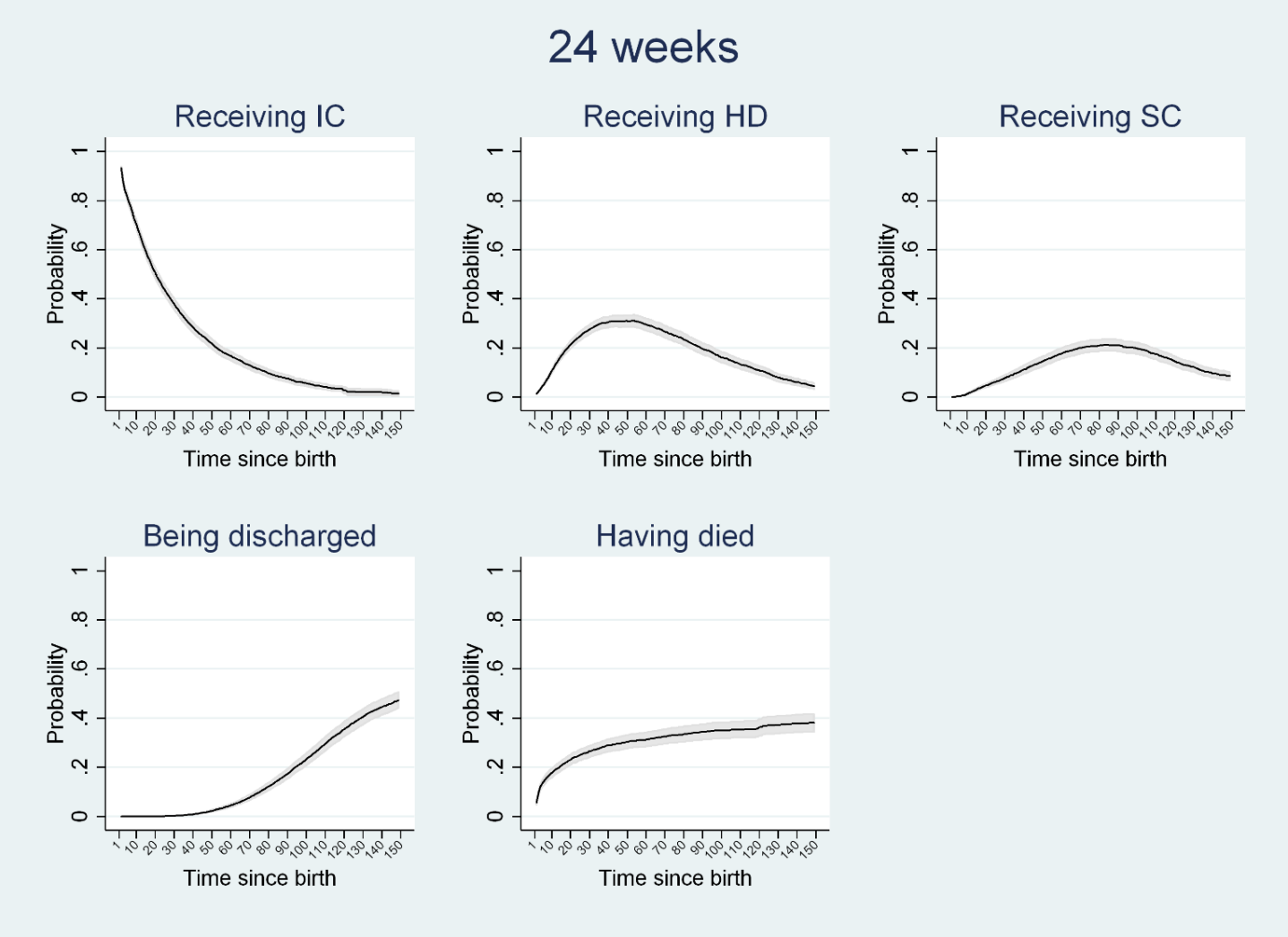
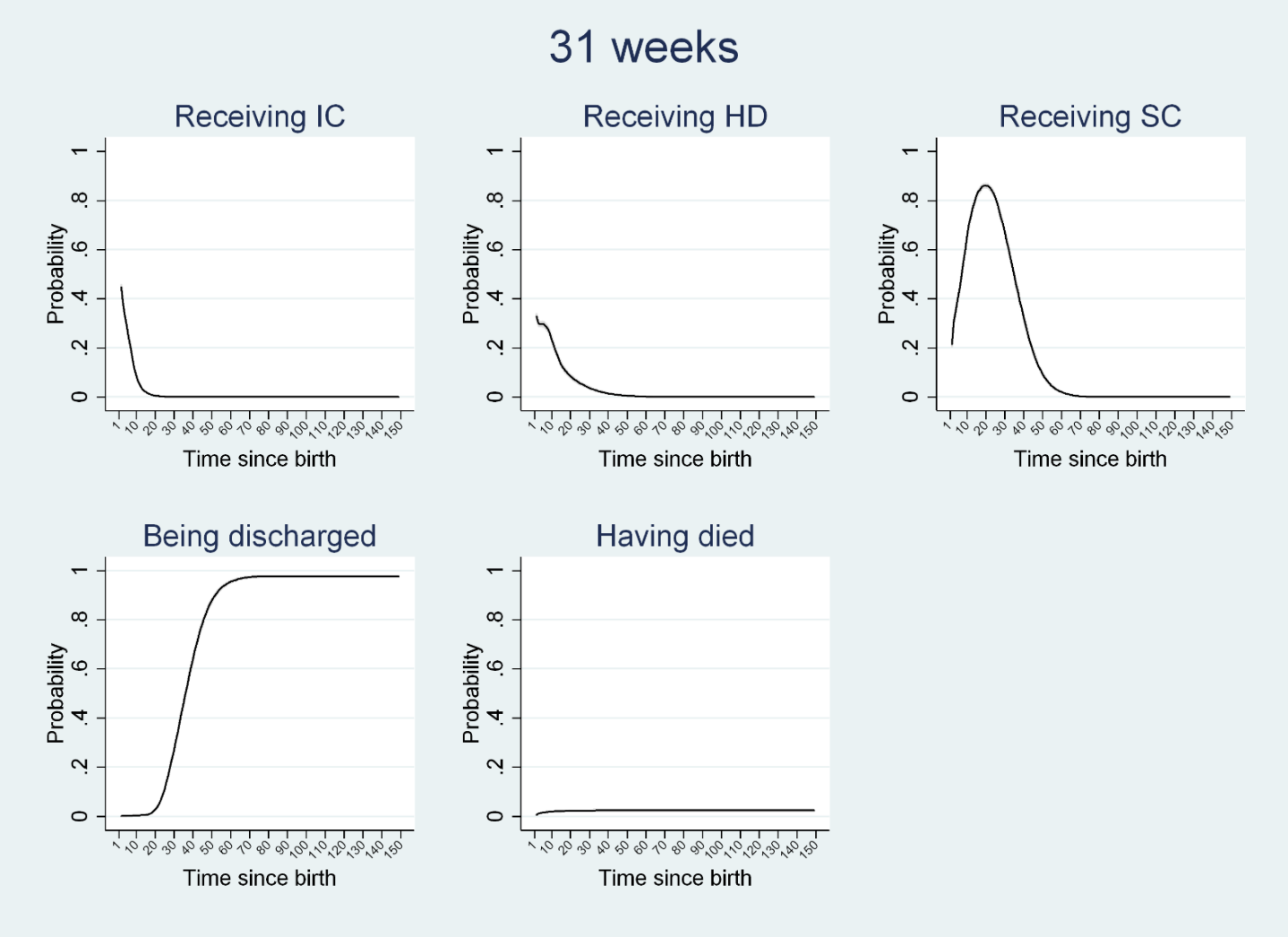


Figure 6-11: Proportion of babies in each state over time with 95% confidence interval for a baby born at 31 weeks gestational age.



6.6.3 TESTING MODEL ASSUMPTIONS

As proportional hazards were assumed within each transition the Therneau-Grambsch test was performed on each covariate within each transition. A p-value of <0.001 indicates a potential violation of the proportional hazards assumption (Table 6-10).

The transition identified as potentially breaching proportional hazards for most weeks of gestational age was the transition from intensive to high dependency care (Table 6-10). This was investigated more fully via use of Schoenfeld residuals (Figure 6-12) for those with p<0.001. Babies born at 30 and 31 weeks appeared to violate the proportional hazards assumption the most, indicated by a non-horizontal lines on the plots. The non-horizontal lines indicates that the hazards are changing over time.

Potential issues were also noted for the transitions of high dependency to special care and special care to home for the babies born at 30 and 31 weeks gestational age and Schoenfeld residual plots are provided in Figure 6-13 and Figure 6-14. The non-horizontal line on these plots indicates violation of the proportional hazards assumption. The transitions identified as breaching the proportional hazards assumption were also the transitions with the most data in the analyses. Stratification will be discussed in Chapter 6.8 to investigate these issues further as seen in Chapter 4.7.3.

Table 6-10: Therneau-Grambsch test for proportional hazards for each transition by week of gestational age.[8]

Gestational age (weeks) Transition	Chi-squared value	P-value
24 weeks		
IC -> HD	27.82	<0.001
IC -> SC	0.10	0.747
IC -> Death	5.24	0.02
HD -> SC	6.21	0.013
HD -> Death	3.04	0.081
SC -> Home	0.31	0.579
SC -> Death	10.25	0.001
25 weeks		
IC -> HD	14.30	<0.001
IC -> SC	1.30	0.255
IC -> Death	0.39	0.532
HD -> SC	6.80	0.009
HD -> Death	1.52	0.217
SC -> Home	0.25	0.618
SC -> Death	5.07	0.024
26 weeks		
IC -> HD	4.42	0.036
IC -> SC	0.00	0.983
IC -> Death	1.89	0.170
HD -> SC	4.72	0.030
HD -> Death	0.59	0.444
SC -> Home	0.11	0.737
SC -> Death	1.89	0.170
27 weeks		
IC -> HD	Baseline	Baseline
IC -> SC	Baseline	Baseline
IC -> Death	Baseline	Baseline
HD -> SC	Baseline	Baseline
HD -> Death	Baseline	Baseline
SC -> Home	Baseline	Baseline
SC -> Death	Baseline	Baseline
28 weeks		
IC -> HD	8.29	0.004
IC -> SC	1.70	0.193
IC -> Death	0.56	0.456
HD -> SC	4.53	0.033
HD -> Death	0.54	0.463

[8] In this table the following acronyms apply: IC intensive care; HD high dependency; SC special care.

Gestational age (weeks) **Transition**	**Chi-squared value**	**P-value**
SC -> Home	0.90	0.342
SC -> Death	0.47	0.491
29 weeks		
IC -> HD	23.92	<0.001
IC -> SC	2.52	0.113
IC -> Death	1.36	0.243
HD -> SC	10.22	0.001
HD -> Death	0.16	0.690
SC -> Home	4.43	0.035
SC -> Death	0.73	0.392
30 weeks		
IC -> HD	66.64	<0.001
IC -> SC	3.50	0.061
IC -> Death	0.08	0.781
HD -> SC	18.95	<0.001
HD -> Death	0.21	0.643
SC -> Home	16.95	<0.001
SC -> Death	1.34	0.245
31 weeks		
IC -> HD	104.19	<0.001
IC -> SC	4.82	0.028
IC -> Death	0.43	0.510
HD -> SC	26.22	<0.001
HD -> Death	2.73	0.10
SC -> Home	43.87	<0.001
SC -> Death	2.82	0.093

Figure 6-12: Schoenfeld residual plots for the transition of intensive care to high dependency for the weeks of gestational age indicated as potentially breaching the proportional hazards assumption.

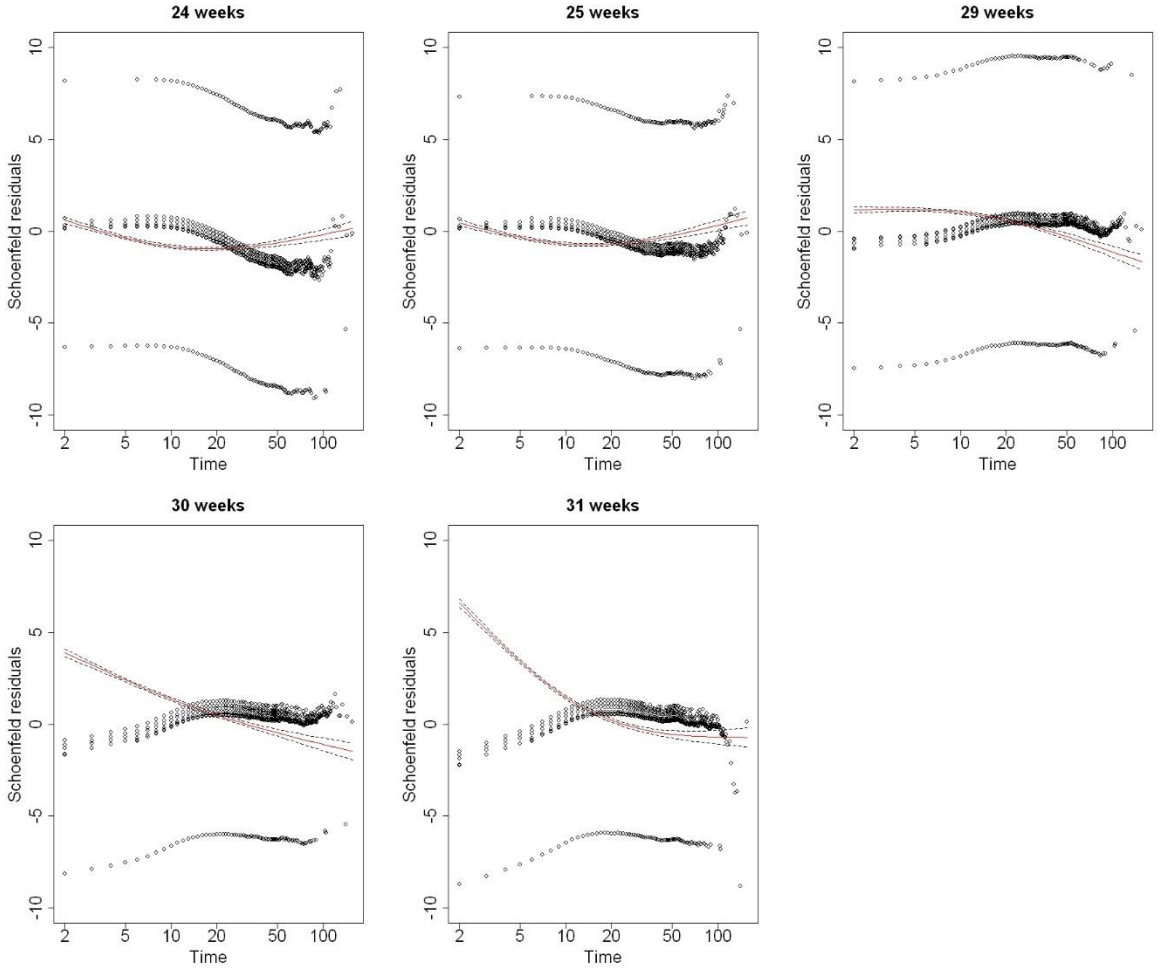


Figure 6-13: Schoenfeld residual plots for the transition of high dependency to special care for the weeks of gestational age indicated as potentially breaching the proportional hazards assumption.

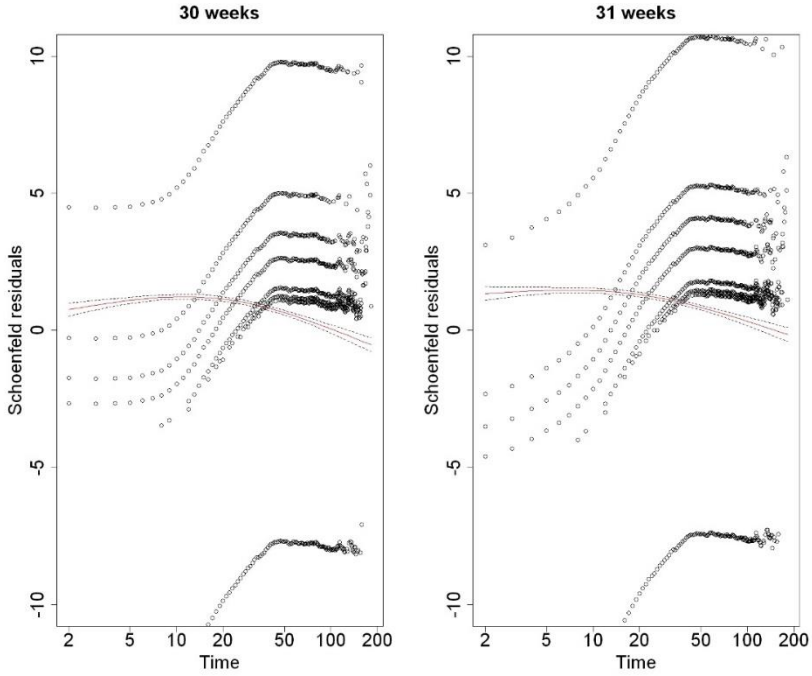
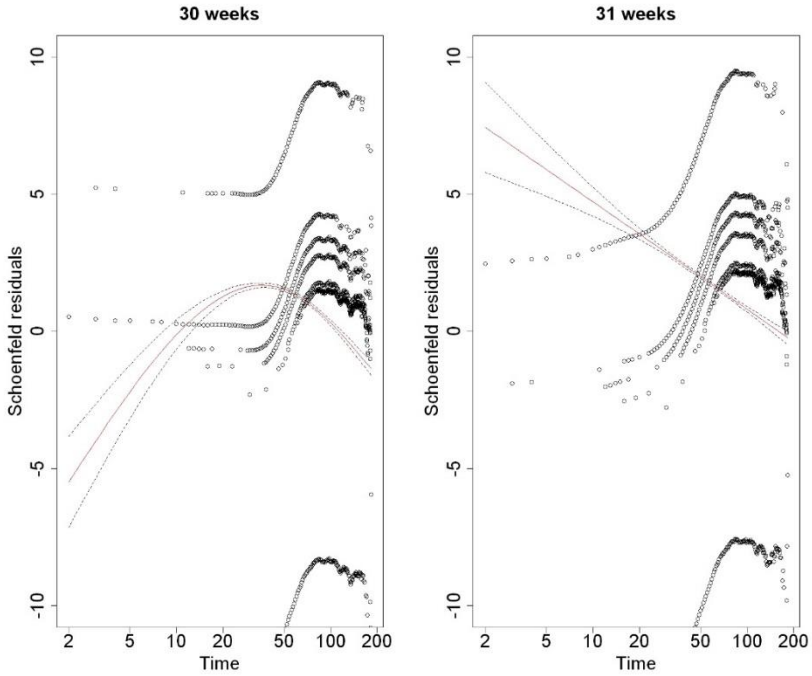


Figure 6-14: Schoenfeld residual plots for the transition of special care to home for the weeks of gestational age indicated as potentially breaching the proportional hazards assumption.



In addition to proportional hazards the Markov assumption was required to allow computation of the probabilities. To test if the assumption holds, Table 6-11 provides the hazard ratios for the entry times into a new state where a hazard ratio of approximately one indicates that the Markov assumption holds. A hazard ratio of more than one indicates that as entry time into the state increases, the hazard of making the next transition increases, whilst a hazard ratio of less than one indicates that as entry time into the state increases, the hazard of making the transition decreases. Two transitions had significant departures from a hazard ratio of one: special care after high dependency and discharged after special care although the hazard ratios remained close to one. Whilst the hazard ratios were still close to one, future work should consider relaxation of the Markov assumption for these transitions. The entry times with a hazard ratio that differed from one more substantially were the occurrence of death after special care and the occurrence of special care after intensive care. However, these transitions have the least events and therefore the hazard ratios slightly different from one do not provide evidence of a departure from the Markov assumption and use of the Markov assumption is therefore appropriate.

Table 6-11: Inclusion of entry time into the multistate model to test the Markov assumption.

Entry time into state	**Hazard ratio**	**95% confidence interval**
High dependency (after intensive care)	1.01	0.96, 1.09
Special care (after intensive care)	1.23	0.91, 1.65
Died (after intensive care)	1.04	0.82, 1.31
Special care (after high dependency)	0.97	0.96, 0.97
Died (after high dependency)	1.08	1.06, 1.09
Discharged (after special care)	0.97	0.96, 0.97
Died (after special care)	1.34	1.11, 1.16

6.6.4 SENSITIVITY ANALYSIS

The endpoint of discharge was an amalgamation of several outcomes as described in Chapter 3. To investigate the robustness of this assumption all babies discharged to another location (Table 3-4, n=734) were censored by the analysis. This provided a new transition matrix:

TO / **FROM**	INTENSIVE	HDU	SPECIAL	DEATH	DISCHARGE	NO EVENT	TOTAL
BIRTH	17,269	2,796	973				21,038
INTENSIVE	–	15,129	824	1,316	–	–	17,269
HDU	–	–	17,665	260	–	–	17,925
SPECIAL	–	–	–	186	18,542	734	19,462
DEATH	–	–	–	–	–	–	–
DISCHARGE	–	–	–	–	–	–	–
TOTAL	17,269	17,925	19,462	1,762	18,542	734	75,694

After censoring these observations, the hazard ratios (Table 6-12) and transition probabilities (Figure 6-15) were re-estimated. There were no differences in the results when compared with Table 6-4 and Figure 6-9 which indicated that amalgamating all of the outcomes together did not bias the results.

Table 6-12: Hazard ratios from sensitivity analysis with outcomes to other care locations censored.[9]

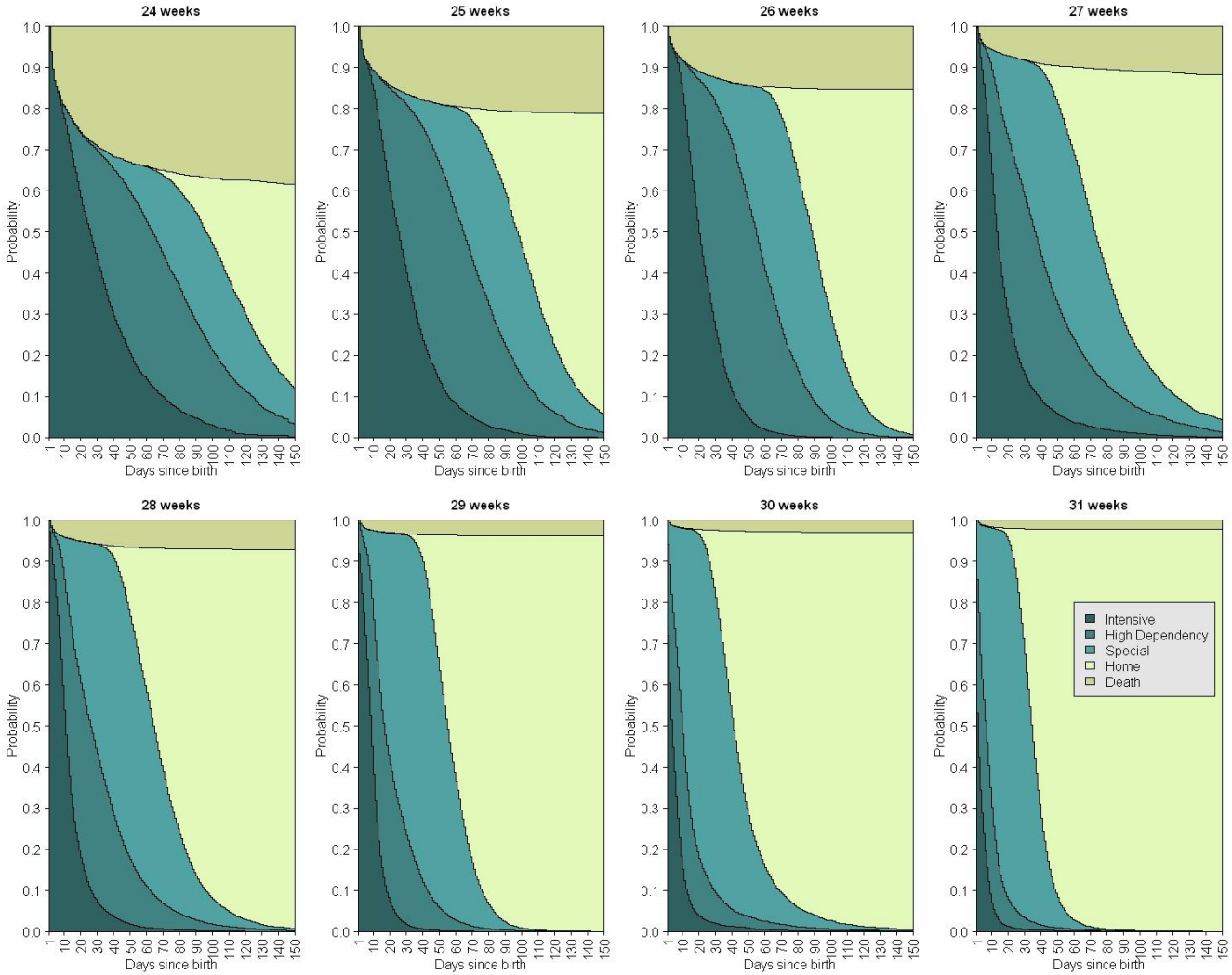
Gestational age (weeks) **Transition**	**Hazard ratio**	**95% Confidence Interval**	**p-value**
24 weeks			
IC -> HD	0.38	0.35, 0.41	<0.001
IC -> SC	0.65	0.26, 1.62	0.36
IC -> Death	3.03	2.51, 3.65	<0.001
HD -> SC	0.51	0.47, 0.56	<0.001
HD -> Death	1.85	1.21, 2.82	0.005
SC -> Home	0.59	0.54, 0.66	<0.001
SC -> Death	4.51	2.15, 9.45	<0.001
25 weeks			
IC -> HD	0.52	0.48, 0.56	<0.001
IC -> SC	0.25	0.07, 0.88	0.03
IC -> Death	1.63	1.33, 1.99	<0.001

[9] In this table the following acronyms apply: IC intensive care; HD high dependency; SC special care.

Gestational age (weeks) **Transition**	**Hazard ratio**	**95% Confidence Interval**	**p-value**
HD -> SC	0.61	0.57, 0.63	<0.001
HD -> Death	1.05	0.68, 1.61	0.84
SC -> Home	0.69	0.64, 0.75	<0.001
SC -> Death	3.40	1.73, 6.71	<0.001
26 weeks			
IC -> HD	0.75	0.70, 0.80	<0.001
IC -> SC	0.55	0.23, 1.32	0.18
IC -> Death	1.25	1.02, 1.53	0.036
HD -> SC	0.80	0.75, 0.86	<0.001
HD -> Death	0.88	0.58, 1.34	0.57
SC -> Home	0.80	0.75, 0.86	<0.001
SC -> Death	2.96	1.68, 5.21	<0.001
27 weeks			
IC -> HD	Reference	Reference	Reference
IC -> SC	Reference	Reference	Reference
IC -> Death	Reference	Reference	Reference
HD -> SC	Reference	Reference	Reference
HD -> Death	Reference	Reference	Reference
SC -> Home	Reference	Reference	Reference
SC -> Death	Reference	Reference	Reference
28 weeks			
IC -> HD	1.38	1.30, 1.47	<0.001
IC -> SC	3.95	2.21, 7.07	<0.001
IC -> Death	0.71	0.57, 0.86	0.003
HD -> SC	1.40	1.32, 1.48	<0.001
HD -> Death	0.79	0.51, 1.22	0.29
SC -> Home	1.26	1.18, 1.34	<0.001
SC -> Death	0.59	0.33, 1.04	0.07
29 weeks			
IC -> HD	2.04	1.92, 2.17	<0.001
IC -> SC	9.93	5.71, 17.3	<0.001
IC -> Death	0.44	0.34, 0.57	<0.001
HD -> SC	1.97	1.85, 2.09	<0.001
HD -> Death	0.49	0.28, 0.86	0.012
SC -> Home	1.87	1.76, 1.99	<0.001
SC -> Death	0.28	0.15, 0.52	<0.001
30 weeks			
IC -> HD	2.86	2.69, 3.04	<0.001
IC -> SC	26.7	15.6, 45.9	<0.001
IC -> Death	0.55	0.42, 0.72	<0.001
HD -> SC	2.45	2.30, 2.60	<0.001
HD -> Death	0.26	0.12, 0.55	<0.001
SC -> Home	3.19	3.00, 3.38	<0.001

Gestational age (weeks) **Transition**	**Hazard ratio**	**95% Confidence Interval**	**p-value**
SC -> Death	0.16	0.08, 0.30	<0.001
31 weeks			
IC -> HD	3.48	3.27, 3.70	<0.001
IC -> SC	53.6	31.4, 91.6	<0.001
IC -> Death	0.65	0.49, 0.84	0.001
HD -> SC	3.10	2.91, 3.30	<0.001
HD -> Death	0.48	0.25, 0.95	0.034
SC -> Home	6.29	5.90, 6.68	<0.001
SC -> Death	0.10	0.05, 0.20	<0.001

Figure 6-15: Proportion of babies receiving each level of care, or of having died or been discharged over time, estimated from the censored sensitivity analysis.



6.7 INVESTIGATING THE PROPORTIONAL HAZARDS ASSUMPTION

To investigate the impact of assuming proportional hazards on the results in Chapter 6.6, a sensitivity analysis was undertaken. One approach for handling non-proportional hazards, and to relax the assumption (89), is to perform an analysis that is stratified by groups that are more likely to be similar to each other, and therefore more likely to share proportional hazards. This was seen in the standard survival analysis in Chapter 4.7.3. However, at the time of writing this thesis it is not currently possible in multistate modelling software to stratify by more than one variable and the analysis has already been stratified for the different transitions. Therefore, separate models were fitted to the groups of babies that were similar to each other, according to the different length of stay groupings seen in Chapter 5.7.4. Gestational age was categorised into three groups of: 24 to 26 weeks (babies discharged around their due date); 27 to 29 weeks (babies discharged shortly before their due date) and 30 to 31 weeks (babies discharged in advance of their due date).

6.7.1 PREDICTION OF PROBABILITIES

The stacked probabilities presented in Figure 6-9 provided a poor fit when the probability of discharge was considered. This was a particular issue for babies born extremely preterm, noticeable for those born at 24 to 26 weeks gestational age when compared with the observed data (Figure 6-8). This is unsurprising as the extremely preterm babies were clearly different in their lengths of stay in the competing risks analysis (see Chapter 5).

Separate models were fitted and the predicted proportion of babies in each state over time was re-estimated (Figure 6-16). This improved the prediction of the proportions and these now reflected the observed data closer when comparing with Figure 6-8.

The predicted proportion of babies in each state is provided at certain time points for babies born at 24 weeks (Table 6-13) and 31 weeks (Table 6-14). These can be compared to the estimates from the single model (Table 6-5 and Table 6-6). In the separate analyses, there was a lower proportion of babies being discharged early in the time period after birth and this reflects the observed data more effectively than

seen previously. However, both approaches have a similar proportion of events by the latest time points. For example, of the babies born at 24 weeks who died, the single proportional hazards model estimated a proportion of 0.380 would die (Table 6-5) whilst the separate models estimated 0.384 (Table 6-13). Similarly, for babies born at 31 weeks, the estimated proportions of babies who survived to discharge from the two models were 0.977 (Table 6-6) and 0.978 (Table 6-14).

Table 6-13: Estimated proportions (95% confidence interval) for babies born at 24 weeks of receiving each level of care or who have died or been discharged from the three separate models.

Day after birth	Intensive care	High dependency	Special care	Discharged	Died
1	1.000 (1.00, 1.00)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)
2	0.939 (0.925, 0.953)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.061 (0.051, 0.071)
3	0.897 (0.883, 0.912)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.103 (0.089, 0.117)
4	0.868 (0.852, 0.884)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.131 (0.115, 0.147)
5	0.853 (0.835, 0.871)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.146 (0.128, 0.164)
6	0.836 (0.818, 0.854)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.157 (0.139, 0.175)
7	0.827 (0.809, 0.845)	0.006 (0.004, 0.008)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.168 (0.148, 0.188)
10	0.776 (0.756, 0.796)	0.029 (0.023, 0.035)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.194 (0.172, 0.216)
30	0.431 (0.407, 0.455)	0.265 (0.241, 0.289)	0.012 (0.004, 0.020)	0.000 (0.000, 0.000)	0.292 (0.263, 0.321)
50	0.204 (0.184, 0.224)	0.392 (0.359, 0.425)	0.073 (0.038, 0.108)	0.000 (0.000, 0.000)	0.330 (0.281, 0.379)
150	0.002 (0.000, 0.006)	0.032 (0.022, 0.042)	0.083 (0.056, 0.110)	0.499 (0.407, 0.591)	0.384 (0.258, 0.470)

Table 6-14: Estimated proportions (95% confidence interval) for babies born at 31 week babies of receiving each level of care or who have died or been discharged from the three separate models.

Day after birth	Intensive care	High dependency	Special care	Discharged	Died
1	0.533 (0.521, 0.545)	0.323 (0.311, 0.335)	0.144 (0.136,0.152)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)
2	0.424 (0.414, 0.434)	0.353 (0.343, 0.363)	0.218 (0.206, 0.230)	0.000 (0.000, 0.000)	0.005 (0.003, 0.007)
3	0.345 (0.335, 0.355)	0.328 (0.316, 0.340)	0.318 (0.306, 330)	0.000 (0.000, 0.000)	0.010 (0.008, 0.012)
4	0.287 (0.277, 0.297)	0.326 (0.314, 0.338)	0.375 (0.363, 0.387)	0.001 (0.000, 0.003)	0.011 (0.009, 0.013)
5	0.240 (0.230, 0.250)	0.325 (0.313, 0.337)	0.422 (0.410, 0.434)	0.002 (0.000, 0.004)	0.012 (0.010, 0.014)
6	0.195 (0.185, 0.205)	0.318 (0.306, 0.330)	0.473 (0.461, 0.485)	0.002 (0.000, 0.004)	0.013 (0.011, 0.015)
7	0.155 (0.147, 0.163)	0.304 (0.292, 0.316)	0.525 (0.513, 0.537)	0.002 (0.000, 0.004)	0.014 (0.012, 0.016)
10	0.070 (0.064, 0.076)	0.231 (0.221, 0.241)	0.680 (0.668, 0.692)	0.003 (0.001, 0.005)	0.016 (0.014, 0.018)
30	0.006 (0.004, 0.008)	0.027 (0.023, 0.031)	0.642 (0.630, 0.654)	0.305 (0.293, 0.317)	0.020 (0.016, 0.024)
50	0.020 (0.018, 0.022)	0.007 (0.005, 0.009)	0.074 (0.068, 0.080)	0.896 (0.888, 0.904)	0.021 (0.017, 0.025)
150	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.000 (0.000, 0.000)	0.978 (0.974, 0.982)	0.022 (0.018, 0.026)

6.7.2 EXPECTED LENGTH OF STAY

The expected length of stay was calculated for each level of care from the three separate models (Table 6-15). There was little difference in the estimated expected lengths of stay in each state between the two different approaches. This indicated that although the probabilities of discharge occurred early in the extreme preterm gestational ages in the analysis presented in Chapter 6.6, the calculation of the overall expected length of stay was unaffected.

Similarly, the estimated number of days in each state (Table 6-16) from the hypothetical total only differs by 1% from that calculated by the single model with gestational age in Table 6-9.

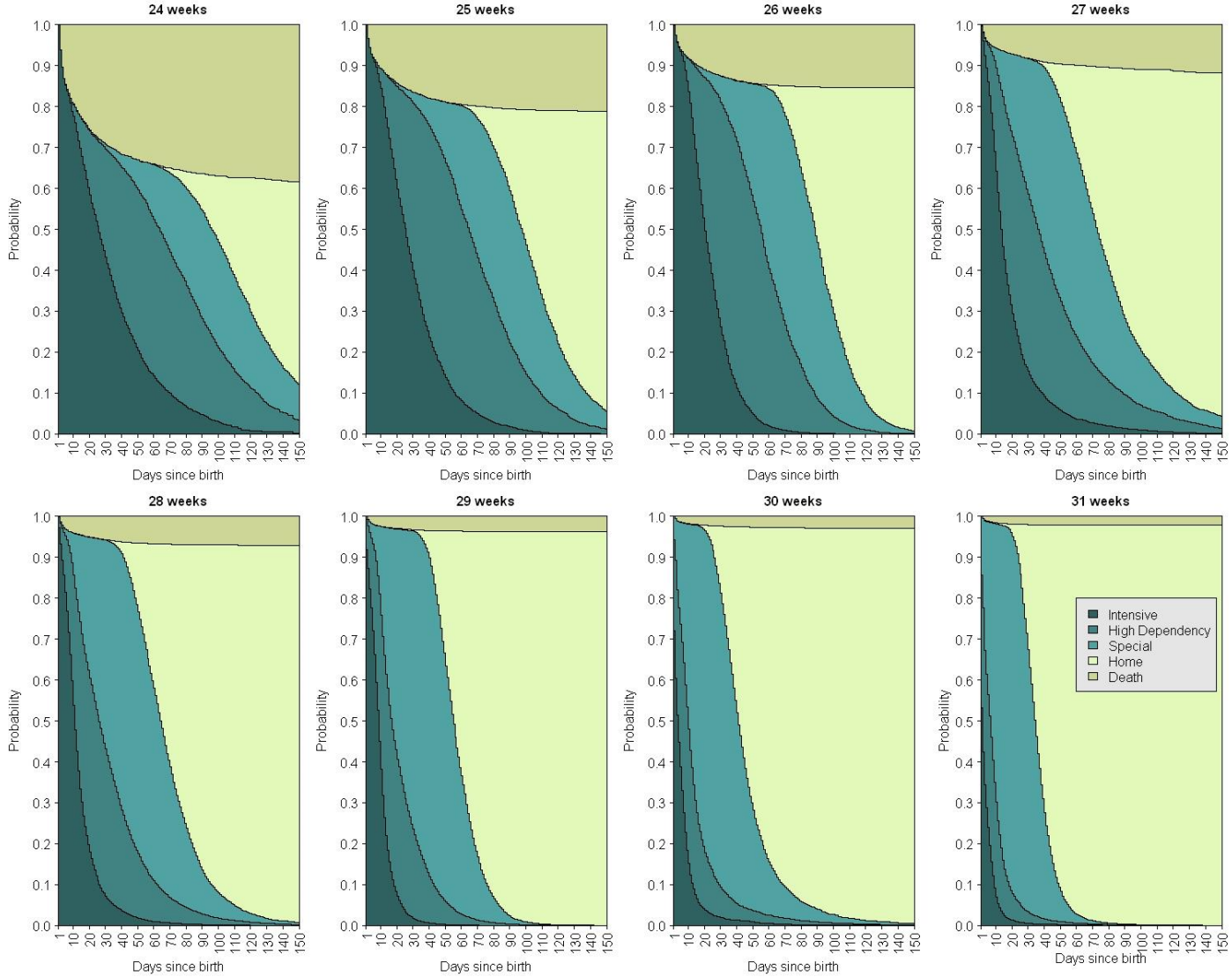
Table 6-15: Expected time (days) receiving each level of care, and total length of stay from the three separate models.

Gestational age (weeks)	Intensive care	High dependency	Special care	Total
24	32 (30, 34)	31 (28, 33)	21 (19, 22)	84 (81, 88)
25	29 (27, 31)	35 (24, 27)	26 (24, 27)	90 (88, 93)
26	23 (22, 24)	31 (29, 32)	29 (27, 30)	83 (82, 85)
27	19 (17, 20)	25 (23, 26)	33 (31, 34)	77 (75, 78)
28	14 (13, 14)	18 (17, 19)	35 (33, 35)	67 (66, 68)
29	10 (9, 10)	12 (11, 13)	35 (33, 35)	57 (56, 58)
30	6 (5, 6)	9 (7, 9)	32 (30, 32)	47 (45, 47)
31	4 (3, 4)	6 (5, 6)	27 (26, 28)	37 (37, 38)

Table 6-16: Estimated days of care for each level of care using a hypothetical number of babies similar to that observed in one year.

Gestational age (weeks)	**Hypothetical total**	**Intensive care (days)**	**High dependency (days)**	**Special care (days)**	**Number of care days (total)**
24	280	8,960	8,680	5,880	23,520
25	320	9,280	11,200	8,320	28,800
26	440	10,120	13,640	12,760	36,520
27	540	10,260	13,500	17,820	41,580
28	700	9,800	12,600	24,500	46,900
29	800	8,000	9,600	28,000	45,600
30	1,000	6,000	9,000	32,000	47,000
31	1,300	5,200	7,800	35,100	48,100
Total					318,020

Figure 6-16: Estimated proportion of babies in each state over time (separate models).



6.8 MULTISTATE COX MODEL ADJUSTED FOR GESTATIONAL AGE, BIRTHWEIGHT AND SEX

Whilst proportional hazards were indicated to be a problem in Chapter 6.6, a separate analysis in Chapter 6.7 indicated that the estimates of length of stay were largely unaffected. Additionally, fitting three separate models and adjusting them further would have led to sparse data relative to the number of parameters. Therefore, an analysis of all babies together assuming proportional hazards was undertaken with adjustments for gestational age, birthweight z-score and sex was fitted.

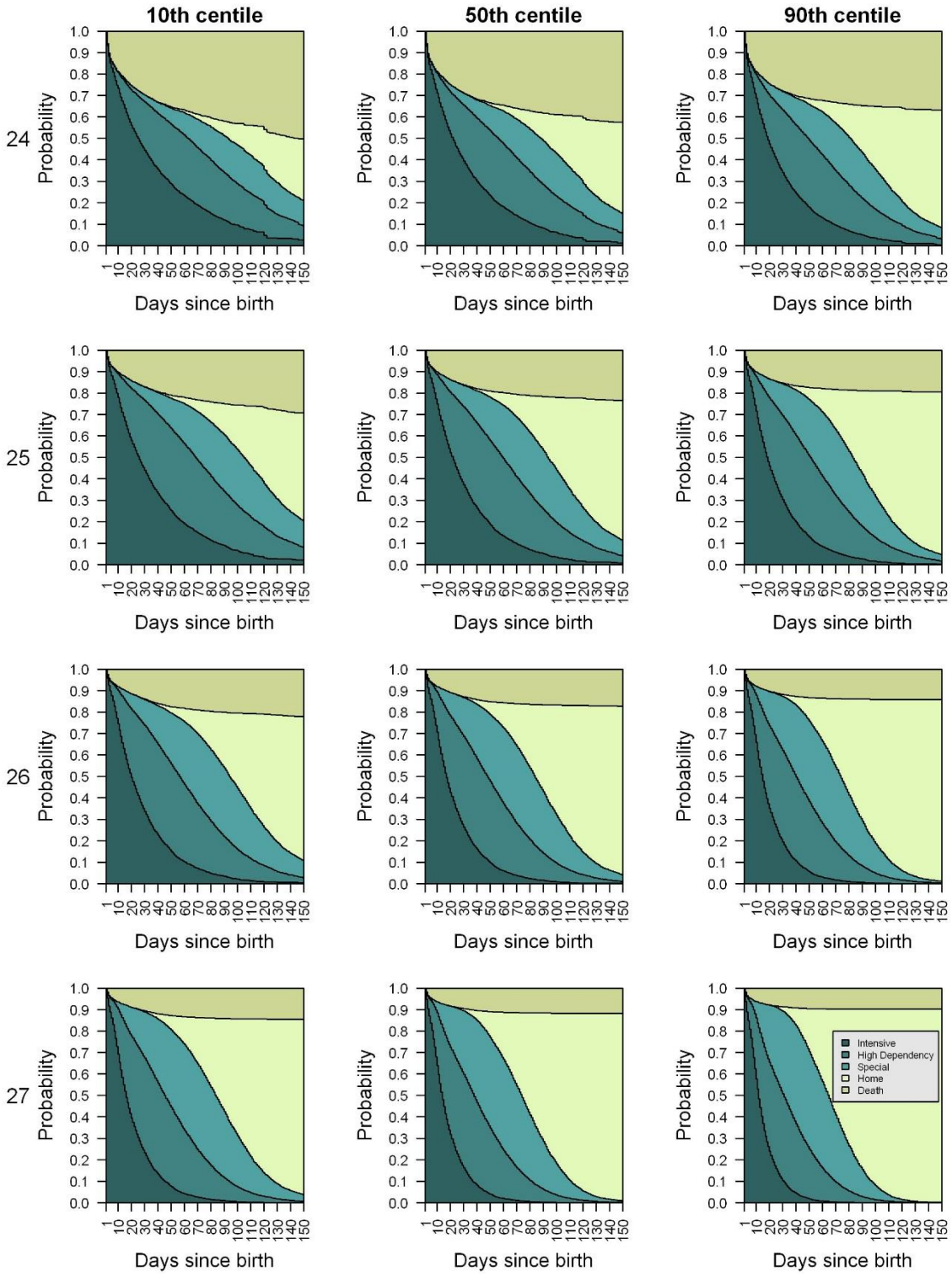
Birthweight z-score was modelled linearly and whilst it is unlikely that birthweight has a linear effect, a more complex approach was likely to result in overfitting. Overfitting would be a particular issue when the transitions were uncommon (e.g. special care to death) or the number of babies was small. This analysis included 20,900 babies.

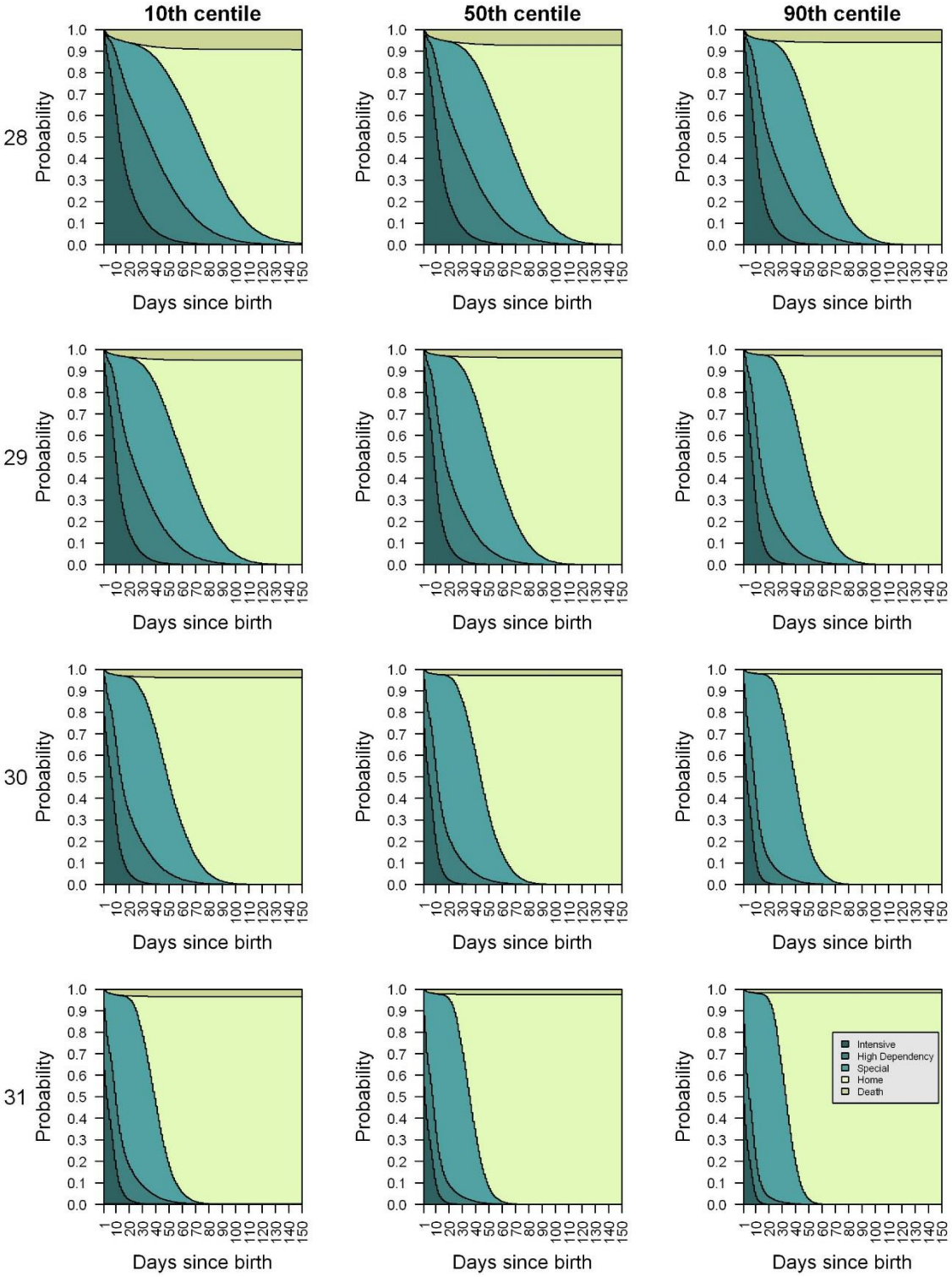
6.8.1 PREDICTION OF PROBABILITIES

The probability of being in each state over time was estimated and is provided in Figure 6-17 for male babies, by week of gestational age with birthweights at the 10th, 50th and 90th centile. These plots can be interpreted as before, where the distance between two lines represents the probability of being in that given state at that time.

As in Chapter 6.6, the probability of discharge in Figure 6-17 occurred too early in the days after birth for the most preterm babies. Whilst previously this impacted on the appearance of the probability plots, it did not appear to bias the estimation of the expected time babies received each level of care.

Figure 6-17: Stacked probabilities for male babies by week of gestational age for those with a birthweight at the 10th, 50th and 90th centiles.




10th centile
50th centile
90th centile
28
29
30
31
Probability
Days since birth
Intensive
High Dependency
Special
Home
Death

6.8.2 EXPECTED LENGTH OF STAY

The expected time spent receiving each level of care was estimated for babies of selected characteristics: female or male babies by week of gestational age and with a birthweight at the 10th, 50 and 90th centile (Table 6-17 and Table 6-18).

The expected times receiving each level of care for male babies born with a birthweight at the 50th centile were similar to that from the model that only accounted for gestational age (Table 6-7). This is because male babies comprise the largest proportion of neonatal admissions (Table 3-1: 55% versus 45%). Therefore, the estimates from the entire neonatal population are heavily influenced by the male babies and makes these estimates comparable with the estimates of the male babies.

As with the competing risks analysis (Chapter 5.8), the impact of sex on time receiving care was limited with similar results seen for babies of the opposite sex at the same birthweight. There was a larger difference in the expected time receiving each level of care between the different birthweight centiles. The difference in the total length of stay between the smallest (10th centile) and largest (90th centile) female babies ranged from 21 days (25 weeks, Table 6-17) to six days (31 weeks, Table 6-17). Confidence intervals are not provided for the expected length of stay as the number of babies at the extremes of birthweight is very small, and on individual bootstraps it is possible to select no data with the desired characteristics, leading to imprecise estimates, or models which fail to converge.

Table 6-17: Estimated expected time receiving each level of care for female babies by week of gestational age and with a birthweight at the 10th, 50th or 90th centile.[10]

Gestational age (weeks)	**Birthweight**											
	10th centile				**50th centile**				**90th centile**			
	IC	**HD**	**SC**	**Total**	**IC**	**HD**	**SC**	**Total**	**IC**	**HD**	**SC**	**Total**
24	38	32	23	93	32	31	24	87	27	29	23	79
25	34	37	30	101	28	34	29	91	23	30	27	80
26	25	33	34	92	21	29	32	82	18	25	29	72
27	19	28	35	82	16	24	32	72	14	20	30	64
28	15	21	36	72	12	17	34	63	10	14	32	56
29	11	15	36	62	9	12	34	55	7	9	32	48
30	7	11	33	51	5	8	32	45	4	6	30	40
31	5	7	28	40	3	5	28	36	3	4	27	34

[10] In this table the following acronyms apply: IC intensive care; HD high dependency; SC special care.

Table 6-18: Estimated expected length of stay for male babies by week of gestational age and with a birthweight at the 10th, 50th or 90th centile.[11]

Gestational age (weeks)	**Birthweight**											
	10th centile				**50th centile**				**90th centile**			
	IC	**HD**	**SC**	**Total**	**IC**	**HD**	**SC**	**Total**	**IC**	**HD**	**SC**	**Total**
24	39	29	19	87	34	29	21	84	29	27	21	77
25	37	36	26	99	31	33	26	90	26	30	25	81
26	28	33	31	92	23	29	29	81	20	25	27	72
27	22	28	32	82	18	25	30	73	15	21	29	65
28	16	22	34	72	14	18	32	64	11	15	31	57
29	12	16	34	62	10	13	33	56	8	10	31	49
30	8	12	32	52	6	9	31	46	5	7	30	42
31	5	8	27	40	4	6	27	37	3	4	27	34

[11] In this table the following acronyms apply: IC intensive care; HD high dependency; SC special care.

6.8.3 TESTING MODEL ASSUMPTIONS: PROPORTIONAL HAZARDS

Issues of proportional hazards were identified earlier in this chapter from the Therneau-Grambsch test and examination of Schoenfeld residuals (Chapter 6.6.3). To investigate these issues previously, three separate models were fitted to groups of babies believed to be similar to each other (Chapter 6.7). However, it was not possible to do this in this analysis as further adjustment would have resulted in overfitting due to the small number of observations compared to the number of parameters in the model.

The Therneau-Grambsch test indicated that some issues of proportional hazard remained in this adjusted model (results in Appendix 6). These issues remained noticeable between different gestational ages, particularly for the transition from intensive care to high dependency and high dependency to special care. There were no issues of proportional hazards between the two sexes.

Previously the impact of assuming proportional hazards seemed to affect the estimation of the probabilities and not the expected length of stay. However, future work should investigate the robustness of the results presented from this fully adjusted analysis.

6.9 DISCUSSION

This chapter has introduced an extension to competing risks methods which allowed consideration of intermediate events before the final outcomes of death or discharge. This approach is known as multistate modelling and allowed the time at each level of care to be estimated to address the third aim of this thesis.

A Cox proportional hazards multistate model was used to describe the probability of needing each level of care over time, and the expected number of days receiving each level of care, by gestational age.

Proportional hazards were assumed to hold between the different gestational age groups within each transition. Proportional hazards were investigated by the Therneau-Grambsch test and Schoenfeld residuals and was identified as an issue in

nine transitions. When comparing the obtained probabilities with the observed outcomes the shaded area in the stacked plot for the probability of discharge occurred too soon after birth, particularly for babies born at 24 and 25 weeks gestational age, as the model was heavily influenced by babies born at 30 and 31 weeks. To investigate and relax the impact of proportional hazards, three separate models were fitted dividing gestational age into: 24 to 26 weeks; 27 to 29 weeks and 30 to 31 weeks. When these separate models were fitted, the probability plots reflected those observed more accurately, although the expected length of stay did not seem to be substantially affected. The result of most interest from this work is the expected length of stay and this provides assurance that the results are robust to the proportional hazards assumption.

A final multistate model was presented, which was adjusted for key predictors of length of stay and mortality: gestational age; sex and birthweight z-score. It was not possible to relax the proportional hazards assumption in this model and future work should investigate this. However, as proportional hazards did not previously impact on the expected length of stay estimates, these results are likely to be unbiased.

The calculation of the expected length of time receiving each level of care is an average for all babies and it is not possible to separate the estimates by the future outcome of death or discharge. For babies where the proportion of deaths is small, particularly from 28 weeks onwards, the estimates of the total expected length of stay will be similar to the average length of stay for babies who survive.

6.9.1 IMPLICATIONS FOR PRACTICE: INVESTIGATING LEVELS OF CARE IN NEONATAL CARE

The length of stay in neonatal care was considered as the number of days spent receiving each level of care. This estimate was obtained by integrating the area under the curve and provided an average for all babies. For example, babies born at 24 weeks gestational age will require approximately 32 days receiving intensive care; 31 days receiving high dependency care and 21 days of special care. The expected total length of stay was 84 days which is lower than the predicted median length of stay of 122 days for 24 week babies who survive to discharge from the competing risks analysis in

Chapter 5. This difference occurs because the multistate analysis estimates an average of all babies, and because approximately 40% of babies born at 24 weeks will die, mostly in a short period of time, this will reduce the expected time. These results reflected that seen in the observed data (Table 6-8).

Whilst it is not possible to provide a breakdown of these results by final outcome, these estimates are informative for the commissioning of specialist neonatal services and the allocation of resources, as it is possible to estimate the number days of care needed for an average baby of certain characteristics.

The estimates of the expected (average) time receiving each level of care are useful for policy makers (124). A median time as presented previously in Chapter 5 is ideal for parental counselling. However, for commissioning and consideration of costs it disregards the effect of the extreme cases of length of stay and the impact these have on total resource use. Mean length of stay is more relevant for an economic decision maker in a budgeting situation as the total time in hospital can be estimated as the average time multiplied by the number of babies (125).

6.9.2 COMPARISON OF THIS RESEARCH WITH OTHER PUBLISHED RESEARCH

The British Association for Perinatal Medicine revised the standards for the levels of neonatal care in 2011 (22). Since 2011, there has been limited research which has investigated the levels of care with regards to length of stay. One study by Battersby et al reported observed length of stay by level of care in the term population (126). A simulation based study investigated the levels of care in the Peninsula network (Devon and Cornwall) (127). However, no study has investigated length of stay and levels of care for the very preterm population using multistate modelling.

6.9.3 STRENGTHS AND LIMITATIONS OF THIS ANALYSIS

This chapter has investigated how the different levels of care contribute towards time spent in neonatal care by very preterm babies using a multistate modelling approach. No previous research has investigated the levels of neonatal care within the context of estimating and understanding length of stay and these results provide a foundation for future work.

Proportional hazards were identified as a potential problem in this analysis, and whilst it was not possible to incorporate time-dependent effects into the analysis separate models were fitted to groups of babies who were similar in terms of their gestational age. Results from this approach indicated that the estimates of the expected time receiving each level of care were robust to issues introduced by assuming proportional hazards. Whilst fitting separate models is one approach for handling non-proportional hazards, it introduces a non-smooth function between the weeks of gestational age. A flexible parametric approach would allow a smooth function, which would allow for borrowing of strength between weeks of gestational age.

There are no formal tests to investigate the fit of the models presented in this chapter. However, efforts to investigate performance of the methods by comparing the observed and the predicted probabilities, and the observed and estimated time receiving each level of care were undertaken. These indicated that the results were robust, however future work should investigate formal approaches to investigate model fit.

6.10 CHAPTER CONCLUSION

This chapter has extended the methods of competing risks to consider intermediate events via use of multistate modelling. The intermediate events were the different levels of care which a baby may need whilst in the neonatal unit. A model which only accounted for gestational age was presented and the probability of being in each state over time and the expected time receiving each level of care was estimated. For example, for babies born at 24 weeks gestational age, the expected number of days spent receiving intensive care; high dependency care and special care was 33, 29 and 22 respectively. To investigate the impact of assuming proportional hazards, separate models were fitted to group similar gestational ages together. Whilst this had the potential to improve the estimation of the probabilities, the estimates of expected time receiving each level of care remained similar.

Finally, an analysis was undertaken which accounted for gestational age, birthweight z-score and sex of the baby and estimates of expected time receiving each level of care were produced for babies according to their gestational age; sex and birthweight.

7 MULTISTATE MODELLING TO COMPARE CARE BETWEEN OPERATIONAL DELIVERY NETWORKS

7.1 CHAPTER OVERVIEW

Differences in healthcare provision, or inequalities in care, across the country have become the focus of increased attention. Previous chapters have investigated approaches for estimating length of stay and the expected need for different levels of care nationally across England. In this chapter the differences that may exist between Operational Delivery Networks (ODNs) in England are considered. Firstly, the observed differences in the lengths of stay between ODNs are investigated. Secondly, the differences will be investigated formally, using a multistate model with an indicator term for the ODN, to examine how the expected length of stay differs and in which levels of care the differences occur.

7.2 INTRODUCTION

Specific units are equipped to provide focussed care to babies of particular characteristics. Babies born at less than 27 weeks gestational age should be cared for in Neonatal Network Units (intensive care units), with the equipment and staff to provide their clinical needs. No babies with these characteristics should be found in the lower level units immediately after birth, or they should be transferred soon after birth (128).

Babies are often transferred between neonatal units during their time in neonatal care and recent reports from the National Neonatal Audit Programme (NNAP) and Bliss suggest that at least 10% of babies are transferred whilst in neonatal care (3, 129). A comparison of units would not be appropriate due to the different populations, the impact of transfers and that the number of extremely preterm babies admitted to most units would be small, even on an annual basis.

The thirteen neonatal ODNs, which existed in 2013 in England, were used for this analysis. These ODNs cover large geographical regions and were established in April

2013 following recommendations by NHS England to sustain and develop clinical networks (130). Prior to this, managed clinical networks had existed as neonatal networks (26). ODNs were developed to focus patient pathways between different neonatal units over a potentially wide geographical area, allowing access to all the specialist resources which might be required by an individual baby.

The large geographical area covered by ODNs provides a population diverse in term of babies cared for, and care provided by neonatal units. Babies generally receive all their care within one ODN where practical (1) and so the ODN selected for analysis is the ODN of first admission. When considering an analysis at ODN level, the sample size in certain ODNs can be quite small and some transitions, for example intensive care to special care, may be poorly estimated. In an attempt to mitigate this somewhat, gestational age is modelled linearly, rather than categorically, in this chapter. ODNs are anonymised in this chapter.

7.2.1 INCLUSION AND EXCLUSION CRITERIA

The 21,038 singleton babies born at 24 to 31 weeks gestational age as described in Chapter 3 were included in this analysis. However, in this chapter babies were excluded if their ODN of birth was recorded as: an unknown location (n=56); born at home (n=22); born in transit (n=14) or born in a non-NHS facility (n=54).

These babies were included in other analyses in this thesis as they were all admitted to NNRD hospitals on the first day of life (Table 3-7). However, in this chapter they are excluded as their birth and initial care, including resuscitation and stabilisation if required, would have been managed in a location other than the ODN of first admission. No exclusions were made for missing data related to birthweight or sex as these variables were not used in this chapter and therefore 20,892 babies are included in this chapter.

7.2.2 MORTALITY RATE IN EACH ODN

The observed proportion of babies who died in each ODN, with 95% confidence intervals, are presented in Table 7-1 overall and for babies born at 24 weeks and 31 weeks gestational age. There was a significant difference in the overall proportion of

deaths between the different ODNs (p<0.001). ODN 5 had the largest number of admissions, with 133 babies born at 24 weeks gestational age and 704 babies born at 31 weeks gestational age. ODN 10 had the smallest number of babies born at 24 weeks (n=53) and 31 weeks gestational age (n=213). ODN 2 had a high proportion of deaths at 24 weeks but overall they were similar to the average (Table 7-1). ODN 10 had a high proportion of deaths overall and this is reflected in later analyses (Table 7-2). These ODNs are both smaller ODNs, so these results may reflect uncertainty in the estimation of death in the small ODNs. However, this chapter focusses on the hazard of discharge rather than the hazard of death.

Table 7-1: Observed proportion of babies (95% confidence interval) who died overall and of those born at 24 weeks or 31 weeks gestational age in each by ODN.

ODN	**Died overall**	**Died (Born at 24 weeks)**	**Died (Born at 31 weeks)**
ODN 1	0.08 (0.06, 0.09)	0.25 (0.17, 0.35)	0.03 (0.02, 0.06)
ODN 2	0.08 (0.06, 0.09)	0.50 (0.37, 0.63)	0.01 (0.00, 0.03)
ODN 3	0.07 (0.06, 0.07)	0.27 (0.20, 0.36)	0.04 (0.02, 0.06)
ODN 4	0.09 (0.08, 0.11)	0.35 (0.24, 0.36)	0.01 (0.00, 0.03)
ODN 5	0.08 (0.07, 0.09)	0.42 (0.34, 0.51)	0.02 (0.01, 0.03)
ODN 6	0.08 (0.07, 0.10)	0.36 (0.25, 0.48)	0.02 (0.01, 0.05)
ODN 7	0.08 (0.07, 0.10)	0.35 (0.25, 0.48)	0.02 (0.01, 0.05)
ODN 8	0.08 (0.07, 0.09)	0.42 (0.32, 0.53)	0.01 (0.00, 0.03)
ODN 9	0.07 (0.05, 0.08)	0.35 (0.24, 0.47)	0.01 (0.00, 0.03)
ODN 10	0.13 (0.11, 0.15)	0.53 (0.40, 0.66)	0.04 (0.02, 0.08)
ODN 11	0.08 (0.07, 0.09)	0.35 (0.26, 0.45)	0.02 (0.01, 0.04)
ODN 12	0.10 (0.09, 0.12)	0.40 (0.30, 0.50)	0.03 (0.02, 0.05)
ODN 13	0.09 (0.07, 0.10)	0.43 (0.34, 0.53)	0.03 (0.02, 0.04)
Overall	0.082 (0.08, 0.09)	0.38 (0.35, 0.41)	0.02 (0.02, 0.03)
Chi-squared p-value	<0.001	0.02	0.04

7.2.3 Observed differences in length of stay between ODNs

The observed mean length of stay for babies surviving to discharge ranged from 57 days to 63 days between the ODNs (Figure 7-1). ODNs 2 and 12 had the lowest mean length of stay and ODNs 3 and 13 had the longest mean length of stay. The mean length of stay of babies who died varied more between the ODNs (Figure 7-2) with the mean varying from ten to 31 days. This was because the number of deaths is much

lower than the number of discharges, and therefore there is likely to be more variation.

The estimates of average length of stay are presented by gestational age for each ODN for babies that survived to discharge (Figure 7-3) and those who died in neonatal care (Figure 7-4). For babies born at 24 weeks who survived to discharge their mean length of stay ranged from 115 days (Figure 7-3: ODN 11) to 131 days (Figure 7-3: ODN 3). For those babies born at 31 weeks, the median length of stay for those who survived ranged from 31 days (Figure 7-3: ODN 12) to 38 days (Figure 7-3: ODN 7).

Figure 7-1: Observed mean length of stay (95% confidence interval) for the babies surviving to discharge from neonatal care for each of the 13 ODNs.

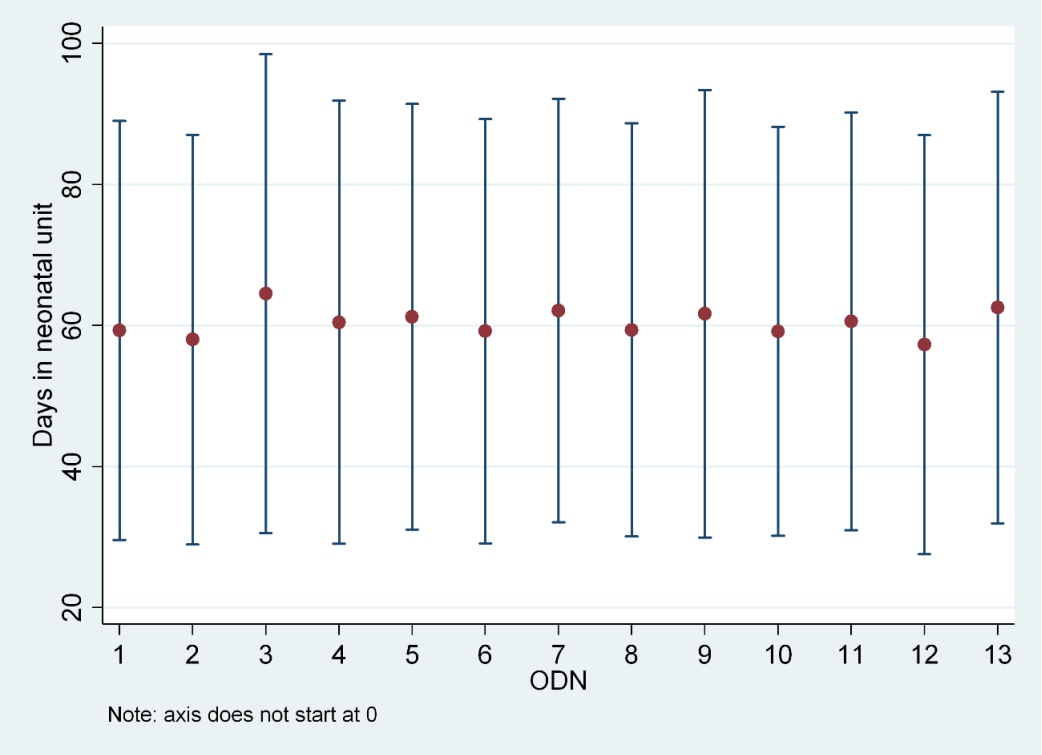


Figure 7-2: Observed mean length of stay (95% confidence interval) for the babies dying in neonatal care for each of the 13 ODNs.

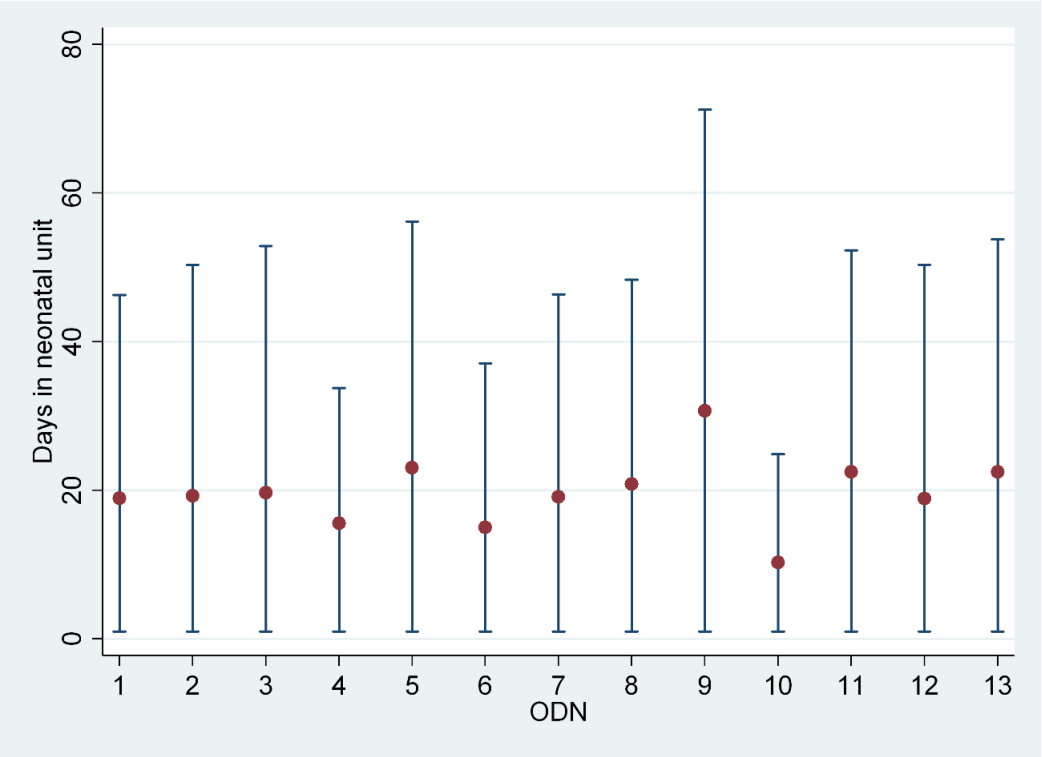


Figure 7-3: Mean length of stay in days for babies who survive to discharge (95% confidence interval) by week of gestational age for births in each of the 13 ODNs.

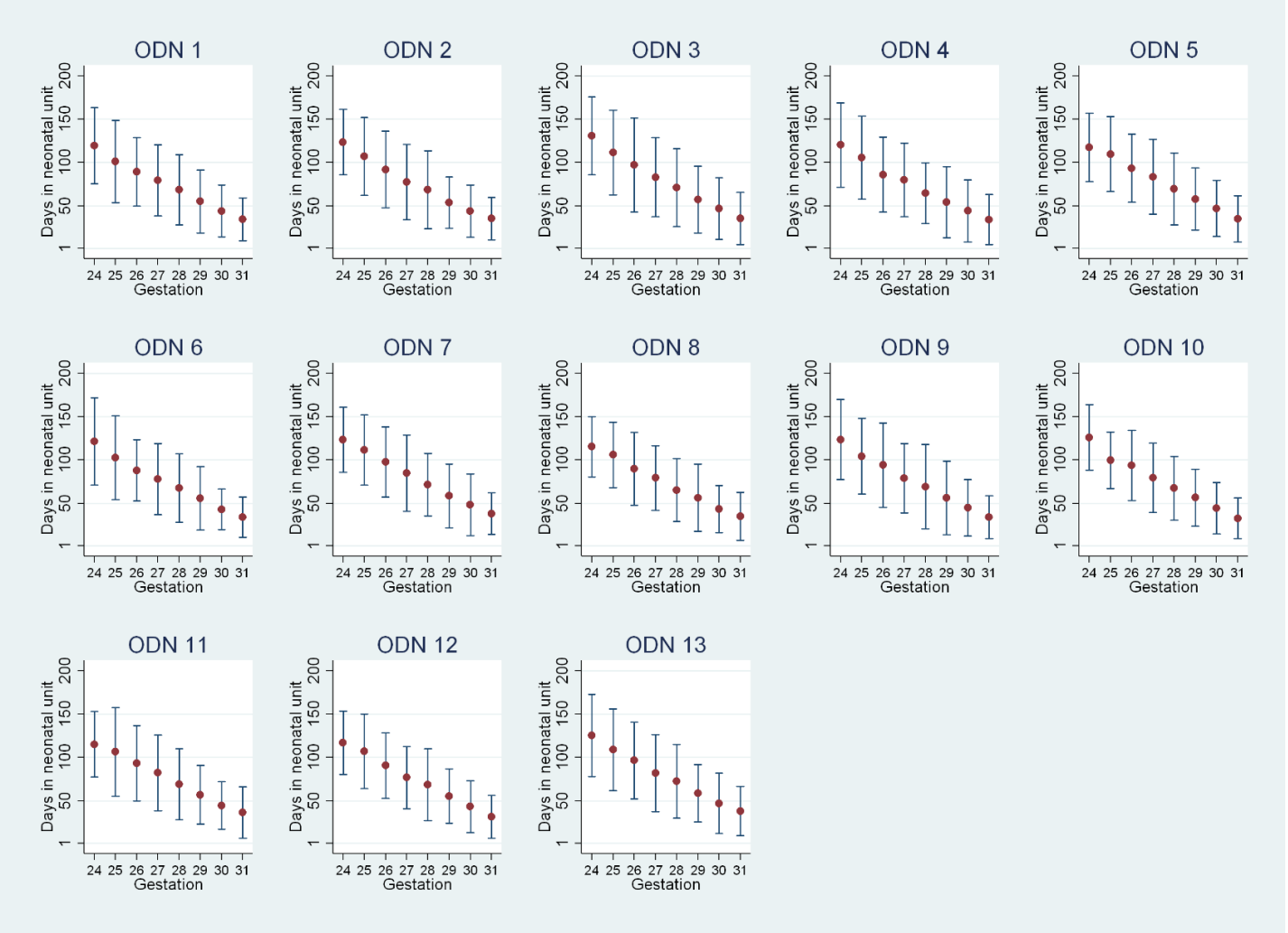
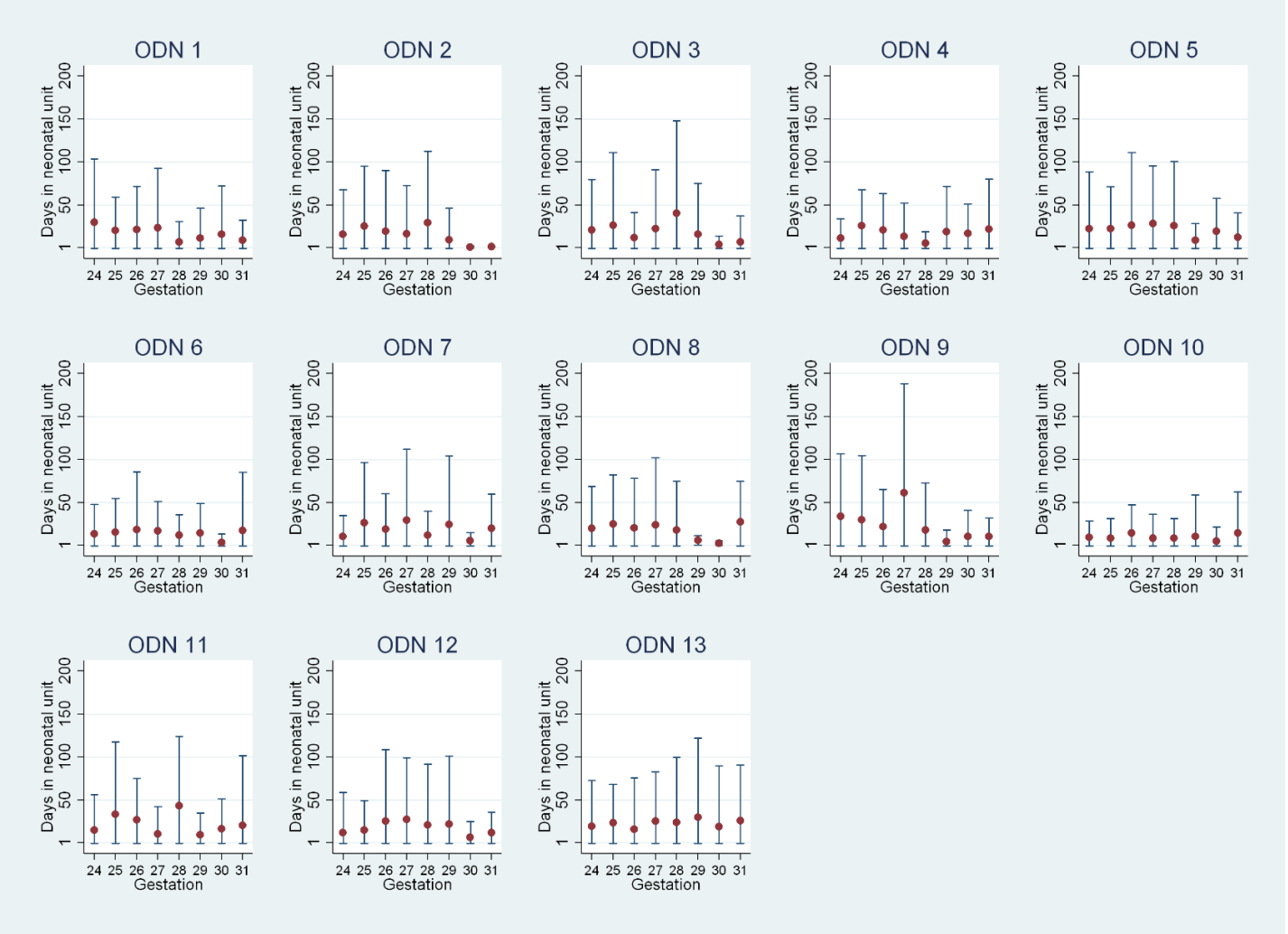


Figure 7-4: Mean length of stay of babies dying in neonatal care (95% confidence interval) by week of gestational age for births in each of the 13 ODNs.



7.2.4 DIFFERENCES IN ODNS FROM COMPETING RISKS ANALYSIS

To investigate differences in length of stay between the ODNs, I fitted a flexible parametric competing risks model with ODN as a covariate using *stpm2* in Stata v 14 using the methods described in Chapter 5 (110). The baseline group is the largest ODN: ODN 5.

The hazard ratios were estimated (Table 7-2) and denote the hazard of death or discharge compared to the baseline. ODN 3 had a significantly reduced hazard of discharge at all times, leading to a longer length of stay compared to ODN 5. ODN 12 had a significantly increased hazard of discharge, leading to a shorter length of stay at all times compared to ODN 5. These two ODNs and the reasons for these differences are investigated further in Chapter 7.3.1 and Chapter 7.3.2.

Only ODN 10 had an increased hazard of death compared to the baseline group of ODN 5 (Table 7-2) and this corresponded with the observed data (Table 7-1). However, this was not investigated further as the hazard of discharge in this ODN was not significant (p=0.03, threshold for significance in this thesis: 0.001).

7.3 CASE STUDIES OF DIFFERENCES IN LENGTH OF STAY

The differences between ODN 3 and ODN 12 were investigated via two multistate models comparing them with the rest of England. I fitted these multistate models using the *mstate* command in R 3.0.2 including a term to indicate if a baby was from the ODN of interest (ODN 3 or ODN 12) or the rest of England excluding that ODN. A linear term was also included for gestational age. Hazard ratios were estimated for each transition (Table 7-3 and Table 7-5).

The expected length of stay overall and for each of the levels of care was estimated (Table 7-4 and Table 7-6) with 95% confidence interval estimated from 1000 bootstrap samples. The probability plots of receiving each level of care or of having died or been discharged, were estimated for England excluding the ODN of interest and for ODN 3 and 12 (Appendix 7).

Table 7-2: Hazard ratios with 95% confidence intervals for discharge and death for each ODN.

	Hazard Ratio	**95% confidence interval**	**p-value**
Died			
ODN 1	0.91	0.73, 1.13	0.41
ODN 2	0.95	0.74, 1.20	0.65
ODN 3	0.78	0.63, 0.96	0.02
ODN 4	1.12	0.88, 1.41	0.36
ODN 5	Baseline	Baseline	Baseline
ODN 6	0.99	0.77, 1.26	0.92
ODN 7	1.00	0.80, 1.25	0.98
ODN 8	0.96	0.77, 1.20	0.74
ODN 9	0.78	0.61, 0.99	0.04
ODN 10	1.61	1.29, 2.00	<0.001
ODN 11	0.95	0.77, 1.17	0.63
ODN 12	1.29	1.07, 1.57	0.01
ODN 13	1.01	0.84, 1.22	0.92
Discharged			
ODN 1	1.07	1.01, 1.14	0.03
ODN 2	1.11	1.03, 1.19	0.003
ODN 3	0.90	0.75, 0.95	<0.001
ODN 4	1.03	0.95, 1.11	0.47
ODN 5	Baseline	Baseline	Baseline
ODN 6	1.07	1.00, 1.15	0.06
ODN 7	0.99	0.92, 1.06	0.70
ODN 8	1.07	1.00, 1.14	0.04
ODN 9	0.97	0.91, 1.04	0.36
ODN 10	1.09	1.01, 1.18	0.03
ODN 11	1.03	0.97, 1.09	0.38
ODN 12	1.12	1.05, 1.19	<0.001
ODN 13	0.96	0.91, 1.02	0.20

7.3.1 ODN 3: LONGER LENGTH OF STAY

Two transitions were significantly different in ODN 3 compared to the rest of England (Table 7-3). Firstly, the hazard of stepping down from intensive care to special care was reduced in ODN 3 compared to the rest of England (Table 7-3, HR: 0.57, 95% confidence interval: 0.42 to 0.77, p<0.001). However, the number of babies experiencing this transition was small, so this result should not be overemphasised.

Secondly, the hazard of the transition from high dependency to special care was reduced in ODN 3 compared to the rest of England, indicating that the time spent receiving high dependency in ODN 3 was likely to be longer than in the rest of England (Table 7-3, HR: 0.84, 95% confidence interval: 0.79 to 0.88, p<0.001).

Table 7-3: Hazard ratios for each of the transitions for those babies born in ODN 3 compared to England with 95% confidence intervals and p-value.[12]

Variable Transition	**Hazard Ratio**	**95% Confidence Interval**	**p-value**
ODN 3			
IC -> HD	0.95	0.90, 1.01	0.088
IC -> SC	0.57	0.42, 0.77	<0.001
IC -> Death	0.73	0.60, 0.90	0.003
HD -> SC	0.84	0.79, 0.88	<0.001
HD -> Death	0.50	0.30, 0.83	0.007
SC -> Home	0.96	0.91, 1.00	0.074
SC -> Death	0.89	0.54, 1.48	0.658
Gestational age (weeks)			
IC -> HD	1.38	1.37, 1.40	<0.001
IC -> SC	2.41	2.27, 2.55	<0.001
IC -> Death	0.76	0.74, 0.78	<0.001
HD -> SC	1.31	1.30, 1.32	<0.001
HD -> Death	0.82	0.75, 0.89	<0.001
SC -> Home	1.46	1.45, 1.48	<0.001
SC -> Death	0.55	0.49, 0.62	<0.001

The total expected length of stay for ODN 3 and the rest of England excluding ODN 3 is provided in Table 7-4. Babies born at 24 weeks were estimated to stay 13 days longer in ODN 3 compared to the rest of England, although part of this may relate to ODN 3 having a low mortality rate for babies born at 24 weeks gestational age (Table 7-1), which may increase the average length of stay.

The expected time receiving intensive care and special care was similar between ODN 3 and the rest of England for most gestational ages (Table 7-4). The largest differences were seen in the time spent receiving high dependency care where differences in

[12] In this table the following acronyms are used: intensive care (IC); high dependency care (HD) and special care (SC).

expected time ranged from one day (31 weeks) to seven days longer in ODN 3 (24 and 25 weeks) than the rest of England.

Table 7-4: Expected number of days (95% confidence interval) receiving each level of care in ODN 3 and England excluding ODN 3

Gestational age (weeks)	**Intensive care**	**High dependency**	**Special care**	**Total**
ODN 3				
24	38 (22, 49)	40 (22, 53)	31 (14, 45)	109 (61, 136)
25	31 (22, 37)	38 (29, 47)	33 (20, 46)	102 (78, 118)
26	24 (18, 29)	34 (29, 41)	32 (25, 44)	90 (78, 104)
27	19 (14, 21)	28 (24, 36)	31 (27, 42)	78 (68, 89)
28	14 (11, 15)	22 (19, 30)	30 (26, 39)	66 (59, 77)
29	10 (8, 11)	17 (13, 25)	29 (25, 39)	56 (49, 65)
30	7 (5, 8)	11 (8, 19)	29 (24, 39)	47 (41, 54)
31	4 (3, 5)	7 (5, 13)	30 (23, 37)	41 (34, 46)
England	**Intensive care *(Difference from ODN 3)***	**High dependency *(Difference from ODN 3)***	**Special care *(Difference from ODN 3)***	**Total *(Difference from ODN 3)***
24	34 (26, 43) *(-4)*	33 (25, 40) *(-7)*	29 (14, 39) *(-2)*	96 (73, 113) *(-13)*
25	28 (27, 35) *(-3)*	31 (23, 38) *(-7)*	32 (23, 38) *(-1)*	91 (80, 99) *(-11)*
26	22 (20, 26) *(-2)*	28 (26, 30) *(-6)*	32 (30, 37) *(0)*	82 (77, 86) *(-8)*
27	17 (16, 19) *(-2)*	23 (21, 25) *(-5)*	31 (30, 35) *(0)*	71 (69, 74) *(-7)*
28	13 (12, 14) *(-3)*	18 (16, 20) *(-4)*	30 (29, 33) *(0)*	61 (59, 63) *(-5)*
29	10 (8, 10) *(0)*	13 (12, 15) *(-4)*	30 (29, 32) *(+1)*	53 (50, 54) *(-3)*
30	7 (5, 7) *(0)*	9 (7, 10) *(-2)*	30 (28, 31) *(+1)*	46 (42, 46) *(-1)*
31	4 (3, 4) *(0)*	6 (4, 7) *(-1)*	30 (28, 32) *(0)*	40 (35, 40) *(-1)*

7.3.2 ODN 12: SHORTER LENGTH OF STAY

In the multistate model for ODN 12, three transitions were significantly different compared to the rest of England (Table 7-5). Two of these transitions, intensive care to special care and special care to death, were uncommon and results should not be overemphasised. However, the third transition, special care to home, indicated an increased hazard of discharge home occurring (HR: 1.16 95% CI: 1.11 to 1.23, p<0.001). Another result of borderline significance was the transition from high dependency to special care, indicating an increased hazard of care stepping down from high dependency to special care at all time points (HR: 1.08 95% CI: 1.02 to 1.16, p=0.002).

These increased hazard ratios indicate a shorter length of stay at those levels of care in this ODN.

Table 7-5: Hazard ratio for each of the transitions for those in ODN 12 compared to the rest of England and gestational age with 95% confidence interval and p-value.[13]

Variable Transition	Hazard Ratio	95% Confidence Interval	p-value
ODN 12			
IC -> HD	1.00	0.94, 1.06	0.991
IC -> SC	1.67	1.36, 2.05	<0.001
IC -> Death	1.25	1.04, 1.49	0.02
HD -> SC	1.08	1.02, 1.16	0.004
HD -> Death	1.07	0.68, 1.69	0.769
SC -> Home	1.16	1.11, 1.23	<0.001
SC -> Death	2.30	1.55, 3.40	<0.001
Gestational age (weeks)			
IC -> HD	1.38	1.37, 1.40	<0.001
IC -> SC	2.42	2.28, 2.56	<0.001
IC -> Death	0.76	0.74, 0.78	<0.001
HD -> SC	1.31	1.30, 1.32	<0.001
HD -> Death	0.81	0.75, 0.88	<0.001
SC -> Home	1.47	1.45, 1.48	<0.001
SC -> Death	0.55	0.49, 0.62	<0.001

The total expected length of stay for ODN 12 and the rest of England is provided in Table 7-6. The expected time receiving intensive care was similar between ODN 12 and the rest of England. Larger differences existed between ODN 12 and England for the days spent receiving high dependency and special care. Babies born at 24 weeks gestational age received high dependency care and special for four and five less days respectively in ODN 12 then the rest of England.

[13] In this table the following acronyms are used: intensive care (IC); high dependency care (HD) and special care (SC).

Table 7-6: Expected number of days (95% confidence interval) receiving each level of care in ODN 12 and England excluding ODN 12

Gestational age (weeks)	**Intensive care**	**High dependency**	**Special care**	**Total**
ODN 12				
24	33 (7, 61)	30 (4, 45)	25 (8, 44)	88 (19, 128)
25	27 (8, 46)	29 (7, 37)	28 (14, 41)	84 (29, 112)
26	22 (14, 36)	26 (8, 32)	29 (19, 39)	77 (43, 95)
27	17 (13, 27)	21 (9, 28)	28 (22, 37)	66 (53, 79)
28	13 (10, 20)	17 (8, 22)	28 (22, 35)	58 (50, 66)
29	10 (7, 15)	12 (6, 16)	28 (22, 33)	50 (44, 55)
30	7 (4, 11)	8 (4, 11)	29 (22, 32)	44 (38, 46)
31	4 (2, 7)	5 (2, 7)	30 (24, 32)	39 (32, 40)
England	**Intensive care (Difference from ODN 12)**	**High dependency (Difference from ODN 12)**	**Special care (Difference from ODN 12)**	**Total (Difference from ODN 12)**
24	35 (17, 42) *(+2)*	34 (22, 41) *(+4)*	30 (14, 39) *(+5)*	99 (61, 111) *(+11)*
25	29 (24, 34) *(+2)*	32 (25, 36) *(+3)*	32 (24, 37) *(+4)*	93 (80, 99) *(+9)*
26	23 (20, 26) *(+1)*	29 (25, 32) *(+3)*	32 (30, 36) *(+3)*	84 (77, 87) *(+7)*
27	18 (16, 19) *(+1)*	24 (22, 26) *(+3)*	31 (31, 34) *(+3)*	73 (69, 75) *(+7)*
28	13 (12, 14) *(0)*	19 (17, 20) *(+2)*	30 (30, 33) *(+2)*	62 (60, 64) *(+4)*
29	10 (9, 10) *(0)*	14 (12, 15) *(+2)*	30 (28, 32) *(+2)*	54 (51, 54) *(+4)*
30	7 (6, 7) *(0)*	9 (8, 11) *(+1)*	30 (28, 32) *(+1)*	46 (42, 46) *(+2)*
31	4 (3, 4) *(0)*	6 (4, 7) *(+1)*	30 (28, 32) *(0)*	40 (35, 40) *(+1)*

7.4 DISCUSSION

The analysis provided here demonstrates how multistate models can be used to compare different ODNs and investigate which levels of care, if any, contribute to the differences in total length of stay. Differences could be as a result of different approaches to healthcare delivery or inequalities in the health of the populations. The Marmot review identified the need to reduce health inequalities as a matter of fairness and social justice (131). Analyses such as that undertaken in this chapter can aid the understanding of differences in the provision of care.

7.4.1 DIFFERENCES IN LENGTH OF STAY AND TIME RECEIVING EACH LEVEL OF CARE

Overall, the observed mean length of stay varied from 57 to 63 days (Figure 7-1) although there was little variation between the ODNs. However, these estimates are heavily influenced by the babies who were born at 30 and 31 weeks gestational age.

Two ODNs were identified to investigate further because they had significantly longer or shorter lengths of stay: ODN 3 and ODN 12. Multistate models were fitted with an indicator term comparing ODN 3 or ODN 12 to the rest of England and gestational age was accounted for as a linear term.

ODN 3 and ODN 12 had a lower and higher mortality rate respectively than the rest of England (ODN 3: 0.067, ODN 12: 0.102 versus England: 0.083, Table 7-1). This may explain some of the differences seen as the estimation of expected length of stay is an average and an increased mortality rate would reduce the average length of stay as deaths often occur in the initial days after birth. A lower mortality rate as seen in ODN 3 could increase length of stay as more surviving babies are staying for longer periods of time in the neonatal unit. However, if the differences could be explained by differences in mortality alone it would be expected that there would be an increase or decrease in the expected number of days across all levels of care, or potentially an altered time receiving intensive care, which cares for the sickest babies.

In ODN 3 the increases in expected stay were not consistent across all the levels of care and were mainly seen in high dependency care. This indicates that this ODN had a different approach to the provision of high dependency care compared to the rest of England. Similarly, in ODN 12 less time was spent receiving special care, and this ODN may be different to the rest of England, for example they may have better community neonatal services and can discharge babies home earlier.

7.4.2 IMPACT OF DIFFERENCES IN LENGTH OF STAY

A shorter length of stay, such as in ODN 12, results in a reduced cost to neonatal services, but there is no evidence to suggest whether a shorter length of stay is beneficial or harmful in the longer term for a baby discharged from neonatal care (5).

Irrespective of this, if survival continues to improve in the future the average length of stay of all preterm babies and the workload required will increase (18, 132).

Identification of ODNs with apparently short lengths of stay should not necessarily indicate that their approach is best practice. Other researchers have also noted this: *"we presume a shorter LOS [length of stay] is a preferred outcome, this may not always be a true assumption as later outcomes are unknown…"* (5) Even from a cost perspective a short length of stay may not ultimately be cheaper for the healthcare service as a whole, as other services such as domiciliary or community care are needed, leading to a cost to the healthcare service from another area. Similarly, it is unclear how many neonatal discharges result in a readmission to paediatric services in a short time frame. Work linking neonatal care with paediatric care would be needed to investigate the long term implications of the initial length of neonatal stay for babies surviving to discharge.

However, what has been identified in this chapter is one ODN that potentially has a longer expected length of stay than the rest of England. This difference seems to be concentrated in their provision of high dependency care. The potential reasons for this increased length of stay could be investigated by that ODN and their commissioners. Similarly, an ODN has been identified with a potentially shorter length of stay, which appears to be explained by a difference in approach to the provision of special and high dependency care.

7.4.3 Choice of ODN for analysis

Any analysis which involves consideration of a care provider needs to determine which provider to allocate the responsibility of the care. For example, when a baby is born in need of neonatal care they are generally admitted to the unit in the hospital of their birth. An antenatal transfer of the mother before birth may have taken place to ensure the unit is appropriate for the baby's care needs. However, if their needs are more complex than that unit can provide, the baby may be transferred to a unit able to provide intensive care. Over time, if a baby survives and improves they may be transferred to a lower levelled unit closer to home in anticipation of discharge. Therefore, there are three healthcare providers in this situation which could be

considered in an analysis: the unit following birth; the unit where the majority of the care is received or the final unit before discharge from neonatal care. The same consideration could apply to ODNs, although due to the large geographical regions and the established network they cover, babies are less likely to be moved between different ODNs. In this dataset, approximately 96% of all days of care were provided within the ODN of birth.

7.4.4 STRENGTHS AND LIMITATIONS OF THIS ANALYSIS

This analysis has presented the first comparison of the different levels of care within ODNs in England. The differences in length of stay in terms of specific levels of care were described. Individual ODNs could use this approach to compare their provision of care to the rest of England and discuss how their approach differs and why this may have occurred.

The effect of ODN was adjusted for by introducing a categorical term to indicate if a baby was from that particular ODN or the rest of England. In other areas of neonatal medicine, mortality rates of ODNs have been compared using multilevel modelling (133). Use of a random effects model, or a Cox model with a frailty term, would allow all ODNs to be compared at once, with the random term representing a given ODN. This would provide shrunken estimates of the effect sizes, which reduces the risk of identifying an ODN with a higher hazard of death or discharge spuriously. However, the use of random effects or frailty models to model the different ODNs (134) was not undertaken here as these approaches have not yet been extensively investigated in a multistate modelling context and routine software does not yet exist (135).

Gestational age was adjusted for as a linear term due to the small number of babies, particularly those born at 24 and 25 weeks, within individual ODNs. The assumption of linearity may not be appropriate but more complex or categorical adjustment was likely to create issues of convergence or cause the model to over-fit and reduce statistical power.

7.5 CHAPTER CONCLUSION

In this chapter the differences in length of stay between ODNs were investigated. Two ODNs were selected as case studies, one with a long length of stay and another with a shorter length of stay. Both were formally compared with the rest of England in two separate multistate models, with the ODN included as an indicator term, and the levels of care where the differences in length of stay occurred were explored. In the ODN with a long length of stay this was due to an increase in the expected days receiving high dependency care. In the ODN with the short length of stay this difference was potentially due to a shorter expected time receiving special care than the rest of England. As demonstrated here, these methods could be useful for an individual ODN to compare their care to the rest of England to see in what ways their delivery of care differs and to aid the discussions of whether these differences are important.

8 DISCUSSION AND CONCLUSIONS

8.1 OVERVIEW OF CHAPTER

Preceding chapters have each contained a brief discussion section and in this chapter a summary of the entire thesis and a discussion of the main findings of this work is provided. The strengths and limitations are discussed, before considering how this work can inform clinical practice, particularly facilitating conversations between parents and clinicians and the commissioning of neonatal services. Finally, the chapter concludes with an overview of future work to be undertaken in this area.

8.2 SUMMARY OF KEY FINDINGS

The four over-arching aims of this thesis were:

1. To identify the factors that predict mortality and length of stay in the neonatal unit, focussing on babies born very preterm (less than 32 weeks gestational age).
2. To investigate the length of stay of very preterm babies (24 to 31 weeks gestational age) admitted for neonatal care to inform parental counselling regarding the risk of mortality and length of stay for a baby of given characteristics.
3. To examine the different levels of care, the neonatal care pathway, required whilst a very preterm baby (24 to 31 weeks gestational age) is in neonatal care to inform commissioning of specialist neonatal services by providing estimates of the number of days of levels of care required.
4. To compare the levels of care provided by different Operational Delivery Networks to babies born at 24 to 31 weeks gestational age in order to identify and investigate differences in care provision.

To answer the first aim of this thesis a systematic review was undertaken to identify the factors that have been used to predict mortality and length of stay in neonatal care. Research was well developed in the area of mortality prediction, with 19 studies identified over the time period of the review. Studies generally focussed on making

predictions of mortality based on the state of the baby at birth, or in the first 24 hours of life.

Whilst a recent systematic review had collated the evidence on the prediction of neonatal mortality (56), before this thesis no research had systematically investigated the prediction of length of stay in neonatal care. Studies investigating neonatal length of stay were more limited in terms of both their number and focus. As in the prediction of mortality, information from the first day of life was considered important to predict length of stay, particularly: gestational age, birthweight and sex. The review highlighted the importance of using a small number of clinical factors to predict length of stay to allow for ease of use in a clinical setting. This review was published in *BMJ Open* (18) and a copy can be found in Appendix 2.

To investigate the second aim of this thesis: the prediction of length of stay for very preterm babies, competing risks methods were explored. These methods allowed the measurement of time until discharge from neonatal care or death occurred. Both Cox proportional hazards models, and flexible parametric models for competing risks were introduced and compared and the advantages and disadvantages of each were discussed. The proportion of babies who survived to discharge and those who died in neonatal care over time were estimated initially from a model adjusted for gestational age and then developed further in a model adjusted for gestational age, birthweight and sex.

The median length of stay for babies who died in neonatal care was estimated to be approximately ten days, whilst the length of stay of babies who survived varied, particularly according to gestational age and birthweight. The babies born the most preterm had a median length of stay equal to approximately the time remaining until their due date. On the other hand, babies born at 30 and 31 weeks gestational age had a median length of stay approximately a month shorter than the time remaining until their due date. The estimates of mortality and length of stay can be used alongside clinician judgement to aid parental counselling. This work is currently being written for peer reviewed publication.

The third aim of this thesis was to investigate the different levels of care required by very preterm babies whilst in the neonatal unit. The extension of competing risks methods to multistate modelling was explored to consider the different levels of neonatal care: intensive care, high dependency care and special care. A Cox multistate model adjusted for gestational age was presented and expected length of stay estimates at each level of care were provided. For example, babies born at 31 weeks, were estimated to require four, six and 28 days of intensive care; high dependency care and special care respectively. The strengths and limitations of this analysis were discussed. This analysis was extended to consider additional covariates as identified from the systematic review: gestational age, birthweight and sex. These estimates can be used to aid the commissioning of specialist neonatal services. Work from this chapter has been published in *PLOS One* (28) and *Infant* (27) and copies can be found in Appendix 6.

Finally, the fourth aim of this thesis was to investigate whether differences exist in the length of stay of very preterm babies and the provision of the levels of care between different Operational Delivery Networks (ODNs). To investigate this, two ODNs were identified as case studies for further investigation due to having significantly longer or shorter lengths of stay. These case studies were investigated in two multistate models to identify how the provision of the different levels of care differed between networks. For the ODN with a longer than expected length of stay, this seemed to be explained by an increase in the expected time receiving high dependency care. This information could be used by the ODN to facilitate discussions about the care provision in their network.

8.3 STRENGTHS AND LIMITATIONS OF THIS THESIS

This study is the largest to date to consider the length of stay of all very preterm babies, and also consider the different levels of care that babies need whilst in hospital. The only similarly sized study (5), undertaken in the United States, highlighted the importance of future research including babies who die in length of stay research. The strengths and limitations of this work are considered in this section.

8.3.1 CHOICE OF STATISTICAL METHODS

A strength of this work is that it uses statistical methods which allow for consideration of the entire neonatal population when considering length of stay. Around 8% of very preterm singleton babies admitted to neonatal care will not survive to discharge (14), so to exclude babies who die, as most research has previously done (18), overlooks an important proportion of the population and neonatal care workload. In this thesis, exclusion of babies who died would have removed 1762 babies from the analysis and provided biased results, as acknowledged in previous research (19). This issue is also relevant for analyses investigating length of stay in other critical care areas with potentially high mortality rates including paediatric and adult intensive care.

The use of competing risks methods to investigate neonatal length of stay is relatively novel, although other researchers have used this approach to investigate length of stay whilst also considering survival in other clinical areas. Examples include investigation of: nosocomial pneumonia in adults (107), trauma patients (136) and burns injuries (137). Use of competing risks methods has been encouraged (108) and the appropriateness of these methods for modelling length of stay has been emphasised because *"treating death as a competing risk gives estimators which address the clinical questions of interest and allows for the modelling of both in-hospital mortality and TCS [time to clinical stability]/LOS [length of stay]…* (108)." Other researchers have also advocated competing risks methods as *"analysis by survival groups restricts the interpretation to only what would occur if the competing risk were not a possibility"* and this does not *"reflect the reality or interrelation between the outcomes* (137)."

Similarly to competing risks methods, the use of multistate modelling to investigate length of stay in neonatal care is new. However, recently studies in other clinical areas have also begun using multistate models for investigating length of stay and aiding the understanding of intermediate events (138-140). For example, Clark et al used multistate modelling to investigate length of stay, readmission to hospital and death following elective general surgery (139). This study considered multistate modelling superior to separate analyses as it provided a *"multi-dimensional"* description of surgical outcomes (139).

Whilst the use of competing risks methods and multistate modelling approaches are still relatively new for the investigation of length of stay, this research joins other researchers in investigating their potential and advocating their use in this context.

8.3.2 STATISTICAL MODELLING AND ASSUMPTIONS

The analyses in this thesis built on the theory of the Cox proportional hazards model. The main criticism of the Cox model is the assumption of proportional hazards, i.e. that the difference in the hazard between two or more groups is assumed to be constant over time.

In the competing risks analysis it was possible to relax the proportional hazards assumption by introducing a time-dependent covariate (110) for gestational age and this improved the model fit. Whilst it was not possible to relax proportional hazards in the multistate model, the results were re-estimated using groups of gestational age where proportional hazards were more likely to hold to investigate the robustness of the results. This improved the prediction of the probabilities and the appearance of the stacked plots (Chapter 6.7). However, it also provided reassurance that the estimation of the expected time receiving each level of care was estimated well despite the potential violation of the proportional hazards assumption. Therefore, whilst the assumption of proportional hazards is a limitation, a strength of this work is the thorough investigation of the impact of the assumption.

One approach not investigated in this thesis was the use of parametric modelling. As explained in Chapter 6.9.3 this approach assumes a parametric form for the hazard, for example by using a Weibull or Gompertz distribution (141). This method has been extended to competing risks and recently, multistate modelling (142). One advantage of this approach is that the distribution is fully specified and can be described numerically. As such, it is possible to calculate estimates at time points when events are not observed and, if desired, it is possible to extrapolate beyond the time frame of follow-up to provide predictions into the future. However, a disadvantage of parametric modelling, and the reason this approach was not used here, is that the distributional assumptions require the hazard to increase or decrease monotonically.

Clinically, it was unlikely that this assumption would hold and to impose this distributional assumption incorrectly would have led to imprecise results.

8.3.3 USE OF ROUTINE DATA

This study made use of a national routine dataset: the National Neonatal Research Database (NNRD) which has collected data related to neonatal care and been managed by the Neonatal Data Analysis Unit since 2007. The advantage of the use of the NNRD is that these data encompass the entire population, and provided a large sample size and information about each day of care for all admissions to neonatal care. Additionally, implementation of the findings from this thesis will not require any additional data collection by neonatal units or healthcare providers.

A limitation of the use of routine data is that it was created for clinical purposes, for example to aid with financial re-imbursement, rather than research, and this can potentially lead to reduced data quality. However, the NNRD has data quality systems in place (68) and variables related to the level of care are collected well, particularly those required by the Neonatal Critical Care Minimum Data Set (143), as they inform the commissioning and costing of care within a neonatal unit. Feedback loops have provided neonatal units with quarterly data quality checks since 2013 for variables used in the National Neonatal Audit Program (NNAP) to improve data completeness (3). Variables used in this thesis had high levels of completeness, with less than 0.5% of data missing for levels of care and 0.7% missing or implausible for birthweight or sex of the baby.

8.3.4 SELECTION OF THE STUDY POPULATION

A limitation of this research is that only singleton babies were considered and this excludes a proportion of very preterm births which are multiple births (144). The inclusion of multiple births born at 24 to 31 weeks of gestation could have increased the sample size by approximately 7,500 babies and increased the total number of care days by approximately 500,000 days. However, multiple births are known to have a have a higher risk of mortality, particularly for those born at 24 to 27 weeks of gestational age (144, 145). Multiple births comprise a different population in terms of

their demographic characteristics, with a different risk profile to singleton babies. These differences include that multiple births are often born to older mothers and a lower prevalence of some pregnancy complications including maternal hypertension (145). Therefore, the selection of the population in this research was appropriate, and future work can extend these findings to other populations.

8.3.5 FINAL OUTCOME FOR AN INDIVIDUAL BABY

A limitation of these data were that as the NNRD only covered England during the time period of this thesis and therefore transfers for neonatal care in units outside of England may have introduced an artificially short length of stay. However, a strength of this work is the robust sensitivity analysis which investigated babies discharged to other hospitals. The results of this sensitivity analysis (Chapter 5.7.5) indicated that the assumption that these babies had survived to discharge did not impact on the results. Inspection of the data also indicated that most discharges to other hospitals seemed to be admissions to postnatal care in England (i.e. discharges from specialist neonatal care), and less than 1% of discharges to other locations were to neonatal units in Wales or Scotland.

The outcome of discharge from the neonatal unit also comprised other outcomes including discharge for surgery and cardiac care. For the commissioning of neonatal care services this outcome reflects the end of the need for specialist neonatal services if there was no subsequent readmission to neonatal care. Many of these babies may have been admitted to paediatric wards before discharge home. A sensitivity analyses indicated the results were not impacted on by the assumption that these babies survived. An additional endpoint of discharge to other services may have been possible, although the number of babies experiencing this was very small and the endpoint does not reflect the neonatal care pathway investigated in this work.

In this analysis the outcomes of death or discharge from neonatal care were considered. Deaths can occur elsewhere including deaths at home and within a hospice. It is not possible to assess the impact of this on these results, however the number of deaths of babies outside of neonatal care is likely to be low with most anticipated neonatal deaths occurring in hospital (71). However, this is not a limitation

of this work as only the care received in neonatal services was considered in this thesis. For other applications, for example considering the longer term morbidity or care needs of a surviving baby throughout infancy and childhood, these endpoints may be important intermediate steps in an analysis (see Chapter 8.6.4). Similarly, when considering how to extend this section of the thesis to aid communication of levels of care to parents in the future these alternative discharge locations will need to be considered to counsel parents about their potential.

8.3.6 ORDERING AND LEVELS OF CARE

In this thesis, care was assumed to occur hierarchically, i.e. that if any intensive care was required it was all received first, followed by all high dependency care, and finally all special care. This hierarchical receipt of care was defined as the *neonatal care pathway* in this thesis. However, in day-to-day neonatal care, this may not be how care is actually received, particularly for intensive care and high dependency care when a baby may alternate between different levels of care as their clinical condition stabilises or deteriorates. It is clinically likely that for a baby a return to a higher level of care will alter the probability of death and discharge. Therefore the hierarchical assumption of this thesis is a limitation for the understanding of clinical care pathways. However, for commissioning purposes the important results are the totals, or expected, number of days of care required at each level rather than the detail of how and in what order the care was required.

Individual clinicians may be interested in specific neonatal care pathways. For example, there may be interest in the probability of death for babies requiring a step up to intensive care after a period of step down to high dependency care. An example of this would be a baby requiring re-intubation after a day of managing on CPAP (continuous positive airways pressure). It is likely that a step up to receiving intensive care would alter the probability of death as well as the probability of being stepped down to high dependency care again. Information about these specific care pathways may be useful for clinical management and this could be modelled by the introduction of additional states, for example: ‘intensive care after high dependency care’ or by allowing the model to be bi-directional (146). This modelling would have led to a violation of the

Markov assumption and so alternative non-Markov models would need to be considered. However, this modelling approach would be problematic as it would lead to over-fitting of sparse data. It would also be important to determine the number of movements between levels of care which would be considered clinically meaningful. An alternative would be to produce separate models to investigate transitions of particular interest after identifying them using clinical knowledge *a priori*. This is a limitation of this thesis and future work should build on the foundation provided here to consider the neonatal care pathway further.

8.4 CLINICAL IMPACT OF THESE FINDINGS

The aim of this thesis was to provide information for two main audiences: 1) for clinicians to use in clinical discussions and in parental counselling and 2) for the commissioners and those responsible for the allocation of specialist neonatal resources and services. There has been interest in the results of this thesis including talks, poster presentations at conferences and peer-reviewed papers. Detailed information about the impact and dissemination of this thesis to date can be found in Appendix 8.

8.4.1 CLINICAL AND PARENTAL DISCUSSIONS ABOUT LENGTH OF STAY

The predicted median lengths of stay for babies who survive to discharge and those who die in neonatal care (Chapter 5) were provided for babies of selected clinical characteristics. These estimates can be used by clinicians in clinical pre-ward round discussions including consideration of likely discharge dates and care planning for an individual baby.

These estimates can also be used by clinicians to inform conversations with parents about their baby's progress throughout the neonatal care journey. In this work, approximately half of all deaths occurred in the first ten days and this may be a time point which clinicians use to discuss length of stay with parents.

Work is underway in collaboration with parents, neonatal charities, clinicians and other stakeholders to disseminate findings from this section of the research. Evidence has suggested that if parents are involved in ongoing discussions about their baby's

progress and care there may be the potential to reduce length of stay of the baby where appropriate (147) and improve parental wellbeing (148, 149). Additionally, discussions about length of stay may improve parental preparedness for discharge (148), with unpreparedness often believed to be a cause of parental distress (150). The information from this thesis will inform the discussions between clinicians and parents.

It is essential that findings from this thesis are accessible. To facilitate this a parent panel meeting was held in June 2017 at Bliss (neonatal charity for babies born preterm and sick) in London, to discuss the most effective approach for disseminating these findings. This panel comprised a diverse group including: bereaved parents; parents of preterm born children with ongoing health concerns and parents of well preterm born children. This group of parents stated that they did not want *'another information sheet'* and would prefer to access the information via conversations with their clinicians. Therefore, work will be undertaken to communicate the results of this study for clinicians to use in conjunction with their clinical judgement when counselling parents. For clinicians, dissemination of these results will be via leaflets summarising the main results, talks to Operational Delivery Network clinical meetings (I have presented to these in the past) and via appropriate professional publications including *Infant* and BAPM newsletters. I am consulting with clinicians about how to present this work in an accessible format. Whilst clinicians find the graphs produced in this thesis informative, it can be difficult to extract information from them quickly, and so a table (similar to Table 5-18) may be more accessible in clinical practice.

Southampton neonatal unit have used results from this work to inform their length of stay estimates (personal communication with Fiona Lawson, neonatal intensive care matron). Other neonatal units are interested in using the results of this study to inform their clinical practice.

8.4.2 Clinical and parental discussions about level of care

Whilst overall length of stay could be considered separately for both babies who die and those who survive, it was not possible to estimate the time spent receiving each level of care for each group. Therefore, these estimates may be of less relevance for the counselling of parents, since the expected length of stay will never reflect a baby

who lives or a baby who dies, but rather an average baby of those characteristics (e.g. a specific week of gestational age). However, parents are interested in the levels of care provided for their baby. The parent panel involved in this research explained that often a stepping down of care was seen as *'less support'* rather than their *'baby getting better'* and that better counselling in this area was needed. This group also indicated that they were often excluded from discussions about the stepping down of their baby's care and so they felt unprepared.

8.4.3 Specialist commissioning and planning of levels of care

Estimates of the expected number of days needed at each level of care (Chapter 6) are provided in this thesis by week of gestational age and other characteristics. These results can be used to provide an average number of days of care needed for babies of certain characteristics. Alternatively, these estimates can be multiplied by the number of babies anticipated in a given time period with those characteristics. This will provide an estimate of the number of days required for each level of care for the entire population of very preterm singleton babies admitted for neonatal care. An example of this was provided in Table 6-9.

8.4.4 Comparison of ODNs

An analysis was undertaken to compare two ODNs to the rest of England in Chapter 7. This comparison may be of interest to ODN managers and commissioners who seek to understand differences in care. In the future, more detailed investigation of the care provided by different ODNs could allow consideration of how it differs from the rest of England and, alongside other parameters of interest, whether these differences are important and any possible modification.

8.5 Other work

In addition to work directly related to this thesis, whilst undertaking this research I have retained other research commitments in the broader area of neonatal care. This has included writing papers with colleagues, and alongside published work from this thesis I have published an additional eight papers and on four of these I am the second

author (list provided in Appendix 8). In 2016 I assisted the MBRRACE-UK (Mothers and Babies: Reducing Risk through Audits and Confidential Enquires across the UK) team with the statistical analyses for their annual report.

I have been involved as a co-applicant in the OPTIPREM (optimising neonatal service provision for preterm babies born between 27 and 31 weeks of gestation) study. This research is funded by the NIHR Health Services and Delivery Research programme (2017-2020, value approximately £925,000).

I have also worked with two charities during the writing of this thesis: Bliss and Together for Short Lives. I have assisted Bliss with the establishment of a funding application process for research. I worked together with them, and other stakeholders including parents, to identify research priorities for Bliss. My work with Together for Short Lives arose from the need to investigate neonatal palliative care. These babies were excluded from my thesis (Chapter 3) but I assisted this charity in the development of guidance for clinicians who assist in neonatal palliative care cases.

8.6 FUTURE WORK

Throughout this thesis areas for potential future research have been identified. These are outlined in this section.

8.6.1 METHODOLOGICAL WORK

Future development of multistate modelling may allow investigation of gestational age as a time-dependent covariate. Alternatively, instead of flexible parametric modelling, recent research has suggested that use of a parametric approach with mixture distributions, a mixture of density functions in different proportions (151), could have the potential to aid the modelling of length of stay (152).

In this thesis, comparisons between ODNs and the rest of England were undertaken using an indicator term in a multistate model. Other approaches may have been preferable, for example in a standard survival analysis, with the measuring of time until one event occurred, a frailty term could be used to create a random effect term for the different networks. However, this approach has not, as yet, been routinely extended to

multistate modelling. Methodological work should consider the comparison of healthcare providers, specifically in time-to-event research, further.

8.6.2 CONSIDERATION OF MULTIPLE BIRTHS AND BABIES BORN NEARER TERM

This thesis has focussed on very preterm singleton babies, specifically those born at 24 to 31 weeks of gestational age. These babies generally have the longest length of stay of all neonatal admissions and the primary reason for admission will be complications related to their prematurity such as the need for respiratory support.

Whilst this thesis has focussed on very preterm babies, others born around term contribute more than half of neonatal unit admissions (7), and even if these babies only stay for a short period, they create a large demand on neonatal units due to the size of the population. In the data source used for this thesis, around 180,000 babies born at term (37 weeks gestational age and beyond) were admitted for neonatal care over the same time period. These babies required a total of approximately 1,000,000 days of care (an average of 5.5 days each). Babies born nearer term are a heterogeneous group (5), and are admitted to neonatal care for various needs, for example babies experiencing congenital anomalies, surgery or complications following birth. These babies may require a long stay in the neonatal unit whilst other babies may only need to be admitted for a short period of care after birth. Examples of shorter lengths of stay may relate to needing phototherapy, antibiotics or a brief period of observation.

Lee et al (5) suggested that length of stay in the population born nearer term should be stratified to take account of their clinical condition and future work should consider this group further because *"there may not be one model that optimally fits all patients…..groups have different morbidities and developmental issues."* An example of this was seen in babies born near term with surgical issues where there was wide variation of length of stay even within one neonatal unit for the same condition. Some of these differences may be explained by different levels of severity of a condition and other acquired clinical factors including sepsis (153). To minimise the differences between groups of term babies, methods such as a cluster analysis could be used to identify similar groups of babies for whom to predict length of stay. It may be

appropriate in this population, with very specific health conditions, that the investigation of length of stay could be considered by the National Neonatal Audit Program (NNAP) (3).

Multiple births make up a large proportion of preterm births, and future work should consider how to include these babies in length of stay analysis. This may require appropriate risk adjustment or stratification of the analysis or it may be that length of stay for multiple babies is similar to singletons. Additionally, consideration is needed to account for the similarities between siblings such as via the use of a hierarchical model.

8.6.3 INCLUSION OF NEONATAL COMMUNITY CARE

The National Institute for Health and Care Excellence (NICE) recommends that a measure of quality of neonatal care is the provision of a coordinated transition to community care (154). Neonatal outreach community teams allow for babies to be discharged home with minimal healthcare needs who would only need special or transitional care if they remained in hospital. For example, babies can be discharged home receiving oxygen or whilst being tube fed. Provision of neonatal community care is not consistently offered throughout the country. In a survey undertaken in 2011 (155), only 45% of neonatal units had a dedicated community team, and of those only 48% had weekend cover. However, 85% of units with community teams felt that the team facilitated the discharge of babies. Research from 2002 suggested that the provision of neonatal community services could reduce median length of stay by two days (156). In this thesis, it was not possible to consider the impact of neonatal community care. A full understanding of the mechanisms of community care is needed in the future to understand differences in care between units and networks, particularly when comparing healthcare providers.

8.6.4 CHOICE OF OUTCOME AND TIME HORIZON

This thesis has not considered whether a shorter length of stay is beneficial or harmful for a baby born very preterm throughout their early childhood and beyond. Whilst a short length of stay may seem beneficial, particularly for healthcare costs, this may

result in a later adverse outcome. This was also noted by Lee et al who stated that whilst a *"shorter length of stay is a preferred outcome, this may not always be a true assumption as later outcomes are unknown…. Nor do we know the potential impact of length of stay (short or long) on longer term outcomes"* (5). A simulation study by DeRienzo et al suggested that reducing length of stay does not result in reduced hospital resource use, and that initiatives to reduce length of stay should also consider the longer term clinical outcomes (157).

Research linking length of stay and subsequent readmission to hospital has been undertaken in babies born nearer to term. However, evidence in this population is conflicting. The results from a Cochrane review investigating whether early discharge was harmful were inconclusive (158). Metcalfe et al reported that readmission rates were lowest following a one to two day stay after a vaginal birth for term births (159). However, other studies of late preterm babies found that a longer length of stay did not result in a reduced readmission rate (160) and that an early discharge, after less than two days in hospital, appeared to increase the risk of readmission, particularly for jaundice and infections (161).

There is no recent research investigating the outcomes of very preterm babies after a short or long length of stay. Nevertheless, some neonatal units have actively attempted to minimise length of stay. For example, an intensive care unit in the California Perinatal Quality Care Collaborative reduced their average length of stay by adjusting their clinical management including assessing if the preterm baby was able to feed at 32 weeks PMA which is two weeks sooner than current policy (162). Future work should consider whether a shorter or longer length of stay is beneficial by considering readmission to the paediatric care or other specialist services as an additional state in this analysis.

In this thesis time to discharge from neonatal services was considered. A small proportion of babies are known to spend additional time in specialist services including cardiac units or surgical centres. Future work should consider the care needs of these babies, which may be different from other very preterm babies. Data linkage may be required between data sources of different clinical services to undertake research into their needs.

8.7 FINAL CONCLUSION

In conclusion, as survival has improved in neonatal care, particularly over the last twenty to thirty years, the need to accurately estimate length of stay has increased. Whilst neonatal survival has improved, in-unit mortality remains high for the babies born very preterm, particularly those born at less than 32 weeks gestational age. Anecdotally, the two largest concerns for parents are: the probability of survival and how long their baby will need to remain in the neonatal unit. At the same time, the healthcare service is under increased pressure to use the limited resources available efficiently. Therefore, research which combines these two concerns is both timely and necessary. Competing risks methods and multistate modelling, as presented in this thesis, offer two approaches which can be used to estimate length of stay whilst considering survival. The results provided in this thesis can be used by clinicians to inform conversations with parents, and by commissioners seeking to understand neonatal care provision.

9 APPENDICES

APPENDIX 1

There are no Appendices from Chapter One

APPENDIX 2

DATA EXTRACTION FORM FOR THE SYSTEMATIC REVIEW

Data extraction form: Clinical predictors of the neonatal in-hospital mortality and length of stay

General information	
Title:	**Author:**
Journal:	**Year published:**
Study design details:	**Second screen (if necessary) conducted by:**
Final status: ☐ Include ☐ Exclude ☐ Further details required (second form to be filled out)	**Exclusion reason:**

Study characteristics	
Research aim:	**Country of study:**
Years of study:	**Setting:** ☐ Intensive care unit ☐ Postnatal ward ☐ Other, specify:

Study population	
Gestational age range:	**Birthweight range:**
Any exclusions: ☐ Multiples ☐ Specific ethnicities ☐ Congenital anomalies ☐ Other (specify)	**Specify exclusions other than those listed:**

Outcome	
What was investigated: ☐ Mortality ☐ Length of stay ☐ Both	**Specify other outcomes considered:**
Study power for outcome: ☐ Statistically powered ☐ Not statistically powered ☐ Not possible to assess **Details:**	**How were variables selected for model:** ☐ Statistically (e.g. stepwise) ☐ Clinical judgement ☐ Previous literature ☐ Other, specify:
Variables in model statistically significant: ☐ Discussed ☐ Not discussed ☐ Other, specify: **Details:**	**Clinical significance:** ☐ Discussed ☐ Not discussed ☐ Not possible to assess **Details:**

Clinical predictors	
Statistical/clinical significance for mortality: ☐ Gestational age ☐ Birthweight ☐ Gender Other baby characteristics. Specify: ☐ Antenatal factors (e.g. antenatal steroids). Specify: ☐ Demographic factors (e.g. maternal age). Specify: ☐ Postnatal factor (e.g. surfactant therapy). Specify: ☐ Other factor/s. Specify full adjusted model:	Statistical/clinical significance for length of stay: ☐ Gestational age ☐ Birthweight ☐ Gender Other baby characteristics. Specify: ☐ Antenatal factors (e.g. antenatal steroids). Specify: ☐ Demographic factors (e.g. maternal age). Specify: ☐ Postnatal factor (e.g. surfactant therapy). Specify: ☐ Other factor/s. Specify: Specify full adjusted model:

Study quality will be discussed, and the following domains will be measured on a scale of low, medium or high risk (L/M/H) of bias using the prompting items given (see paper for full description). This is based on the Quality in Prognosis Studies (QUIPS) tool. As prognostic studies are known to often be of poor quality, this is likely to form a discussion point more than a reason for exclusion.

Study quality		
Domain	**Prompting items**	
Study participation Whether the study population represents the population of interest	(a) Adequate participation (b) Description of population (c) Description of baseline (d) Description of recruitment (e) Description of time period (f) Inclusion and exclusion criteria Note: for secondary analyses of routine datasets score low levels of bias here	
Study attrition Whether the data available adequately represents the sample of interest	(a) Adequate response rate (b) Attempts to collect information on drop out (c) Reasons for loss to follow up investigated (d) Description of patients lost to follow up (e) No difference between patients lost and those not Note: for secondary analyses of routine datasets score low levels of bias here	
Prognostic factor measurement The factor is measured the same way for all participants	(a) Clear definition of the factors (b) Method of measurement is valid and reliable (c) Appropriate cut points are used (if necessary) (d) Method for measurement is the same for all participants (e) Adequate proportion of study sample has complete data (f) Methods for imputation are used for missing data	
Outcome measurement The outcome is measured in a similar way across all participants	(a) Clear definition of outcome is given (b) Measurement is valid and reliable (c) Method and setting is the same for all participants	
Study confounding Confounding is accounted for appropriately	(a) All important confounders are measured (b) Definitions of how they are measured is included	

	(c) Measurement is valid and reliable (d) Measurement is the same for all participants (e) Imputation is used for missing data (f) Potential confounding is accounted for in the study design (g) Confounding is accounted for in the analysis	
Statistical analysis and reporting Analysis is appropriate and all primary outcomes are reported	(a) Sufficient data presented (b) Model building strategy is appropriate (c) Selected model is adequate for the study design (d) No selective reporting of results Addition from me (e) External or internal validation is discussed	

Was reference list investigated for further references of interest?

☐ Yes
☐ No

Reviewer comments:

WHAT FACTORS PREDICT LENGTH OF STAY IN A NEONATAL UNIT: A SYSTEMATIC REVIEW – PUBLISHED IN BMJ OPEN

Open Access Research

BMJ Open

What factors predict length of stay in a neonatal unit: a systematic review

Sarah E Seaton,[1] Lisa Barker,[2] David Jenkins,[1] Elizabeth S Draper,[1] Keith R Abrams,[1] Bradley N Manktelow[1]

To cite: Seaton SE, Barker L, Jenkins D, *et al*. What factors predict length of stay in a neonatal unit: a systematic review. *BMJ Open* 2016;**6**:e010466. doi:10.1136/bmjopen-2015-010466

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2015-010466).

Received 5 November 2015
Revised 2 August 2016
Accepted 16 August 2016

[1]Department of Health Sciences, University of Leicester, Leicester, UK
[2]Neonatal Unit, University Hospitals of Leicester NHS Trust, Leicester, UK

Correspondence to
Sarah E Seaton;
sarah.seaton@le.ac.uk

ABSTRACT

Objective: In the UK, 1 in 10 babies require specialist neonatal care. This care can last from hours to months depending on the need of the baby. The increasing survival of very preterm babies has increased neonatal care resource use. Evidence from multiple studies is crucial to identify factors which may be important for predicting length of stay (LOS). The ability to predict LOS is vital for resource planning, decision-making and parent counselling. The objective of this review was to identify which factors are important to consider when predicting LOS in the neonatal unit.

Design: A systematic review was undertaken which searched MEDLINE, EMBASE and Scopus for papers from 1994 to 2016 (May) for research investigating prediction of neonatal LOS. Strict inclusion and exclusion criteria were applied. Quality of each study was discussed, but not used as a reason for exclusion from the review.

Main outcome measure: Prediction of LOS in the neonatal unit.

Results: 9 studies were identified which investigated the prediction of neonatal LOS indicating a lack of evidence in the area. Inherent factors, particularly birth weight, sex and gestational age allow for a simple and objective prediction of LOS, which can be calculated on the first day of life. However, other early occurring factors may well also be important and estimates may need revising throughout the baby's stay in hospital.

Conclusions: Predicting LOS is vital to aid the commissioning of services and to help clinicians in their counselling of parents. The lack of evidence in this area indicates a need for larger studies to investigate methods of accurately predicting LOS.

Strengths and limitations of this study

- There is little research in the area of predicting length of stay (LOS) and this review investigates the limited evidence for the first time. The same articles were independently identified by two authors.
- This review draws together the limited evidence about predicting LOS and discusses the future work needed.
- A variety of settings, gestational groups and types of analysis were considered in the different studies in this review, and it was not possible to conduct a meta-analysis.

BACKGROUND

In the UK, 1 in 10 babies[1] will require specialist neonatal care. Although the most preterm and smallest babies have the highest risk of mortality, if they survive their length of stay (LOS) in the neonatal unit will be very long. As neonatal survival has improved over recent years, particularly for very preterm babies,[2] the number of babies requiring long-term neonatal care has increased. Consequently, the workload of the healthcare service, including the total number of days of care required has increased.

The ability to accurately predict LOS in neonatal care is vital for resource planning, commissioning of services and to aid clinicians in their counselling of parents. However, there is a paucity of evidence related to predicting LOS. Much of the limited evidence which does exist is from observational studies which may suffer from bias. Similarly, factors which are identified from a single study or hospital as being important for predicting LOS may be biased by local medical practice within that study or simply be chance findings. Therefore, it is vital that information about the factors which predict LOS is identified from multiple studies to provide robust evidence for future research.

The objective of this review was to identify factors which are important when predicting LOS, and to draw together and discuss the evidence which currently exists.

METHODS

Selection of studies

MEDLINE, EMBASE and Scopus were searched systematically for papers from 1994 to 2016 (May) which investigated the prediction of mortality and/or LOS. All articles were screened by one author, and a random 10% were screened by a second author to ensure reliability of the reviewing process.

Any differences in identified articles were discussed between the two authors. The results presented here relate to the prediction of LOS. The full search strategy is provided in the online supplementary table.

Inclusion criteria

Studies were included which reported risk factors for LOS in the neonatal unit, irrespective of the outcome for the baby, from a multivariable model (eg, logistic regression, linear regression). To be included studies needed to have been undertaken in a human population and have been published in English. Neonatal survival dramatically improved in 1994 with the introduction of routine surfactant use[3] and antenatal steroids and therefore the search was started from this year. Studies which included data from before and after 1994 were included.

Exclusion criteria of studies

Exclusion criteria were determined in advance and included:

- Conference proceedings, as these were not peer-reviewed, although efforts were made to investigate if the conference abstract was subsequently published;
- Review articles, letters and editorials as these did not contain original research;
- Countries which were outside the Organisation for Economic Co-operation and Development in 1994 to identify countries with a different demographic profile and healthcare service;[4]
- Clinical trials, as the population would be unlikely to be representative of other babies in neonatal care;
- Wrong study population, for example, investigation of a paediatric or maternal population, or outcome, for example, predicting readmission;
- Specific disease areas (eg, *Escherichia coli* outbreaks or infections) as these babies are very different to other babies in neonatal care;
- Work that was subsequently updated or validation studies.

Data extraction

A data extraction form was prepared in advance to aid extraction of all necessary information. Information extracted related to: general details of the study (to determine eligibility); study characteristics; study population; outcome; clinical predictors and the quality of the study. Reference lists of included studies were examined for any additional studies which were relevant. Identified prognostic factors were grouped into broad categories of: inherent factors; antenatal treatment and maternal factors; conditions of the baby; treatment of the baby and organisational factors.

Study quality

The quality of research is known to often be poor in prognostic studies,[5] and therefore quality was not used as a reason for exclusion from this review. However, study quality was considered and discussed using an adaptation of Quality In Prognostic Studies (QUIPS) tool.[6] Domains of quality included consideration of: study participation; study attrition; prognostic measurement (eg, measurement, validity, completeness of data); outcome measurement (eg, definition and measurement); risk adjustment and predictors (eg, discussion of missing data) and statistical analysis and reporting (eg, was the model building appropriate, validation considered). A study was considered to be of reasonable quality if potential bias introduced by these domains was minimised as far as practical.

This review was registered with PROSPERO (registration number: CRD42013006020). Ethical approval was not required for this review.

RESULTS

A total of 7996 studies were identified from a systematic search of MEDLINE, EMBASE and Scopus (see figure 1). After removing duplicates, 5042 studies were screened for inclusion in this review. For 4978 articles it was clear from the title and abstract that they did not satisfy the inclusion criteria. The remaining 64 articles were read in full and manual searching of references, led to a final total of 24 being identified. Of these nine studies investigated the prediction of LOS and are included in this review. Summary characteristics of the studies are provided in table 1.

Of the nine identified articles, eight were identified by both authors performing the screening, and the ninth was agreed on after discussion between the authors.

Description of LOS studies

Inclusion and exclusion criteria of LOS studies

Exclusions within the nine studies, included: (major) congenital anomalies (as defined by study authors as no standard exists);[7–11] deaths in hospital[7 9 11–13] or before admission to intensive care;[10] babies who were admitted for comfort care (neither intubation or cardiorespiratory resuscitation was provided);[10] step down care;[14] surgery;[7 9 11] ambiguous sex;[15] implausible birth weight;[15] non-normal care pathways;[12] in hospital >1 year;[8] previously discharged and readmitted,[11] transfers,[13] and transfers to long-term care facilities.[8]

Although most studies excluded infants who died in hospital; two papers included deaths in the calculation of LOS. One paper accounted for this in the methodology implemented[15] and another acknowledged 'mortality rates may have introduced bias, since non-survival truncates observed LOS'.[10] One study which excluded deaths[11] acknowledged that accounting for deaths in LOS 'may be particularly complex…'

Study populations within LOS studies

Studies investigated a variety of gestational ages and a range of different study settings (table 1) leading to varied populations. Studies appear to have been largely

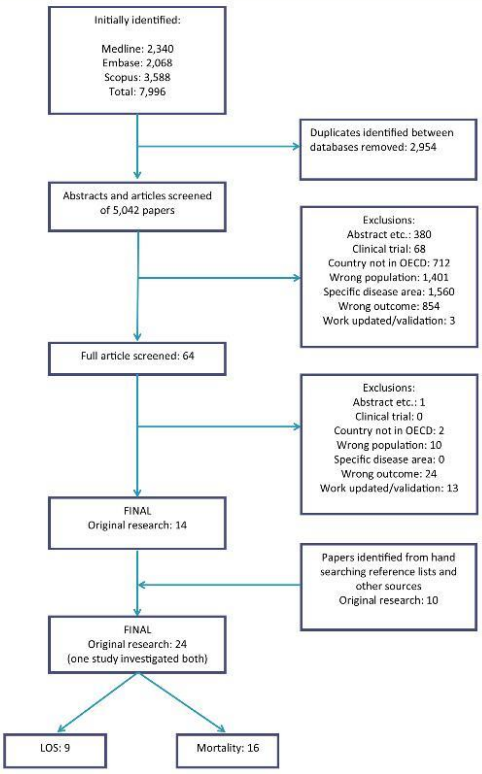


Figure 1 Flow chart documenting the results of the systematic review search. This review focuses on the articles identified which investigated the prediction of LOS. LOS, length of stay; OECD, Organisation for Economic Co-operation and Development.

based in intensive care units, although it is difficult to comment on whether individual babies within a study required or received intensive care (eg, mechanical ventilation) as no study stated this explicitly.

Prognostic factors in LOS studies

The nine identified studies investigating the prediction of LOS presented a total of 39 prognostic factors. These variables were grouped into broad categories of: inherent factors; antenatal treatment and maternal factors; conditions of the baby; treatment of the baby and organisational factors. Details of the prognostic factors identified by each study are given in table 2.

All nine studies accounted for some form of inherent factor, with the most common being birth weight (88.9%, 8/9), gestational age (55.5%, 5/9) and sex (55.5%, 5/9). Seven studies attempted to account for the condition of the baby. However, there was little consensus on what factor would be appropriate, with variables ranging from those occurring early in the care pathway (eg, admission reason) to those potentially occurring later on (eg, Retinopathy of Prematurity). Similarly, variables such as congenital anomalies were only accounted for by three studies (33.3%, 3/9); however, this often comprised part of the exclusion criteria (55.5%, 5/9).

Organisational factors were considered in 5 (55.5%) studies, with most relating to the setting of the care being received including transfers between units.[14]

Study quality of the LOS studies

Quality of research is well acknowledged as an issue in prognostic or prediction studies.[5] Therefore, an adapted form of the QUIPS tool was used to discuss the quality of the studies (table 3), although poor quality was not used as a reason for exclusion from the review. Domains of bias which were examined included: level of study participation; exclusion and attrition; how the outcome was measured; details about risk adjustment and information about the analyses, specifically if validation was conducted. Study participation was not an issue as all studies used data from routine sources and none actively recruited participants. Attrition caused by infants being transferred out of the area covered by the hospital/study was potentially an issue in all studies except one[10] which included LOS in other facilities. However, this study[10] was based in a single centre, and although they lost no infants to attrition, the details about the population they recruited only included care received while within that hospital site.

Seven studies used continuous LOS/postmenstrual age (PMA) as their outcome.[7 9–13 15] Two studies categorised LOS, one by dichotomising into <21 days and ≥21 days[14] and the other by classifying discharge as early or late (lowest and highest quartile of PMA).[8] The decision of how to model LOS was based on the statistical analysis being implemented. There were no issues in the measuring of LOS, as this is an objective, simple measurement.

Five studies had validated their results by splitting the sample during the initial analysis and holding some data back for validation purposes.[8–11 13] Two studies acknowledged that further validation was needed before results could be generalised[12 14] and one acknowledged that further work was needed to assess the modelling techniques.[15] One study, as part of their analyses, had conducted a preplanned external validation on a model presented in their paper, but concluded that the non-validated model was statistically superior.[10] Only one study did not mention validation of the results.[7] Therefore, a strength of these studies was that most addressed the issue of validation in some way.

In general, study quality was considered to be good with low levels of potential bias. There were few issues with study participation as most studies obtained data from medical notes which would introduce a low risk of bias. All studies had a defined outcome which could be

Table 1 Summary characteristics of the nine studies included in this review

	Country of study	Year of publication (data)	Exclusions in study	Number of patients in study	Population investigated	Physical location of study	Model selection	Statistical methods	Model fit methods
Altman *et al*[7]	Sweden	2009 (2004–2005)	Congenital anomalies; death; surgery.	2388	30–34 weeks gestational age	Neonatal units of varying levels of care	Univariate analysis then significant (p<0.2) entered into stepwise	Linear regression	R^2
Bender *et al*[10]	USA	2013 (1999 and 2002)	Congenital anomalies; death; admitted for comfort care.	293 (validated on 615)	All gestations	Neonatal intensive care unit	Prior knowledge	Accelerated failure time parametric models	Cross validation R^2
Berry *et al*[14]	Canada	2008 (2002)	Admitted for step down care.	604	All gestations	Neonatal intensive care unit	Prior knowledge	Logistic regression	None, but validation in other centres recommended
Hinchliffe *et al*[15]	UK	2013 (2006–2010)	Ambiguous sex; implausible birth weight.	2723	24–28 weeks gestational age	Neonatal intensive care unit	Prior knowledge	Competing risks	None (acknowledged as weakness)
Hintz *et al*[8]	USA	2010 (2002–2005)	Congenital anomalies; in hospital >1 years; transferred to long-term care.	2254	<27 weeks gestational age	Unclear but likely to be neonatal intensive care due to gestational age	Prior knowledge	Linear mixed model	R^2
Lee *et al*[9] (2013)	USA	2013 (2008–2010)	Congenital anomalies; death; surgery.	2012	401–1000 g birth weight	Neonatal intensive care unit	Stepwise selection	Linear mixed model	R^2
Lee *et al*[11] (2016)	USA	2016 (2008–2011)	Congenital anomalies; death; surgery; readmitted.	23 551	All babies 401 g–1500 g or 22–29 weeks gestational age plus larger babies meeting specified criteria	Neonatal intensive care units	Prior knowledge then minimum AIC	Negative binomial model with hospital as random effect	Root mean-square error (RMSE)
Manktelow *et al*[12]	UK	2010 (2005–2007)	Death; non-normal care.	4702	23–32 weeks gestational age	Neonatal unit.	Prior knowledge and then change in deviance to decide how to model variables	Quantile regression	Observed vs predicted comparison
Zernikow *et al*[13]	Germany	1999 (1989–1996)	Transfers; deaths.	2144	23–36 weeks gestational age	Unclear but single centre.	Forward stepwise	Artificial neural networks Multiple linear regression	Observed vs predicted comparison

Table 2 Prognostic factors for predicting length of stay included in the analysis of each study

	Altman *et al*[7]	Bender *et al*[10]*	Berry *et al*[14]	Hinchliffe *et al*[15]	Hintz *et al*[8]	Lee *et al*† (2013)[9]	Lee *et al* (2016)[11]	Manktelow *et al*[12]	Zernikow *et al*[13]	Number of studies
Inherent factors										
Birth weight (modelled in multiple ways including categorised, SGA, z score)	X (SGA)	X		X	X	X (+SGA)	X	X	X	8
Congenital anomalies			X				X		X	3
Date/year of birth							X		X	2
Ethnicity/race/nationality						X	X		X	3
Gestational age	X	X		X				X	X	5
Head circumference									X	1
Length of baby at birth									X	1
Multiplicity	X						X			2
Sex		X		X		X	X	X		5
SNAPPE-II‡			X							2
Any inherent factor	X	X	X	X	X	X	X	X	X	9
Antenatal treatment and maternal factors										
Antenatal steroids						X	X			2
Diabetes							X			1
Emergency delivery									X	1
Fetal distress						X	X			2
Hypertension						X	X			2
Maternal age	X						X			2
Mode of delivery							X			1
Other maternal/obstetric condition							X			1
Received prenatal care							X			1
Any antenatal treatment or maternal factor	X					X	X		X	4
Conditions of the baby										
Admission reason								X		1
Apgar score						X	X			2
Bronchopulmonary Dysplasia					X					1
Hyperbilirubinaemia	X									1
Hypoglycaemia	X									1
Infection	X									1
Respiratory distress syndrome	X									1
Retinopathy of prematurity (stage 3 or higher)					X					1
Sepsis episode or NEC					X					1
Severe morbidity§	X									1
SNAP¶		X								1
SNAPPE-II			X							2

Continued

Table 2 Continued

	Altman *et al*[7]	Bender *et al*[10]*	Berry *et al*[14]	Hinchliffe *et al*[15]	Hintz *et al*[8]	Lee *et al*† (2013)[9]	Lee *et al* (2016)[11]	Manktelow *et al*[12]	Zernikow *et al*[13]	Number of studies
Any condition of the baby	X	X	X		X	X	X	X		7
Treatment of the baby										
Surgery while in hospital			X							1
Surgery for patent ductus arteriosus, necrotising enterocolitis, or retinopathy of prematurity					X					1
Umbilical vein catheter									X	1
Ventilation									X	1
Any treatment of the baby			X		X				X	3
Organisational factors										
Centre (random effect)					X	X		X		3
Domiciliary care	X									1
Fixed discharge criteria	X									1
Level 3 centre	X									1
Transferred/outborn status			X				X			2
Any organisational factor	X		X		X	X	X			5

*The final model is taken to be the SNAP one as this model was validated.
†This study stratified analyses by birth weight, and different variables were used for each stratification. All variables from all models are listed here.
‡The calculation of the SNAPPE-II score includes: MBP; lowest temperature; Po2/FIO2 ratio; lowest serum pH; multiple seizures; urine output; birth weight; SGA and Apgar score. These are a combination of inherent and conditions of baby factors and so SNAPPE II appears in both categories.
§Severe morbidity is defined as: any of: IVH 3-4; ROP>=3; BPD.
¶This is the original SNAP score, devised in 1993, and comprised of 34 items, largely related to the condition of the baby. Examples of items belonging to the score include: heart rate, blood pressure and platelet count.
BPD, bronchopulmonary dysplasia; IVH, intraventricular haemorrhage; MBP, mean blood pressure; ROP, retinopathy of prematurity; SGA, small for gestational age; SNAP, Score for Neonatal Acute Physiology; SNAPPE, Score for Neonatal Acute Physiology Perinatal Extension II.

Table 3 Quality assessment of the included studies using a modified version of the QUIPS tool

	Domains of quality				
	Study participation	**Study exclusion/attrition**	**Outcome measurement (eg, definition and measurement)**	**Risk adjustment and clinical predictors* (eg, missing data)**	**Statistical analyses and reporting (eg, validation considered)**
Altman *et al*[7]	Study is population based (and included 21/34 units in Sweden) but infants were excluded if moved to a hospital not included in study. Data is collected	Infants discharged to other clinics were excluded.	Continuous postmenstrual age at discharge.	Detailed information about how factors were measured.	None mentioned
Bender *et al*[10]	Single centre study.	Transfers were included in the analysis and their LOS in other facilities was included in the total LOS. Sensitivity analyses excluded them.	Continuous LOS (days).	Made use of mortality scores with large number of elements included. Potential issues if there was missing data.	Split sample.
Berry *et al*[14]	Study based in two hospitals. Data extracted from ward registers, charts and patient records.	LOS days after transfer to another centre were not included.	LOS categorised into: <21 days or ≥21 days. No justification for these cut points.	Made use of mortality scores with large number of elements included. Potential issues if there was missing data.	Acknowledgement that future validation required.
Hinchliffe *et al*[15]	Population-based study covering a region of hospitals. Data is extracted from medical records and stored in a routine database used for research purposes.	Minimal losses to follow-up when discharged out of region covered by study. Included in analysis as censored observations.	Continuous LOS (days).	Detailed information about how factors were measured.	Acknowledged that further work is required to assess model.
Hintz *et al*[8]	Population-based study within a large network containing multiple hospitals. Data extracted from a routine database set up for research.	Attrition of infants transferred out of the region covered by study.	Early (lowest quartile of age at discharge) or late discharge (high quartile of age at discharge). No justification for these cut points.	Variables clearly defined. Some factors subjective in measurement (eg, Bells staging for NEC).	Split sample
Lee *et al*[9] (2013)	Population-based study of a large number of intensive care units.	Attrition from transfers to lower levels of care (acknowledged as causing bias).	Continuous LOS in days (log transformed).	Limited details about variables but most could be measured objectively.	Split sample
Lee *et al*[11] (2016)	Population-based study in 90% of intensive care units in large American state	Only babies inborn or transferred to unit in study within one day of life.	Continuous LOS (days).	Variables clearly defined and objectively measured. Missing data not discussed.	Split sample
Manktelow *et al*[12]	Population-based study covering a region of hospitals. Data is extracted from medical records and stored in a routine database used for research purposes.	Minimal attrition: when discharged out of region covered by study.	Continuous LOS (days).	Some factors subjectively measured (eg, reason for admission to intensive care).	Acknowledged that future validation needed.
Zernikow *et al*[13]	Single centre study	Transfers excluded from the study.	Continuous LOS (days).	Limited information about variables but most objective to measure.	Split sample

*Unmeasured and unknown confounders are always a potential issue within observational research, so no study has this specifically mentioned.
LOS, length of stay; NEC, necrotising enterocolitis.

objectively measured and so was unlikely to differ between studies indicating no issues of bias. Only one study[7] did not mention validation of the results, indicating that statistical analyses were well reported. While no formal scoring of study quality has been undertaken here, all studies had a level of quality which indicated there was a low level of bias given the constraints of the study designs.

DISCUSSION

In recent years, the ability to accurately predict LOS in the neonatal unit has become increasingly important. As neonatal survival has improved, the number of babies requiring long stays in the neonatal unit has increased. However, there has been limited evidence on how to predict LOS, and what factors are important to aid the prediction. This review has provided a systematic search of the literature to consider what factors should be considered in future analyses of LOS.

All of these studies investigated the prediction of LOS, although two studies categorised the outcome,[8 14] which leads to less informative estimates and therefore using a method which can appropriately model continuous LOS is more useful clinically. It is likely that the choice of how to measure LOS is decided by the selection of statistical method. A variety of methods were used, although surprisingly only one study used a survival analysis approach,[15] which is often the most popular methodology when measuring time to event.

Prognostic factors for LOS

All studies accounted for some form of inherent factors which have the advantage of being generally simple and objective to measure, and being present at birth. A prediction for LOS on the first day of life can be made using these factors. However, this prediction may change over time depending on the clinical progress of the baby, and the quality of care provision, during the baby's stay. However, there were a variety of study populations in this review even before adjustment for inherent factors, with predictions for extremely preterm[15] and for all babies.[10 14] A prediction model for all babies, such as that proposed by Bender[10] or Berry[14] is unlikely to perform well as the babies born near term may have very different reasons for being in the neonatal unit to those born preterm. This was discussed by Lee[11] who stratified their analyses by different birth weight groups to attempt to group similar babies together. They acknowledged that babies born at a normal birth weight may need further stratification by the reason for their admission, for example: sepsis or respiratory disease.[11] The approach appears reasonable, and future LOS predictions should focus on groups of babies with similar characteristics, for example, very preterm or very low birth weight, or analyses should be stratified by clinical condition.

It has been acknowledged that while this information from the first day of life is useful,[13] prediction is generally poor unless perinatal factors[8] or severity of illness[10] factors are also considered. However, there was little consensus on what this factor should be, with potential factors ranging from early occurring conditions (eg, reason for admission to intensive care) to those that occurred later in the care pathway (eg, retinopathy of prematurity). Therefore, while it may be important to account for the condition of the baby, there is little agreement over which factors should be used to do so. It is difficult to adjust for conditions which will only be experienced by surviving children. To provide an early prediction of LOS the clinical condition should be an event which occurs early in the care pathway, for example, Apgar score.

Congenital anomalies were not accounted for by many studies, but often formed part of the exclusion criteria within a study, indicating the importance of their consideration. However, there is no accepted list of what constitutes a major anomaly, and the term is often used to refer to a wide and varied range of conditions, making statistical adjustment or exclusions from a study difficult. Some congenital anomalies are unlikely to impact on LOS at all, whereas some severe anomalies or those that require surgery (eg, gastroschisis) may have a significant impact on LOS. Consequently even when studies exclude or adjust for major anomalies it can never be guaranteed that it is a comparison of 'like with like'. Thus, while congenital anomalies may have an impact on LOS, it is likely too broad a term to include in a LOS prediction model, but it should be considered by clinicians when revising LOS estimates using their clinical judgement.

It is difficult to account for organisational factors, although around half the studies attempted to do this in some way. However, one major issue with organisational factors is the variation between countries. Similarly, even within a country, the level of the unit may not indicate the type of care given to the infant. Despite this, these factors were seen by some authors to be equally or even more important than perinatal risk factors.[7] This demonstrates the importance of considering the varying levels of care provision within the country of the study. Studies focused in one or two centres such as those by Berry[14] or Bender[10] are likely to be inappropriate to draw definitive conclusions from as they may have high levels of loss to follow-up or loss of detail related to the baby's care, causing issues with estimating LOS. Within the UK, neonatal services are focused in clinical networks,[16] with each network providing the full range of neonatal care. Therefore, it may be appropriate to focus analysis and prediction at a network level to cover all varieties of care, attempting to avoid some of the issues presented by differing organisational factors, and to allow generalisability of the findings. Population-based studies may assist with this; however, these should potentially investigate the use of a random effect term for hospital or

equivalent to allow for variation between different healthcare services. Future work should consider the impact of a baby transferring between hospitals on their LOS.

Thresholds for discharge

Thresholds for determining the timing of discharge informally exist within neonatal medicine. Babies are rarely discharged before they gain the ability to suck and feed (around 35 weeks of gestational age). Irrespective of clinical conditions experienced, most preterm born babies (particularly <32 weeks) are likely to have matured and recovered enough to be discharged at this point, their prematurity being the overwhelming reason for their LOS. For a small number of babies, later occurring conditions (eg, late occurring sepsis, surgical needs) may cause a dramatic increase in their LOS. However, these will not be identifiable for a long period after birth and so potentially, prediction of LOS should be adapted in light of these conditions, if appropriate.

While the LOS of preterm babies is largely determined by their prematurity, normal birthweight babies[11] and those born closer to term are likely to have varied reasons for their LOS making predictions complex. These babies should be considered separately or adjustment or stratification should be made in any prediction model.

Clinical use of prediction models

Clinically, prediction models with a smaller number of factors are easier to use,[8] and this also reflects the concept of statistical parsimony ('simplicity'). This was seen in the area of predicting neonatal mortality, where complex risk scores, such as the Score for Neonatal Acute Physiology (SNAP), were developed and subsequently simplified to allow easier use.[17 18] Even following simplification, these risk scores are, at times, still difficult to implement. For example, the simplified SNAP score still requires the assumption that where medical tests are not performed, the results should be considered normal.[17] Therefore, while accurate prediction is needed, this must be balanced against the need for a simple model, suitable for 'bedside use'.

Clinical judgement is important and potentially informative for predicting LOS, although this was not possible to investigate here. However, prediction models, such as those identified, are useful because they can provide estimates that are more accurate then clinical judgement and assessment alone.[19] It is likely that a statistical estimate of LOS, used in conjunction with clinician judgement, for example, when considering congenital anomalies, may provide the best estimate.

Strengths and limitations of this review

There is little research in the area of predicting LOS and this review investigates the limited evidence for the first time. However, it was difficult to identify a clearly defined population for whom to predict LOS. A variety of settings and gestational groups were considered in the different studies in this review, and it is likely that different gestational ages will require different prediction models, incorporating very different factors. Future research will need to specifically investigate this in large studies.

A meta-analysis of the data presented in this review was not undertaken, due to the varying analyses and adjustments made in each study. Theoretically, an individual patient data meta-analysis could have been undertaken in order to overcome these issues; however, this is known to be difficult, particularly with acquiring the necessary data.[20] Similarly, it was not possible to investigate publication bias due to the varying analyses and potentially this could have been an important issue. Owing to these limitations, as suggested in other medical areas, a large-scale study may be important and clinically useful.[21]

CONCLUSIONS

The ability to predict LOS would be valuable to parents and families, clinicians and service providers, but it is a complex issue. Inherent factors appear to be the most important to account for, particularly birth weight, gestational age and sex. This information from the first day of life is informative for predicting LOS in a simple model and these estimates are a useful early indicator of LOS.

It may be important to consider revising this initial estimate over time if a late occurring condition dramatically adds to the initial LOS prediction. However, it is hypothesised that many medical conditions will resolve before the point at which the baby is well enough in terms of their prematurity to be discharged. In cases where this assumption is unrealistic more complex (dynamic) risk-prediction models would possibly be required.[22] Studies predicting LOS should be at a population level to avoid the issue of organisational factors, and to allow generalisability of the findings.

Contributors SES conceived and designed the study. She undertook the literature search and assessed studies for eligibility. DJ conducted the second review and discussed all results with SES. Results were interpreted and critiqued by SES with assistance from BNM. LB provided clinical input and assisted with the discussion of the results. All authors reviewed and revised the manuscript and provided input to interpreting the results, drafting the manuscript and revising it critically. All authors read and approved the final version of the manuscript.

Funding SES is funded by a National Institute for Health Research Doctoral Research Fellowship award (DRF-2013-06-011). This paper presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. KRA is partially supported by the National Institute for Health Research (NIHR) as a Senior Investigator (NF-SI-0512-10159).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement There is no additional data available.

REFERENCES

1. National Institute for Health and Clinical Excellence (NICE), Quality Standards Programme. *NICE cost impact and commissioning assessment: quality standard for specialist neonatal care*. 2010 https://www.nice.org.uk/guidance/qs4
2. Field DJ, Dorling JS, Manktelow BN, *et al*. Survival of extremely premature babies in a geographically defined population: prospective cohort study of 1994-9 compared with 2000-5. *BMJ* 2008;336:1221–3.
3. Hintz SR, Poole WK, Wright LL, *et al*. Changes in mortality and morbidities among infants born at less than 25 weeks during the post-surfactant era. *Arch Dis Child Fetal Neonatal Ed* 2005;90: F128–33.
4. Organisation for Economic Co-operation and Development (OECD). *OECD: better policies for better lives*. 2014. http://www.oecd.org
5. University of York Centre for Reviews Dissemination. *Systematic reviews: CRD's guidance for undertaking reviews in healthcare*. Centre for Reviews and Dissemination, 2009.
6. Hayden JA, Van Der Windt DA, Cartwright JL, *et al*. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158:280–6.
7. Altman M, Vanpée M, Cnattingius S, *et al*. Moderately preterm infants and determinants of length of hospital stay. *Arch Dis Child Fetal Neonatal Ed* 2009;94:F414–18.
8. Hintz SR, Bann CM, Ambalavanan N, *et al*. Predicting time to hospital discharge for extremely preterm infants. *Pediatrics* 2010;125:e146–54.
9. Lee HC, Bennett MV, Schulman J, *et al*. Accounting for variation in length of NICU stay for extremely low birth weight infants. *J Perinatol* 2013;33:872–6.
10. Bender GJ, Koestler D, Ombao H, *et al*. Neonatal intensive care unit: predictive models for length of stay. *J Perinatol* 2013;33:147–53.
11. Lee HC, Bennett MV, Schulman J, *et al*. Estimating length of stay by patient type in the neonatal intensive care unit. *Am J Perinatol* 2016;33:751–7.
12. Manktelow B, Draper ES, Field C, *et al*. Estimates of length of neonatal stay for very premature babies in the UK. *Arch Dis Child Fetal Neonatal Ed* 2010;95:F288–92.
13. Zernikow B, Holtmannspötter K, Michel E, *et al*. Predicting length-of-stay in preterm neonates. *Eur J Pediatr* 1999;158:59–62.
14. Berry MA, Shah PS, Brouillette RT, *et al*. Predictors of mortality and length of stay for neonates admitted to children's hospital neonatal intensive care units. *J Perinatol* 2008;28:297–302.
15. Hinchliffe SR, Seaton SE, Lambert PC, *et al*. Modelling time to death or discharge in neonatal care: an application of competing risks. *Paediatr Perinat Epidemiol* 2013;27:426–33.
16. Developing Operational Delivery Networks: The Way Forward. 2012. http://www.england.nhs.uk/2012/12/21/odn/ (23 Mar 2015).
17. Richardson DK, Corcoran JD, Escobar GJ, *et al*. SNAP-II and SNAPPE-II: simplified newborn illness severity and mortality risk scores. *J Pediatr* 2001;138:92–100.
18. Richardson DK, Gray JE, Mccormick MC, *et al*. Score for neonatal acute physiology: a physiologic severity index for neonatal intensive care. *Pediatrics* 1993;91:617–23.
19. Einhorn HJ. Accepting error to make less error. *J Pers Assess* 1986;50:387–95.
20. Abo-Zaid G, Sauerbrei W, Riley RD. Individual participant data meta-analysis of prognostic factor studies: state of the art? *BMC Med Res Methodol* 2012;12:56.
21. Mcshane LM, Altman DG, Sauerbrei W. Identification of clinically useful cancer prognostic factors: what are we missing? *J Natl Cancer Inst* 2005;97:1023–5.
22. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* 2011;67:819–29.

APPENDIX 3

ETHICAL APPROVAL FROM LANCASTER RESEARCH ETHICS COMMITTEE

1

NHS
Health Research Authority
National Research Ethics Service

NRES Committee North West - Lancaster
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0161 625 7818
Fax:0161 625 7299

22 May 2014

Ms Sarah E Seaton
Department of Health Sciences
22-28 Princess Road West
Leicester
LE1 6TP

Dear Ms Seaton

Study title:	**Modelling neonatal care pathways: costs and consequences for the future**
REC reference:	**14/NW/0349**
Protocol number:	**0415**
IRAS project ID:	**148248**

The Proportionate Review Sub-committee of the NRES Committee North West - Lancaster reviewed the above application on 22 May 2014.

We plan to publish your research summary wording for the above study on the NRES website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact the REC Manager Mrs Carol Ebenezer, nrescommittee.northwest-lancaster@nhs.net.

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. The REC will acknowledge receipt and provide a final list of the approved documentation for the study, which can be made available to host organisations to facilitate their permission for the study. Failure to provide the final versions to the REC may cause delay in obtaining permissions.

2

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to contest the need for registration they should contact Catherine Blewett (catherineblewett@nhs.net), the HRA does not, however, expect exceptions to be made. Guidance on where to register is provided within IRAS.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion").

Approved documents

The documents reviewed and approved were:

Document	*Version*	*Date*
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only)		23 April 2014
Letters of invitation to participant	1.3	10 April 2014
Other [Study information leaflet]	7	24 April 2014
REC Application Form	3.5	12 May 2014
Research protocol or project proposal	1.4	10 April 2014

3

Summary CV for Chief Investigator (CI)	Manktelow	
Summary CV for Chief Investigator (CI)	Seaton	12 May 2014
Summary CV for Chief Investigator (CI)	Abrams	
Summary CV for Chief Investigator (CI)	Draper	

Membership of the Proportionate Review Sub-Committee

The members of the Sub-Committee who took part in the review are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

14/NW/0349 **Please quote this number on all correspondence**

Yours sincerely

pp.

Dr Lisa Booth
Chair
Email: nrescommittee.northwest-lancaster@nhs.net

4

Enclosures:	*List of names and professions of members who took part in the review*
	"After ethical review – guidance for researchers"
Copy to:	*Mrs Wendy Gamble,*

5

NRES Committee North West - Lancaster

Attendance at PRS Sub-Committee of the REC meeting on 22 May 2014

Committee Members:

Name	*Profession*	*Present*	*Notes*
Dr Lisa Booth	Senior Lecturer / Chair	Yes	
Mrs Valerie Skinner	Nurse (Retired)	Yes	
Professor Jois Stansfield	Professor of Speech Pathology	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Mrs Carol Ebenezer	REC Manager

RESEARCH & DEVELOPMENT (R&D) APPROVAL FROM CHELSEA AND WESTMINSTER HOSPITAL NHS FOUNDATION TRUST

Chelsea and Westminster Hospital NHS
NHS Foundation Trust

Research and Development Support Office

17/06/2014

Professor Neena Modi
Imperial College London/Chelsea & Westminster NHS Foundation Trust
Section of Academic Neonatal Medicine,
Imperial College London
Chelsea and Westminster Hospital,
369 Fulham Ro
London
SW10 9NH

Dear Professor Neena Modi

R&D Number: C&W14/043
IRAS ID: 148248
REC Ref: 14/NW/0349
Study Title: Modelling neonatal care pathways: costs and consequences for the future

I am pleased to inform you that the R&D review of the above project is now complete, and the project has been formally approved to be undertaken at Chelsea and Westminster Hospital NHS Foundation Trust under the terms of the enclosed Site Investigator Agreement. The documents reviewed are as follows:

Document	Version	Date
Evidence Of Sponsor Insurance		23/04/2014
Letters of Invitation to Participant	1.3	10/04/2014
Other (Study information leaflet)	7	24/04/2014
Site Investigator Agreement		12/06/2014
Research Protocol	1.4	10/04/2014
R&D Form		06/05/2014
REC Favourable opinion Letter		22/05/2014

Permission is granted on the understanding that the study is conducted in accordance with the Research Governance Framework, GCP and applicable NHS Trust policies and procedures. R&D standard operating procedures are available to download from the intranet or can be requested by emailing research.development@chelwest.nhs.uk.

The R&D approval applies for the duration of the research except where action is taken to suspend or terminate the approval early. Where the duration of the study is to be extended beyond the period specified in the in the R&D form, you must notify the R&D Support Office prior to the extension. Also please be reminded that you must notify us of any amendments and the study closure.

Please note that the NHS organisation is required to monitor research to ensure compliance with the Research Governance Framework and other legal and regulatory requirements. This is achieved by random audit of research.

I wish you well in your research. Please do not hesitate to contact us should you need any guidance or assistance.

Yours sincerely

Version 3, 26.03.2012

Page 1 of 2

INSURANCE LETTER PROVIDING PROFESSIONAL INDEMNITY INSURANCE FROM UNIVERSITY OF LEICESTER

Our Ref: sdb 2013-2014 – 214

23rd April 2014

University of Leicester
ESTATES AND FACILITIES MANAGEMENT DIVISION
Fielding Johnson Building
University Road
Leicester
LE1 7RH
Tel: +44 (0)116 229 7631
Fax: +44(0)116 229 7633

To whom it may concern,

UNIVERSITY OF LEICESTER CLINICAL TRIAL/PROFESSIONAL INDEMNITY INSURANCE

Title of Study: Modelling care pathways in neonatal care: costs and consequences for the future

Chief Investigator: Sarah Seaton

I confirm that the University of Leicester will provide Clinical Trials and Professional Indemnity insurance cover in respect of its legal liability in relation to the above trial within the UK only.

Any significant departure from the programme of research as outlined in the application (such as changes in methodological approach, large delays in commencement of research, additional forms of data collection or major expansions in sample size) must be communicated to us.

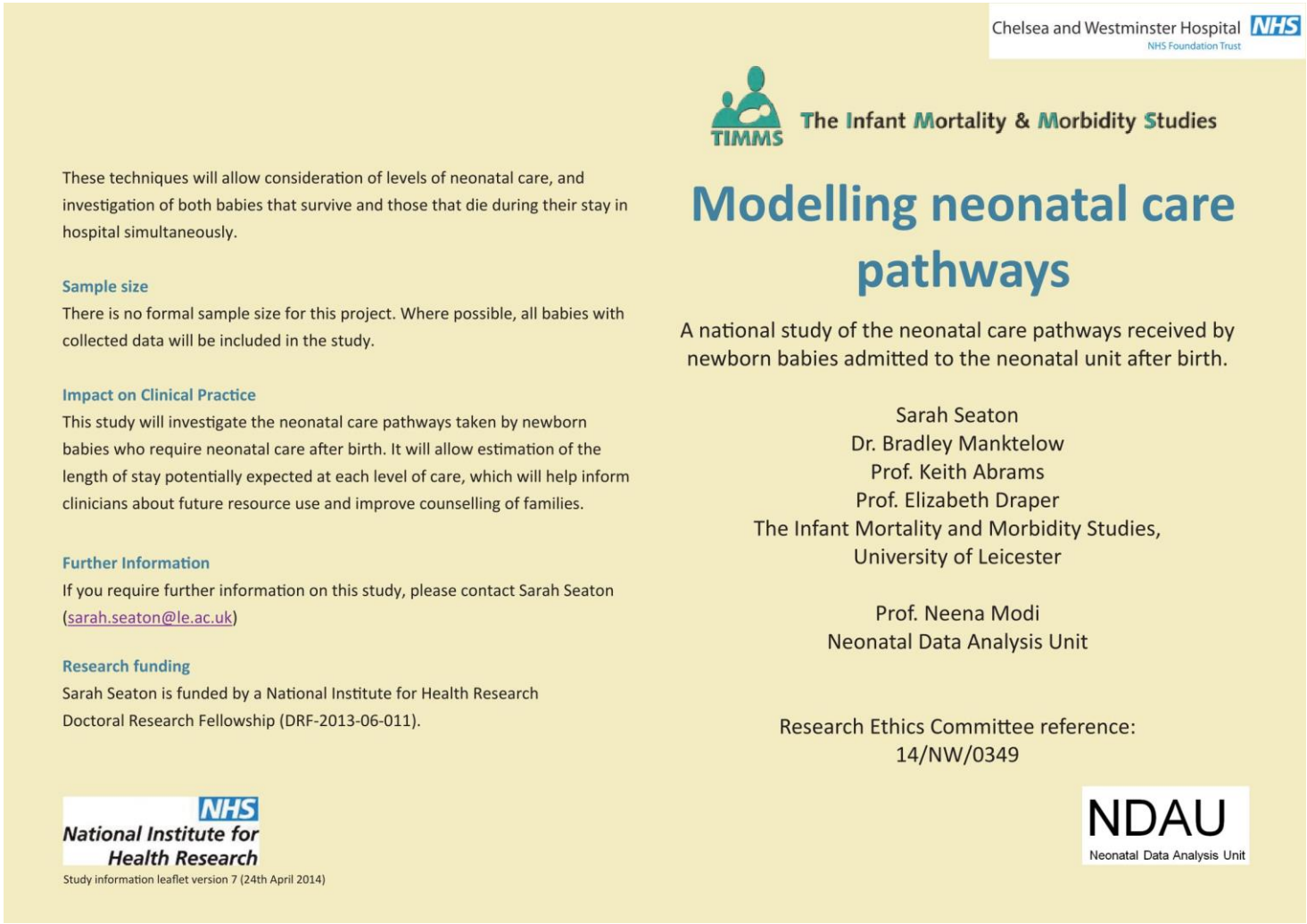
The cover is provided subject to normal policy terms and conditions.

Sue Banbury

Sue Banbury
Insurance & Risk Manager
University of Leicester

Sdb16@le.ac.uk

STUDY INFORMATION LEAFLET PROVIDED TO NEONATAL UNITS



Chelsea and Westminster Hospital NHS
NHS Foundation Trust

TIMMS The Infant Mortality & Morbidity Studies

Modelling neonatal care pathways

A national study of the neonatal care pathways received by newborn babies admitted to the neonatal unit after birth.

Sarah Seaton
Dr. Bradley Manktelow
Prof. Keith Abrams
Prof. Elizabeth Draper
The Infant Mortality and Morbidity Studies,
University of Leicester

Prof. Neena Modi
Neonatal Data Analysis Unit

Research Ethics Committee reference:
14/NW/0349

These techniques will allow consideration of levels of neonatal care, and investigation of both babies that survive and those that die during their stay in hospital simultaneously.

Sample size

There is no formal sample size for this project. Where possible, all babies with collected data will be included in the study.

Impact on Clinical Practice

This study will investigate the neonatal care pathways taken by newborn babies who require neonatal care after birth. It will allow estimation of the length of stay potentially expected at each level of care, which will help inform clinicians about future resource use and improve counselling of families.

Further Information

If you require further information on this study, please contact Sarah Seaton (sarah.seaton@le.ac.uk)

Research funding

Sarah Seaton is funded by a National Institute for Health Research Doctoral Research Fellowship (DRF-2013-06-011).

NHS National Institute for Health Research

NDAU Neonatal Data Analysis Unit

Study information leaflet version 7 (24th April 2014)

Aim

To investigate the neonatal care pathways taken by babies admitted to a neonatal unit after birth.

Objectives

1. To determine the length of stay, at the various levels of care, within the neonatal unit following birth. This will be investigated for babies that survive to discharge or die during their neonatal stay.

2. To compare neonatal networks to investigate whether there are differences in care between networks which treat similar groups of babies.

3. To undertake a preliminary health economics analysis to investigate methods for allocating costs within neonatal care.

The National Neonatal Research Database (NNRD)

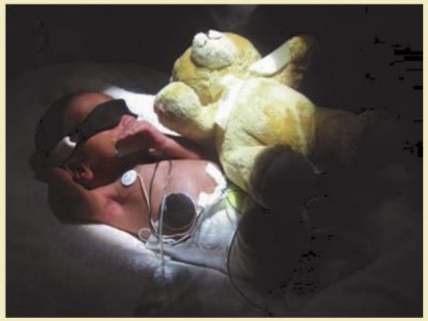
The National Data Analysis Unit (NDAU) is an independent academic unit based at the Chelsea & Westminster campus of Imperial College London. NDAU receives electronic data from contributing neonatal units and has permission to hold these in a National Neonatal Research Database (NNRD). The NNRD is a resource for research and service evaluations to improve newborn care. Contributing neonatal units form the UK Neonatal Collaborative. The National Research Ethics Service (ref 10/H0803/151) and the Ethics & Confidentiality Committee of the National Information Governance Board (ref ECC 8-05(f)/2010) have approved the use of the NNRD for NHS service evaluations and research.

Donovan James Lynch, born at 31 weeks on 15/03/2010. Photos reproduced with kind permission of his father, Jim Lynch.

Methods

Babies born from 01/01/2010
onwards will be identified from the NNRD. Anonymised data on all babies will be extracted.

Information will be extracted on daily events throughout the babies' time in hospital, in particular the level of neonatal care received on each day. Anonymised information related to the baby and the parents will be used.



No additional data collection will be required as part of this study. A full list of variables being extracted is available upon request.

Within neonatal medicine there are three broad levels of care as defined by the British Association of Perinatal Medicine (BAPM): intensive, high dependency and standard care. This research aims to estimate the length of time a baby will spend at each of these levels of care and the total amount of time they will spend in hospital.

Survival analysis is an area of statistics which estimates the amount of time which passes before an event (usually death) occurs. An extension of this method is used in this project: multistate modelling. This method also calculates the amount of time which passes before an event occurs, but it allows consideration of what happens before the event, for example, the types of care received. It also allows the consideration of multiple potential outcomes, for example death during hospital stay or survival to discharge.

LETTER AND FORM OF AGREEMENT TO PARTICIPATE IN THIS STUDY

Dear Colleagues,

MODELLING NEONATAL CARE PATHWAYS: COSTS AND CONSEQUENCES FOR THE FUTURE

(Research Ethics Committee Reference: 14/NW/0349)

I am writing to all lead clinicians who are members of the UK Neonatal Collaborative to invite them to participate in this research study jointly undertaken between The Infant Mortality and Morbidity Studies (TIMMS) with Prof. Elizabeth Draper and Dr. Bradley Manktelow, and the Neonatal Data Analysis Unit.

This research project aims to investigate the care pathways which babies take throughout their neonatal stay, considering the different levels of care they can receive along the way: intensive care, high dependency and standard care amongst others. It will investigate time spent at the different levels of care. Differences in care pathways will be investigated between neonatal networks. Finally, a preliminary health economics analysis will be undertaken. The research is funded by the National Institute for Health Research, and its findings will form part of my PhD in Medical Statistics.

We are seeking your agreement to extract information from the National Neonatal Research Database on babies admitted to your neonatal unit from 1st January 2011 until 31st December 2016. We will use data related to the baby, the care the baby received and demographic data about the parents. No identifiable data will be needed and you will not be required to provide any additional data or obtain R&D site approval for this project.

For more information about this project please see the enclosed leaflet. Printed copies of the leaflet are available upon request and a copy of the Research Ethics approval is also included.

Please sign and return the attached form to me by [INSERT DATE] at the latest. If you have queries, please contact me, contact details can be found overleaf.

Yours sincerely,

Sarah Seaton

AGREEMENT

MODELLING NEONATAL CARE PATHWAYS: COSTS AND CONSEQUENCES FOR THE FUTURE

(Research Ethics Committee Reference: 14/NW/0349)

☐ I am happy for data from the neonatal unit named below to be included in this study

☐ I DO NOT want data from the neonatal unit named below to be included in this study

Name of neonatal unit: _____

Name of NHS Trust: _____

Name of Lead Clinician: _____

Signature of Lead Clinician: _____

Date: _____

Authorship of publications arising from this study will be "[NAMED AUTHORS] and members of the UK Neonatal Collaborative". As this is a study using data held in an existing database, participation does not require NHS R&D permission from Trusts contributing data, but only from the NHS Trust holding the database, Chelsea & Westminster NHS Foundation Trust.

Please return this form by post or scanned email to:

Sarah Seaton

Department of Health Sciences, University of Leicester, 22-28 Princess Road West, Leicester, LE1 6TP

sarah.seaton@le.ac.uk

0116 2525434

LIST OF PARTICIPATING NEONATAL UNITS AND THE LEAD CLINICIAN

This thesis would not have been possible without the neonatal units who allowed their data to be used in this work.

Thank you to the Lead Clinicians of the UK Neonatal Collaborative: Dr Matthew Babirecki (Airedale General Hospital), Dr Liza Harry (Alexandra Hospital), Dr Oliver Rackham (Arrowe Park Hospital), Dr Tim Wickham (Barnet Hospital), Dr Sanaa Hamdan (Barnsley District General Hospital), Dr Aashish Gupta (Basildon Hospital), Dr Ruth Wigfield (Basingstoke & North Hampshire Hospital), Dr L M Wong (Bassetlaw District General Hospital), Dr Anita Mittal (Bedford Hospital), Dr Julie Nycyk (Birmingham City Hospital), Dr Phil Simmons (Birmingham Heartlands Hospital), Dr Vishna Rasiah (Birmingham Women's Hospital), Dr Sunita Seal (Bradford Royal Infirmary), Dr Ahmed Hassan (Broomfield Hospital, Chelmsford), Dr Karin Schwarz (Calderdale Royal Hospital), Dr Mark Thomas (Chelsea & Westminster Hospital), Dr Ainyne Foo (Chesterfield & North Derbyshire Royal Hospital), Dr Aravind Shastri (Colchester General Hospital), Dr Graham Whincup (Conquest Hospital), Dr Stephen Brearey (Countess of Chester Hospital), Dr John Chang (Croydon University Hospital), Dr Khairy Gad (Cumberland Infirmary), Dr Abdul Hasib (Darent Valley Hospital), Dr Mehdi Garbash (Darlington Memorial Hospital), Dr Nicci Maxwell (Derriford Hospital), Dr David Gibson (Dewsbury & District Hospital), Dr Pauline Adiotomre (Diana Princess of Wales Hospital), Dr Jamal S Ahmed (Doncaster Royal Infirmary), Dr Abby Deketelaere (Dorset County Hospital), Dr Ramnik Mathur (Ealing Hospital), Dr K Abdul Khader (East Surrey Hospital), Dr Ruth Shephard (Epsom General Hospital), Dr Abdus Mallik (Frimley Park Hospital), Dr Belal Abuzgia (Furness General Hospital), Dr Mukta Jain (George Eliot Hospital), Dr Simon Pirie (Gloucester Royal Hospital), Dr Phil Simmons (Good Hope Hospital), Dr Stanley Zengeya (Great Western Hospital), Dr Timothy Watts (Guy's & St Thomas' Hospital), Dr C Jampala (Harrogate District Hospital), Dr Cath Seagrave (Hereford County Hospital), Dr Michele Cruwys (Hillingdon Hospital), Dr Hilary Dixon (Hinchingbrooke Hospital), Dr Narendra Aladangady (Homerton Hospital), Dr Hassan Gaili (Hull Royal Infirmary), Dr Matthew James (Ipswich Hospital), Dr M Lal (James Cook University Hospital), Dr Ambadkar (James Paget Hospital), Dr Patti Rao (Kettering General Hospital), Dr Khalid Mannan (King George Hospital), Dr Ann Hickey (King's

College Hospital), Dr Dhaval Dave (King's Mill Hospital), Dr Nader Elgharably (Kingston Hospital), Dr Meera Lama (Lancashire Women and Newborn Centre), Dr Lawrence Miall (Leeds Neonatal Service), Dr Jonathan Cusack (Leicester General Hospital), Dr Venkatesh Kairamkonda (Leicester Royal Infirmary), Dr Jayachandran (Leighton Hospital), Dr Kollipara (Lincoln County Hospital), Dr J Kefas (Lister Hospital), Dr Bill Yoxall (Liverpool Women's Hospital), Dr Jennifer Birch (Luton & Dunstable Hospital), Dr Gail Whitehead (Macclesfield District General Hospital), Dr Bashir Jan Muhammad (Manor Hospital), Dr Aung Soe (Medway Maritime Hospital), Dr I Misra (Milton Keynes General Hospital), Dr Tilly Pillay (New Cross Hospital), Dr Imdad Ali (Newham General Hospital), Dr Mark Dyke (Norfolk & Norwich University Hospital), Dr Michael Selter (North Devon District Hospital), Dr Nagesh Panasa (North Manchester General Hospital), Dr Lesley Alsford (North Middlesex University Hospital), Dr Alan Fenton (North Tyneside General Hospital), Dr Subodh Gupta (Northampton General Hospital), Dr Richard Nicholl (Northwick Park Hospital), Dr Steven Wardle (Nottingham Neonatal Service), Dr Tim McBride (Ormskirk District General Hospital), Dr Naveen Shettihalli (Oxford University Hospitals, Horton Hospital), Dr Eleri Adams (Oxford University Hospitals, John Radcliffe Hospital), Dr Seif Babiker (Peterborough City Hospital), Dr Margaret Crawford (Pilgrim Hospital), Dr David Gibson (Pinderfields General Hospital), Dr Minesh Khashu (Poole General Hospital), Dr Caitlin Toh (Princess Alexandra Hospital), Dr M Hall (Princess Anne Hospital), Dr P Amess (Princess Royal Hospital), Dr Elizabeth Sleight (Princess Royal University Hospital), Dr Charlotte Groves (Queen Alexandra Hospital), Dr Sunit Godambe (Queen Charlotte's Hospital), Dr Dennis Bosman (Queen Elizabeth Hospital, Gateshead), Dr Barbara Piel (Queen Elizabeth Hospital, King's Lynn), Dr Banjoko (Queen Elizabeth Hospital, Woolwich), Dr N Kumar (Queen Elizabeth the Queen Mother Hospital), Dr A Manzoor (Queen's Hospital, Burton on Trent), Dr Wilson Lopez (Queen's Hospital, Romford), Dr Angela D'Amore (Rosie Maternity Hospital, Addenbrookes), Dr Shameel Mattara (Rotherham District General Hospital), Dr Christos Zipitis (Royal Albert Edward Infirmary), Dr Peter De Halpert (Royal Berkshire Hospital), Dr Paul Settle (Royal Bolton Hospital), Dr Paul Munyard (Royal Cornwall Hospital), Dr Gitika Joshi (Royal Derby Hospital), Dr David Bartle (Royal Devon & Exeter Hospital), Dr D Schapira (Royal Hampshire County Hospital), Dr Joanne Fedee (Royal Lancaster Infirmary), Dr Natasha Maddock (Royal

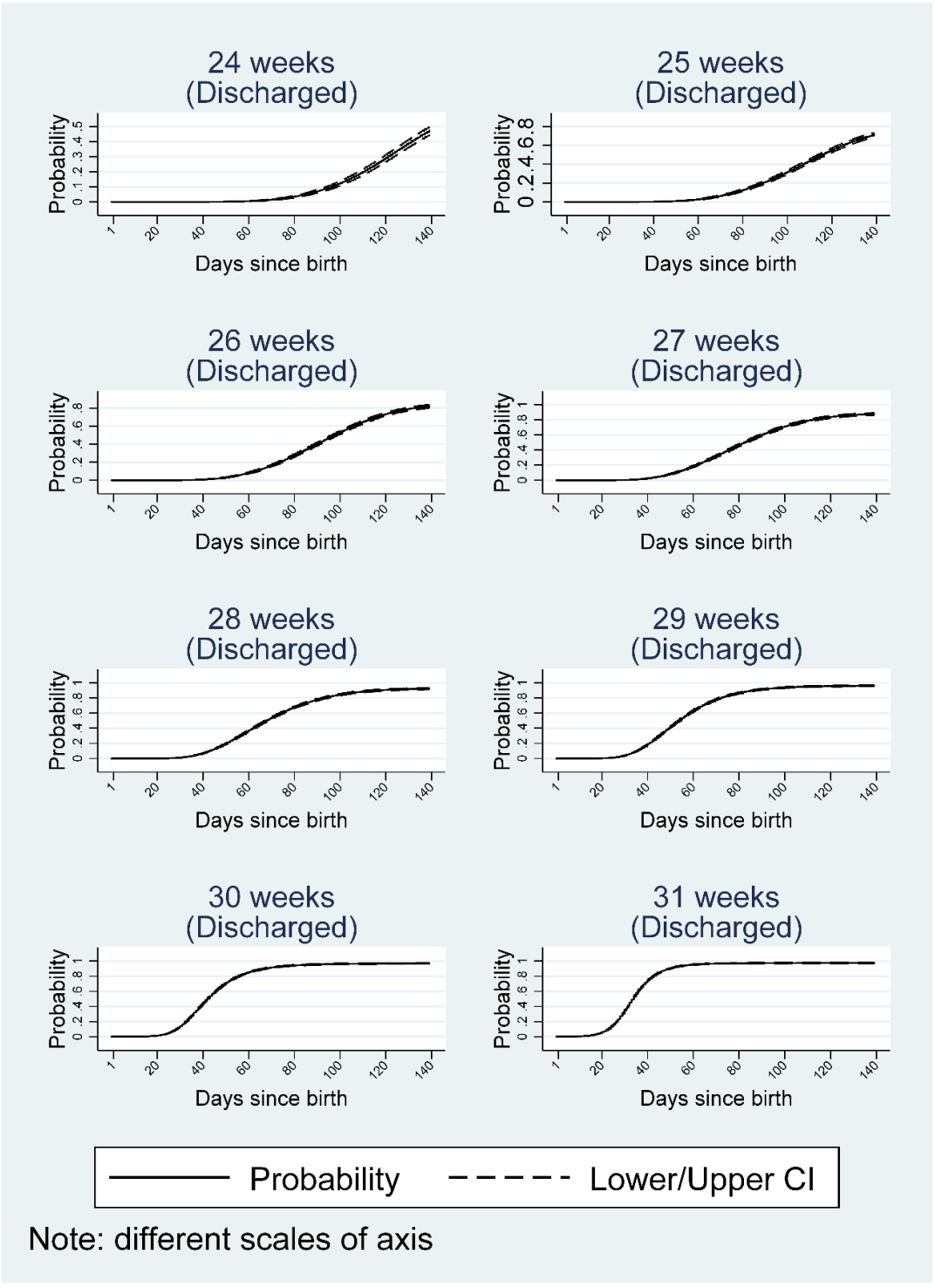
Oldham Hospital), Dr Richa Gupta (Royal Preston Hospital), Dr Deshpande (Royal Shrewsbury Hospital), Dr Charles Godden (Royal Surrey County Hospital), Dr P Amess (Royal Sussex County Hospital), Dr Stephen Jones (Royal United Hospital), Dr Alan Fenton (Royal Victoria Infirmary), Dr Mahadevan (Russells Hall Hospital), Dr Nick Brown (Salisbury District Hospital), Dr Kirsten Mack (Scarborough General Hospital), Dr Pauline Adiotomre (Scunthorpe General Hospital), Dr Rob Bolton (South Tyneside District Hospital), Dr A Khan (Southend Hospital), Dr Paul Mannix (Southmead Hospital), Dr Charlotte Huddy (St George's Hospital), Dr Salim Yasin (St Helier Hospital), Dr Sian Butterworth (St Mary's Hospital, Isle of Wight), Dr Sunit Godambe (St Mary's Hospital, London), Dr Ngozi Edi-Osagie (St Mary's Hospital, Manchester), Dr Bala Thyagarajan (St Michael's Hospital), Dr Peter Reynolds (St Peter's Hospital), Dr Nick Brennan (St Richard's Hospital), Dr Carrie Heal (Stepping Hill Hospital), Dr Sanjay Salgia (Stoke Mandeville Hospital), Dr Majd Abu-Harb (Sunderland Royal Hospital), Dr Jacqeline Birch (Tameside General Hospital), Dr Chris Knight (Tameside General Hospital), Dr Simon Clark (The Jessop Wing, Sheffield), Dr V Van Sommen (The Royal Free Hospital), Dr Nandiran Ratnavel (The Royal London Hospital, Constance Green), Dr Mala Raman (Torbay Hospital), Dr Hamudi Kisat (Tunbridge Wells Hospital), Dr Sara Watkin (University College Hospital), Dr Kate Blake (University Hospital Coventry), Dr Jauro Kuna (University Hospital Lewisham), Dr Mehdi Garbash (University Hospital of North Durham), Dr Alison Moore (University Hospital of North Staffordshire), Dr Hari Kumar (University Hospital of North Tees), Dr Gopi Vemuri (University Hospital of South Manchester), Dr Chris Rawlingson (Victoria Hospital, Blackpool), Dr Delyth Webb (Warrington Hospital), Dr Bird (Warwick Hospital), Dr Sankara Narayanan (Watford General Hospital), Dr Jason Gane (West Cumberland Hospital), Dr Elizabeth Eyre (West Middlesex University Hospital), Dr Ian Evans (West Suffolk Hospital), Dr Rekha Sanghavi (Wexham Park Hospital), Dr Caroline Sullivan (Whipps Cross University Hospital), Dr Laweh Amegavie (Whiston Hospital), Dr Wynne Leith (Whittington Hospital), Dr Vimal Vasu (William Harvey Hospital), Dr Andrew Gallagher (Worcestershire Royal Hospital), Dr Katia Vamvakiti (Worthing Hospital), Dr Megan Eaton (Yeovil District Hospital), Dr Guy Millman (York District Hospital).

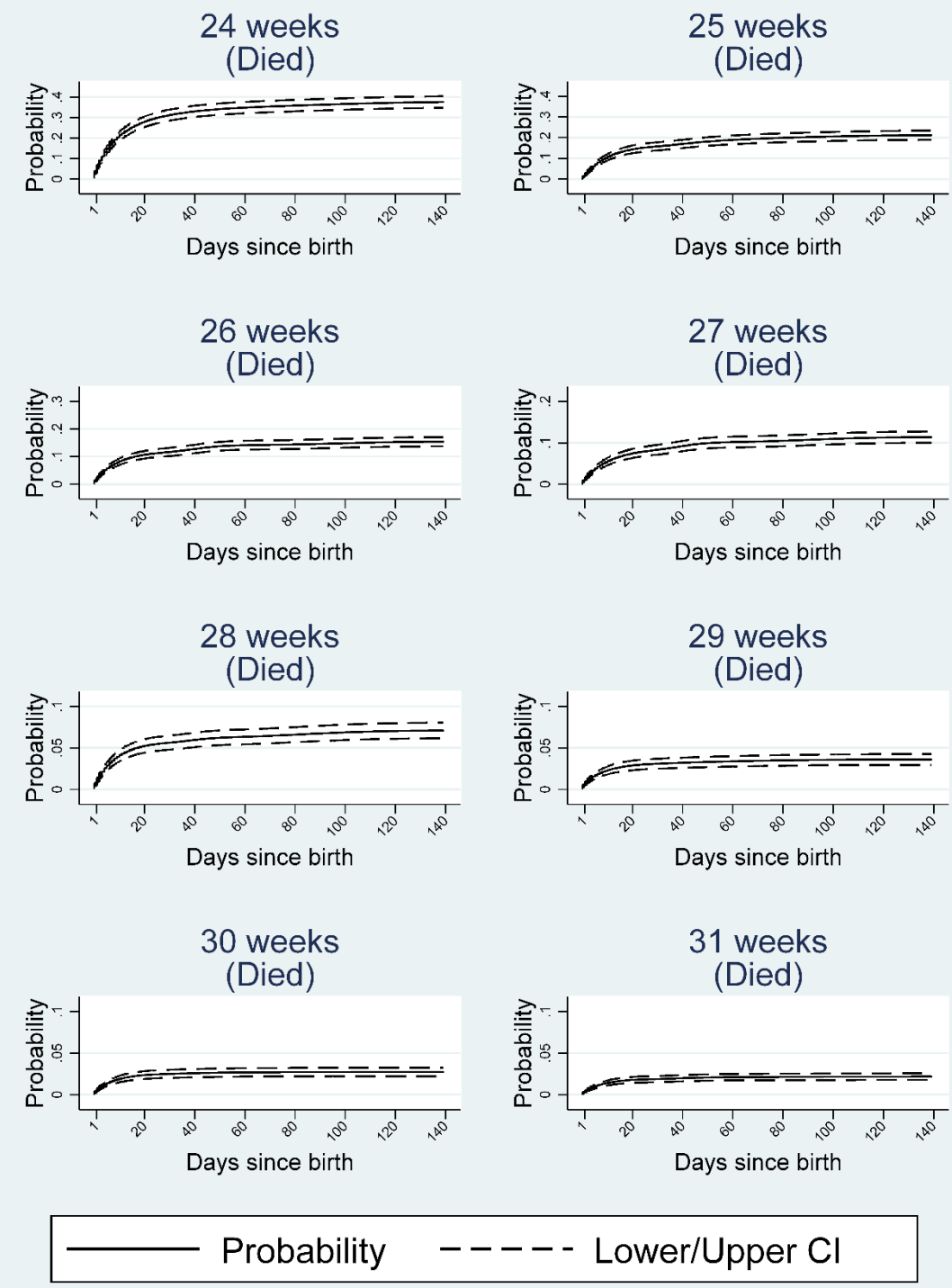
APPENDIX 4

There are no Appendices for Chapter Four.

APPENDIX 5

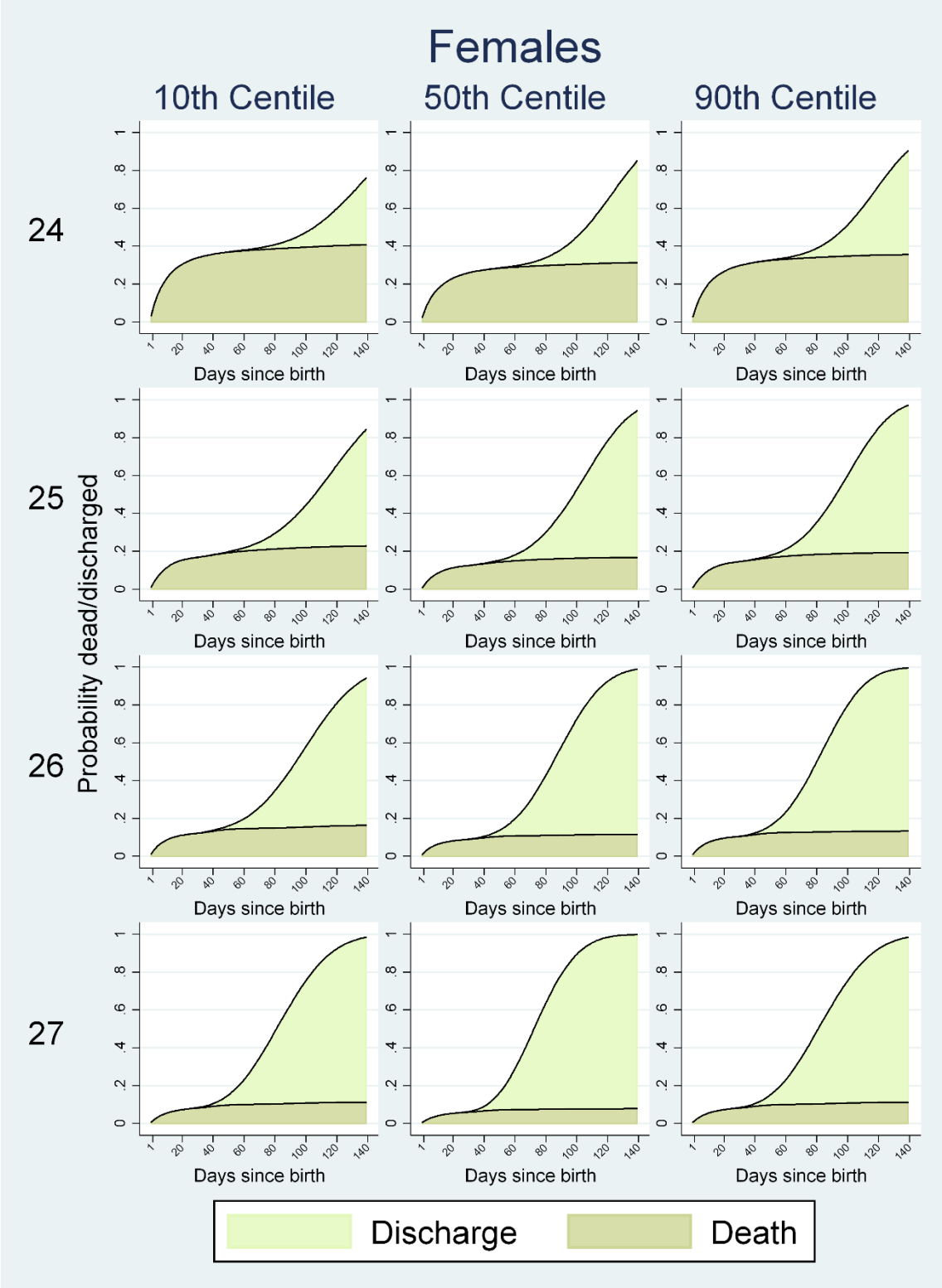
CONFIDENCE INTERVALS FOR THE CUMULATIVE INCIDENCE FUNCTION FOR DEATHS AND DISCHARGE

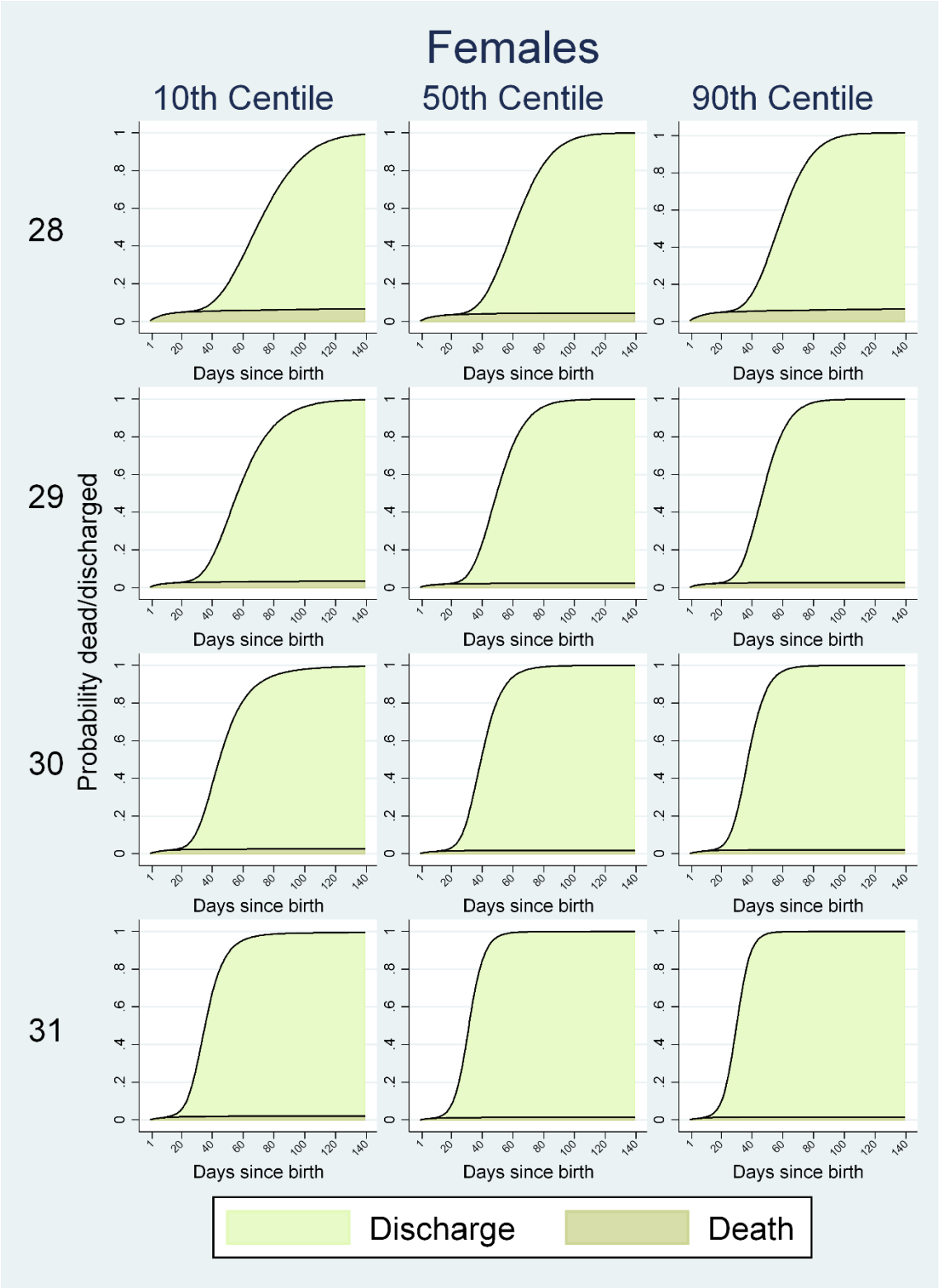




Note: different scales of axis

STACKED CUMULATIVE INCIDENCE PLOTS FOR FEMALES




Females
10th Centile
50th Centile
90th Centile
28
29
30
31
Probability dead/discharged
0
.2
.4
.6
.8
1
1
20
40
60
80
100
120
140
Days since birth
Discharge
Death

APPENDIX 6

MODELLING NEONATAL CARE PATHWAYS: INVESTIGATING LENGTH OF STAY FOR PRETERM BABIES – PUBLISHED IN INFANT

Modelling neonatal care pathways: investigating length of stay for preterm infants

An overall estimate of total length of stay does not fully describe the complex journey, involving different treatments and care, that a preterm infant experiences during its time in hospital. This article presents preliminary results using a statistical approach to account for the levels of care, and the competing outcomes of death or discharge from neonatal care. Detailed information about length of stay in neonatal care is important for counselling families and could be used to aid calculating the true cost of neonatal care, informing neonatal service provision, commissioning and funding.

Sarah E. Seaton
BSc, MSc
Medical Statistician, The Infant Mortality and Morbidity Studies, University of Leicester
sarah.seaton@leicester.ac.uk

Lisa Barker
BMedSci, BMBS, MRCPCH
Neonatal Senior Clinical Fellow, Leicester Neonatal Unit, Leicester Royal Infirmary

Keywords

neonatal care; intensive care; length of stay; mortality; preterm

Key points

Seaton S.E., Barker L. Modelling neonatal care pathways: investigating length of stay for preterm infants. *Infant* 2016; 12(3): 87-90.

1. Predicting length of stay in the neonatal unit is difficult.
2. Previous estimates have not considered the levels of care an infant requires, nor accounted for infants who die.
3. An approach that provides detailed information about length of stay is described.
4. More accurate estimates of length of stay will be useful for counselling parents and commissioning services.

Following birth, around one in eight infants require specialist neonatal care.[1] These infants have varying needs with some requiring simply a few hours of observation, while others require ventilation and full intensive care, with a length of stay of potentially several months. The most preterm, particularly those born at less than 32 weeks' gestation, are the smallest and sickest babies with the highest risk of mortality.[2] While survival rates have improved dramatically over the last 20 years, in-unit mortality remains high in infants born at the earliest gestational ages.[3] For infants that survive, their stay in hospital is likely to be long, lasting weeks to months and requiring the highest level of care. Some of these babies will experience long-lasting effects of their preterm birth including neurodevelopmental problems,[4] and life-long morbidities.[5]

The impact of a having an infant requiring specialist neonatal care on parents and families should not be underestimated. "When can my baby come home?" is a frequently asked question by parents to healthcare professionals at all points during their child's care. Answering this question can be difficult and a prediction of the time a preterm infant is likely to spend in hospital is hard to provide. Rough estimates are commonly given such as: "They'll be home by their due date" or: "They'll go home once they can keep themselves warm and feed." It is likely that this is not far from reality, with most infants discharged home around or after their due date. However, none of this captures the complex journey preterm infants encounter in the neonatal unit, including a potentially high risk of mortality, and the different types and levels of care that will be received. More detailed information could be useful for counselling parents and preparing them for the long journey ahead. Estimates of length of stay and the pathway through different levels of care would also be valuable to individual units and networks, for calculating the true cost of neonatal care for commissioning and allocation of resources including staffing.

Background

What is already known about length of stay?

Limited research has investigated how to predict length of stay in the neonatal unit for infants born preterm. Previous studies have produced overall estimates of total length of stay.[6-10] Providing overall estimates of length of stay does not fully describe the complicated journey, involving different treatments and care, that preterm infants experience during their time in hospital.

Infants that die while receiving neonatal care are often excluded from length of stay analyses,[6-8] or they are included in a way

that does not take account of the fact they have died.[9] These infants are important to include in research to help counsel and care for families and excluding them underestimates costs and service provision to neonatal services caring for preterm infants.

Levels of neonatal care

Levels of neonatal care are defined by the British Association of Perinatal Medicine (BAPM) and relate to the treatment received rather than the physical location, or characteristics of the infant.[11] BAPM guidelines provide a recommendation for the staff to infant ratio required for each level of care. However, despite this, a recent report by Bliss highlighted the issues of providing neonatal care, most notably a lack of nurses, to provide the level of support recommended by NHS England.[12] More detailed information about the pathway for preterm infants through different levels of neonatal care could be useful to neonatal services to support service provision and staffing.

What predicts length of stay?

Length of stay is complex and little research has investigated what predicts it. The limited research that has been undertaken in this area highlights the information available on the first day of life (eg gestational age, birth weight and sex) that can be very useful for predicting length of stay. Anecdotally, we know that infants are rarely discharged home before they have gained the ability to suck and feed (around 34-35 weeks corrected gestational age) and most go home later than this. Irrespective of the clinical conditions they have experienced, it is likely that most preterm infants will have matured and recovered enough to be discharged between this point and their due date, their prematurity having been the overwhelming reason for their long length of stay.

Some infants may experience specific late occurring conditions (eg sepsis), which may add dramatically to their length of stay or result in death. Therefore estimates of the discharge date will need to be adjusted in light of this. However, although more likely in infants of the smallest gestational ages, whether an infant is going to experience such complications cannot be identified at birth and so any initial prediction of length of stay will need to be revised if appropriate.

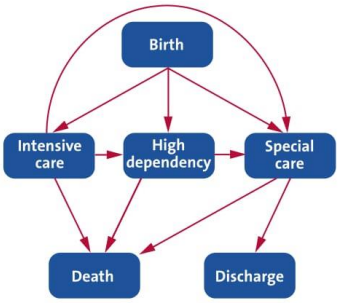


FIGURE 1 An example of a multistate model used in the analysis. Each red arrow denotes a potential movement along the neonatal care pathway.

Aims

This project represents a collaboration between the University of Leicester and the Neonatal Data Analysis Unit (NDAU) at Imperial College, London, and it is funded by the National Institute for Health Research (NIHR).

The project has two main aims; firstly to determine the length of stay within the neonatal unit for all preterm infants, investigating the different levels of care required by an infant following birth irrespective of whether the infant survives to discharge or not. Secondly, it compares different operational delivery networks (ODNs) to investigate whether any differences exist between them. This second aim is not discussed in this article.

Method

The project focuses on singleton infants born at 24-31 weeks' gestation. Preliminary results are presented in this article using information about infants that were admitted to an English neonatal unit on the first day of life and were discharged in 2014.

Data about these infants were taken from the National Neonatal Research Database (NNRD). Infants with unusual care, for example those discharged home before 34 weeks corrected gestational age, were excluded. All 162 neonatal units that existed in England, and contributed data to the NNRD, were approached in 2014 to ask their permission to use their data and 100% of neonatal units agreed to participate. No additional data were needed and neonatal units were not required to do any additional work.

The levels of care investigated were intensive care, high dependency care and special care. Other types of care exist, most notably transitional care. However, this is not offered routinely across the country and so here it is amalgamated with special care. It is also possible for infants to receive care in non-neonatal services, such as cardiac specialist centres or surgical units. When an infant receives multiple levels of care on one day, the highest level of care is used for commissioning purposes.

Statistical analysis

The statistical analysis used in this project is known as multistate modelling, which is an extension of standard survival analysis, using the Cox model. A standard survival analysis allows the measuring of time until an event, usually death, occurs. The extension of this to a multistate model allows the measuring of time until a variety of events occur. **FIGURE 1** provides an example of the pathways considered in this analysis. Statistical details about this approach can be found elsewhere.[13] The boxes denote 'states' that an infant can be in, while the arrows are 'transitions' through the neonatal care pathway. Some states are possible to enter but not exit again, for example death or discharge from neonatal care. Other states are intermediate and it is possible to enter and exit them, for example the levels of care. Certain movements through neonatal care are very unusual, for example being discharged immediately from intensive care, and so these are not considered here.

An infant can take a variety of pathways through neonatal care. For example, following birth a preterm infant could be admitted to the neonatal unit and be put on to a ventilator for several weeks (intensive care). With improvement, the respiratory support could be reduced to continuous positive airway pressure (CPAP) and the infant would move from the intensive care state to the high dependency state. Finally this infant could improve further and need a few days of phototherapy (special care) before being discharged from neonatal care and reaching the final state of 'discharge,' which it will not exit again. It is possible to measure the time between these different movements along the neonatal care pathway and predict the probability of being in any of the states over time.

Results

In 2014, approximately 5,000 singleton infants born at 24-31 weeks' gestation were discharged from neonatal care or died

during their stay in a unit in England that submitted data to the NNRD. The infants who survived to discharge from neonatal care often had a very long length of stay before being discharged, while the infants who died often only survived a short period of time. The small number of surviving infants discharged in the initial days of life were treated in other specialist services and so their length of stay is longer than the time spent in neonatal care, although this is not considered here. **FIGURE 2** demonstrates the differences in the distributions of length of stay between infants who survive and those who do not.

Of the days of care given to infants discharged in 2014, approximately:

- 61,000 were intensive care (20%)
- 85,000 were high dependency (29%)
- 149,000 were special care (51%).

While this group are the most preterm infants who have the highest need and are the sickest infants, it is worth noting that the majority of the total care is still special care.

FIGURE 3 shows the percentage of infants in each category of care in the days following their birth. This analysis has not been adjusted for any characteristics of the infant. On the first day of life:

- 84% of infants were receiving intensive care
- 12% were receiving high dependency
- 4% were receiving special care.

It is possible to read any point on the plot to see the percentage of infants in each category. For example, by day ten of life, 42% of infants are receiving intensive care; 26% are receiving high dependency; 28% are receiving special care and 4% of infants have died.

As time continues, the area of the curve related to death plateaus at the point that is the average mortality percentage across all infants born at 24-31 weeks' gestation (around 7%) and the percentage being discharged home increases until 150 days when nearly all infants have been discharged home or died. The graph does not describe beyond 150 days.

Of course, in reality the need for the different levels of care described here varies drastically depending on the gestational age of the infant and potentially other characteristics, including birth weight and sex. As would be expected, the percentage of infants that require intensive care is much higher the lower the gestational age. Similarly, the risk of death is much higher in infants born at the earlier weeks of gestation. This can be seen more clearly in **FIGURE 4**, which shows the length of stay and outcome for infants born at 24-28 weeks' gestation and 29-31 weeks' gestation. The percentage of infants in these lower gestational ages that require higher levels of intensive care is much higher, seen by the larger area on the graph. The percentage of infants that die in the 24-28 week group is also higher, with the percentage of death plateauing at around 14% compared to 3% in the infants born at 29-31 weeks' gestation.

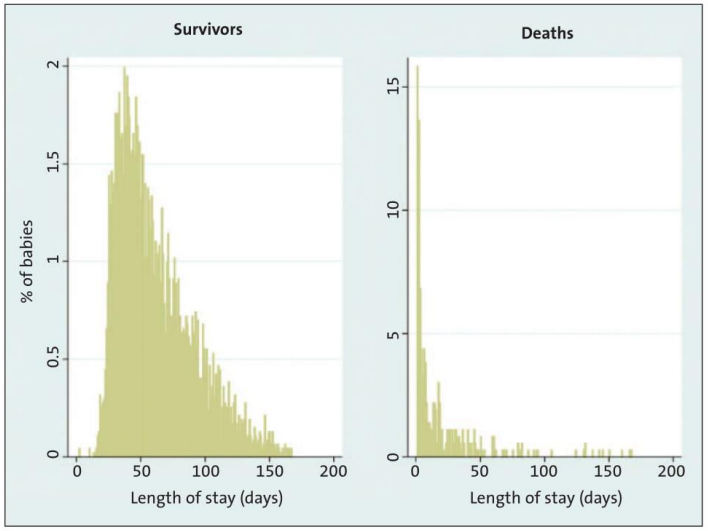


FIGURE 2 Length of stay of all infants. A comparison of the length of stay of infants that survived to discharge and the length of stay of infants that died during neonatal care. Note there is a different scale on each y axis.

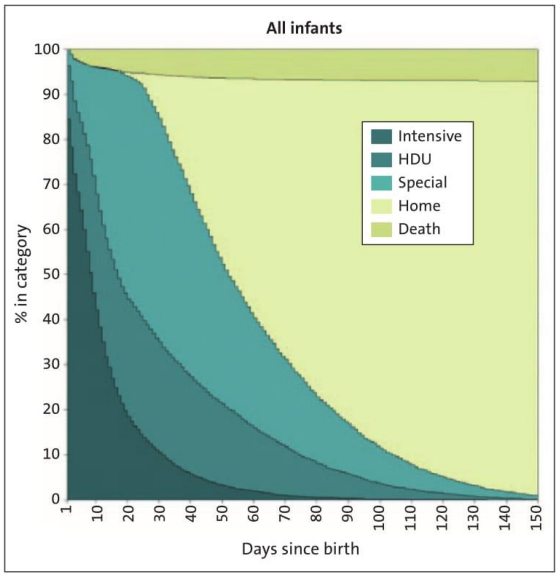


FIGURE 3 All infants and the level of care received and outcome (death or discharge) over time since birth. Key: HDU = high dependency unit.

Discussion

This article presents preliminary results using a statistical approach to account for the levels of care, and the competing outcomes of death or discharge from neonatal care. Estimating total length of stay alone does not provide a clear picture of the experiences an infant will have during their

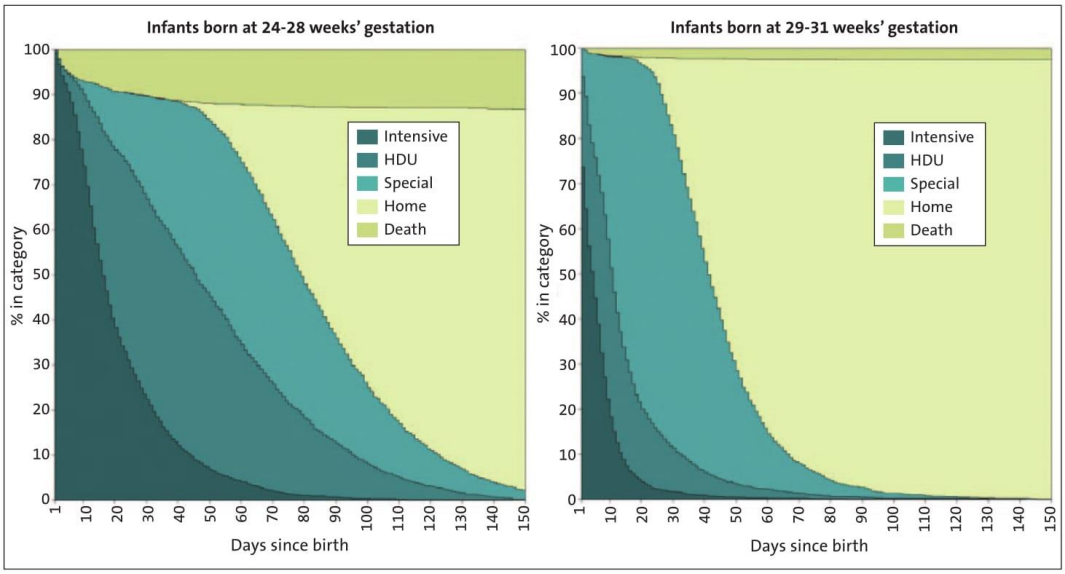


FIGURE 4 Infants divided into 24-28 weeks' gestation and 29-31 weeks' gestation and the level of care received and outcome (death or discharge) over time since birth.

neonatal stay. Detailed information about length of stay in neonatal care is important for counselling families and could be used to aid calculating the true cost of neonatal care, informing neonatal service provision, commissioning and funding.

The journey of an infant through different levels of neonatal care is complicated, and an infant may move up and down different levels of care on different occasions and at different times. The approach discussed in the model here is thus an over-simplification of reality as it assumes that care is received in a hierarchy, ie that all intensive care occurs before all high dependency, which occurs before special care. However, this work is still informative as long as it is interpreted in that context. No adjustments for non-hierarchical movement, or characteristics of the infant, have been made in the analysis presented here. Subsequent work will explore this in more detail including the infant's journey through different levels of care, length of stay and outcome related to gestational age and the presence of different complications and morbidities. Approaches for presenting this in a way to aid communication with healthcare providers and families will also be explored.

Updates of this work will be produced on an ongoing basis; contact sarah.seaton@leicester.ac.uk.

Acknowledgement

The authors would like to thank all the neonatal units in England that agreed for the data they submit to the NNRD to be used as part of this study. They would also like to thank the other members of the study team: Bradley Manktelow, Elizabeth Draper, Keith Abrams and Neena Modi.

Funding

Sarah Seaton is funded by a NIHR Doctoral Fellowship Award. This article presents independent research funded by the NIHR. The views expressed are those of the authors and not necessarily those of the NHS, NIHR or the Department of Health.

References

1. **NNAP.** *National Neonatal Audit Programme Annual Report 2015.* RCPCH. 2015: London.
2. **Manktelow B.N., Seaton S.E., Field D.J., Draper E.S.** Population-based estimates of in-unit survival for very preterm infants. *Pediatrics* 2013;131:e425-32.
3. **Field D.J., Dorling J.S., Manktelow B.N., Draper E.S.** Survival of extremely premature babies in a geographically defined population: prospective cohort study of 1994-9 compared with 2000-5. *Br Med J* 2008;336:1221-223.
4. **Allen M.C.** Neurodevelopmental outcomes of preterm infants. *Curr Opin Neurol* 2008;21:123-28.
5. **Saigal S., Doyle L.W.** Preterm birth 3: An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet* 2008;371:261-69.
6. **Altman M., Vanpée M., Cnattingius S., Norman M.** Moderately preterm infants and determinants of length of hospital stay. *Arch Dis Child Fetal Neonatal Ed* 2009;94:F414-18.
7. **Lee H.C., Bennett M.V., Schulman J., Gould J.B.** Accounting for variation in length of NICU stay for extremely low birth weight infants. *J Perinatol* 2013;33:872-76.
8. **Manktelow B., Draper E.S., Field C., Field D.** Estimates of length of neonatal stay for very premature babies in the UK. *Arch Dis Child Fetal Neonatal Ed* 2010;95:F288-92.
9. **Bender G.J., Koestler D., Ombao H. et al.** Neonatal intensive care unit: predictive models for length of stay. *J Perinatol* 2013;33:147-53.
10. **Hinchliffe S.R., Seaton S.E., Lambert P.C. et al.** Modelling time to death or discharge in neonatal care: an application of competing risks. *Paediatr Perinat Epidemiol* 2013;27:426-33.
11. **British Association of Perinatal Medicine.** *BAPM Categories of Care.* [Online];2011. Available from: www.bapm.org/publications [Accessed 22 April 2016].
12. **Bliss.** *Bliss baby report 2015: hanging in the balance.* [Online]; 2015. Available from: www.bliss.org.uk/babyreport [Accessed 22 April 2016].
13. **de Wreede L.C., Fiocco M., Putter H.** The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. *Comput Methods Programs Biomed* 2010;99:261-74.

MODELLING NEONATAL CARE PATHWAYS FOR BABIES BORN PRETERM: AN APPLICATION OF MULTISTATE MODELLING – PUBLISHED IN PLOS ONE

PLOS | ONE

RESEARCH ARTICLE

Modelling Neonatal Care Pathways for Babies Born Preterm: An Application of Multistate Modelling

Sarah E. Seaton[1]*, Lisa Barker[2], Elizabeth S. Draper[1], Keith R. Abrams[1], Neena Modi[3], Bradley N. Manktelow[1], UK Neonatal Collaborative[¶]

1 Department of Health Sciences, University of Leicester, Leicester, United Kingdom, **2** Leicester Neonatal Unit, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom, **3** Neonatal Data Analysis Unit, Section of Neonatal Medicine, Department of Medicine, Imperial College, London, United Kingdom

¶ Membership of the UK Neonatal Collaborative is provided in the Acknowledgments.
* sarah.seaton@le.ac.uk

OPEN ACCESS

Citation: Seaton SE, Barker L, Draper ES, Abrams KR, Modi N, Manktelow BN, et al. (2016) Modelling Neonatal Care Pathways for Babies Born Preterm: An Application of Multistate Modelling. PLoS ONE 11(10): e0165202. doi:10.1371/journal.pone.0165202

Editor: Umberto Simeoni, Centre Hospitalier Universitaire Vaudois, FRANCE

Received: May 24, 2016

Accepted: October 7, 2016

Published: October 20, 2016

Data Availability Statement: The anonymised data used to produce this manuscript were obtained from a third party. Readers can request access to the data by contacting the Neonatal Data Analysis Unit (ndau@imperial.ac.uk).

Funding: SS is funded by a National Institute for Health Research Doctoral Research Fellowship award (DRF-2013-06-011). This paper presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the

Abstract

Modelling length of stay in neonatal care is vital to inform service planning and the counselling of parents. Preterm babies, at the highest risk of mortality, can have long stays in neonatal care and require high resource use. Previous work has incorporated babies that die into length of stay estimates, but this still overlooks the levels of care required during their stay. This work incorporates all babies, and the levels of care they require, into length of stay estimates. Data were obtained from the National Neonatal Research Database for singleton babies born at 24–31 weeks gestational age discharged from a neonatal unit in England from 2011 to 2014. A Cox multistate model, adjusted for gestational age, was used to consider a baby's two competing outcomes: death or discharge from neonatal care, whilst also considering the different levels of care required: intensive care; high dependency care and special care. The probabilities of receiving each of the levels of care, or having died or been discharged from neonatal care are presented graphically overall and adjusted for gestational age. Stacked predicted probabilities produced for each week of gestational age provide a useful tool for clinicians when counselling parents about length of stay and for commissioners when considering allocation of resources. Multistate modelling provides a useful method for describing the entire neonatal care pathway, where rates of in-unit mortality can be high. For a healthcare service focussed on costs, it is important to consider all babies that contribute towards workload, and the levels of care they require.

Introduction

In the UK, 1 in 8 babies require specialist neonatal care after their birth[1] and the needs of these babies can vary dramatically, both in clinical approach and the length of time they require in hospital. For the most preterm babies who survive, this care can last several months or longer. Understanding the time babies spend in neonatal care is vital to aid the counselling of

Department of Health. KA is partially supported by the National Institute for Health Research (NIHR) as a Senior Investigator (NF-SI-0512-10159). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

parents, plan care provision and ensure the appropriate funding of services. However, there are two potential problems which mean that such information has not been readily available.

First, there has been little research in modelling length of stay in neonatal care, and the research that does exist has understandably focussed on time to discharge for survivors,[2–4] or has considered all babies together irrespective of their outcome.[5] Excluding babies who die during their time in neonatal care overlooks this important group, who contribute to the workload of the health service during the time they are alive and should therefore be included in estimates of length of stay. In particular, babies born preterm have a high rate of in-hospital mortality, particularly for those born at <28 weeks gestation[6] and any analysis which is only based on survivors does not fully describe neonatal care provision or requirements. However, recent work has illustrated a statistical approach, 'competing risks', which can be used to appropriately include all babies when estimating the length of stay in neonatal care.[7]

Second, broad estimates of overall length of stay are of limited use as they provide no information about the type of care a baby requires during their stay in neonatal care. Within the UK, neonatal care is defined into three main levels, using criteria developed by the British Association of Perinatal Medicine (BAPM) in 2011.[8] The levels defined by BAPM are: intensive care (e.g. ventilation); high dependency care (e.g. drug infusion) or special care (e.g. phototherapy). This classification of care provision is similar to that seen in other countries.[9–11] A baby born very preterm is likely to need care provided at different levels throughout their time in hospital, and will move between these levels of care during their in-patient stay (we refer to this as the neonatal care pathway). Whilst other levels of neonatal care do exist, most notably transitional care,[12] these are not offered consistently.

Use of standard statistical approaches, which measure time until an event occurs, to measure length of stay, are unable to capture the complexity of the nature of neonatal care, whilst also modelling babies irrespective of their outcome. Recent advances such as competing risks have been used to model the length of stay for babies, and including those who die before discharge.[7] However, it is of interest to model the neonatal care pathway that a baby follows from birth, considering the different levels of care they require, before being discharged or dying in neonatal care, which previous work has not considered. Here we propose the use of a statistical method known as multistate modelling which allows the complete care pathway to be investigated for all babies: i.e. both those who survive to discharge and those who do not.

Methods

Data were extracted from the National Neonatal Research Database (NNRD) which holds data related to the daily neonatal care, demographics and outcomes of all babies admitted to neonatal units throughout England, Wales and Scotland. The NNRD is maintained by the Neonatal Data Analysis Unit (NDAU) and created from electronic health records. Only data from England was used in the analysis of this paper because full data collection was not available for Wales and Scotland for the years analysed.

Permission was granted to use anonymised data from the NNRD for research purposes (ethics reference: 14/NW/0349, North West—Lancaster ethics committee) and agreement was obtained from all neonatal units in England (n = 162 units) that existed in 2014 to use their data for this project.

All singleton babies born at 24–31 weeks gestational age that were discharged from a neonatal unit from 2011 to 2014 were included in this analysis. Babies born prior to 24 weeks gestational age were excluded as their care relates heavily to local policies.[13] Babies were not excluded for missing or ambiguous data on any other variables. Babies who were discharged home before 34 weeks corrected age were excluded by all analyses, as this is the point at which

babies begin to learn to suck and feed and a discharge home before this point is clinically implausible. Babies with a total length of stay greater than six months were also excluded as although some babies do stay in hospital this long, it is a rare occurrence and the numbers are small producing unreliable estimates. Babies were also excluded if they were discharged from neonatal care having only received intensive care or were discharged having never received special care (i.e. after receiving high dependency and potentially intensive care) as these care pathways are clinically unusual and there was insufficient statistical power to produce reliable estimates.

If a baby was not recorded as having died on a neonatal unit, they were considered to have survived to discharge. However, as the NNRD only holds data related to neonatal care, these babies might have died, or spent substantially periods of time, in other specialist services, such as paediatric wards or cardiac surgical centres. Therefore, two final outcomes were considered: death and discharge alive from the neonatal unit.

As transitional care is not offered consistently, this group has been amalgamated with special care in all analyses (the most similar level of care)[12]–to otherwise exclude this level of care would underestimate the total length of stay of many babies. In this analysis, all care is assumed to have occurred in a hierarchical manner, that is, all intensive care comes before high dependency care, and all high dependency care comes before special care. Certain transitions are clinically unlikely, although not impossible, for example, being discharged from neonatal care having only received intensive care and these are not considered in the analyses presented as the numbers experiencing them are very small and reliable estimates cannot be obtained.

Statistical analysis

A multistate model[14] was used to describe the time period from an initial event (in this case birth) until a final endpoint (i.e. death or discharge from neonatal care) whilst also considering intermediate changes in the level of care received. The multistate model used is described in Fig 1. Each of the boxes (known as states) describes a level of care in the neonatal care pathway that a baby can potentially experience. States are 'absorbing' if upon entry they are impossible to exit again (i.e. death and discharge from neonatal care). Alternatively, they are known as 'intermediate' states if it is possible to exit them and move to a different state (i.e. the different levels of neonatal care).

The arrows denoted in Fig 1 represent 'transitions' a baby can make through the neonatal care pathway. Here, a baby is born and, if considered viable, admitted to the neonatal unit where they are assumed to immediately receive the highest level of care they will require during their neonatal in-patient stay, i.e. intensive care, high dependency care or special care. This is considered to be day 1 of life, so all admitted babies spend at least one day receiving neonatal care.

A Cox proportional hazards model was stratified on transition (depicted by the arrows in Fig 1) and used to estimate the transition specific hazard rates[14,15] and was used to calculate the probabilities of receiving each of the levels of care, or of having died or been discharged by any given time. These probabilities represent the observed data, and are presented as percentages in a graphical form overall without adjustment for any covariates.

The probabilities were then estimated by adjusting for week of gestational age at birth, with 27 weeks gestational age as the baseline group. The care pathway was adjusted by gestational age at birth, as this is known to be an important predictor of length of stay and mortality.[16] Every transition through the model was adjusted for gestational age as clinically it is thought to impact throughout the care pathway. The estimated probabilities of receiving each level of care, or of having died or been discharged, are presented graphically, and tabulated estimates

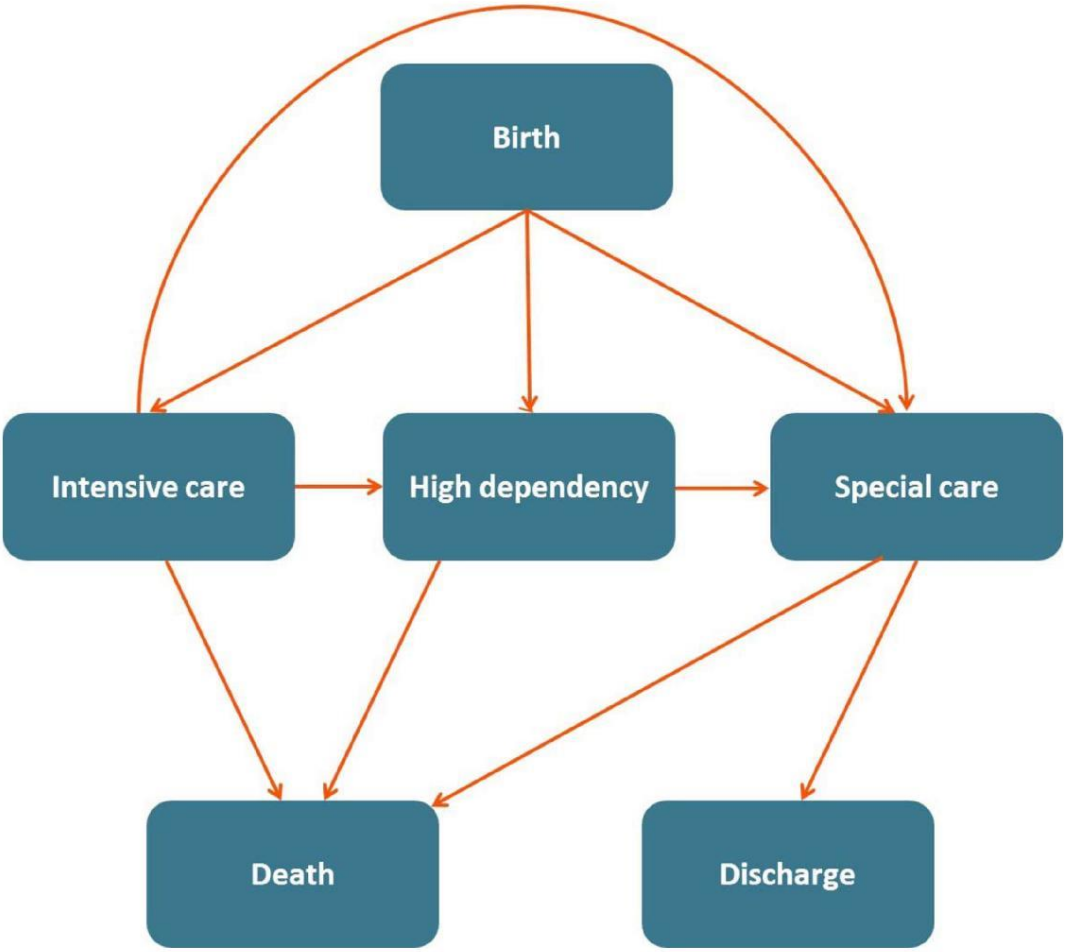


Fig 1. The multistate model used here to describe the neonatal care pathway. All babies begin in the birth state before following transitions (arrows) throughout the model until reaching death or discharge.

doi:10.1371/journal.pone.0165202.g001

are provided, for each week of gestational age over time since birth. Additionally, it is possible to calculate the expected length of stay[17] at each level of care for each week of gestational age by integrating the area under the probability curve, and these are provided.

The assumption of proportional hazards was tested using the Therneau-Grambsch test with $p<0.001$ indicating potential issues and when this occurred the Schoenfeld residuals were plotted against time to visually examine any trends.[18]

All analyses were undertaken in R 3.0.2 using the mstate command.[15,19]

Table 1. Summary statistics of babies born at 24^{+0} to 31^{+6} weeks gestational age and discharged between 2011 and 2014.

	Year of discharge/death from neonatal care			
	2011	**2012**	**2013**	**2014**
Total babies admitted, n	5,368	5,343	5,228	5,099
Gestational age, n (%)				
24	284 (5.3)	287 (5.4)	276 (5.3)	268 (5.3)
25	327 (6.1)	336 (6.3)	316 (6.0)	325 (6.4)
26	466 (8.6)	465 (8.7)	417 (8.0)	437 (8.6)
27	537 (10.0)	579 (10.8)	480 (9.2)	468 (9.2)
28	690 (12.9)	707 (13.2)	702 (13.4)	685 (13.4)
29	758 (14.1)	791 (14.9)	807 (15.4)	748 (14.7)
30	983 (18.3)	937 (17.5)	976 (18.7)	944 (18.5)
31	1,325 (24.7)	1,241 (23.2)	1,254 (24.0)	1,224 (24.0)
Sex of baby				
Male	2,953 (55.0)	2951 (55.2)	2937 (56.2)	2756 (54.1)
Female	2,411 (44.9)	2389 (44.7)	2287 (43.7)	2334 (45.8)
Indeterminate	4 (0.01)	3 (0.01)	4 (0.01)	9 (0.01)
Total days of care	305,150	306,267	295,828	298,177
Days of intensive care	60,995	63,040	60,348	62,058
Days of HDU	77,707	83,726	81,346	86,789
Days of special	166,448	159,501	154,134	149,330
Birthweight (g) Mean (SD)	1227.1 (383.2)	1215.0 (374.5)	1231.9 (378.7)	1217.2 (376.4)
Died in neonatal care, n (%)	492 (9.2)	487 (9.1)	416 (8.0)	367 (7.2)

doi:10.1371/journal.pone.0165202.t001

Results

There were 21,631 singleton babies born at 24–31 weeks gestational age, admitted to neonatal in-patient care in England from 2011 to 2014. The total length of stay was calculated for all babies; except those discharged home before 34 weeks corrected age (n = 205, 0.9%) or who stayed in hospital longer than six months (n = 199, 0.9%). Babies were also excluded if they were discharged from neonatal care having only received intensive care (n = 57, 0.3%) or were discharged having never received special care (i.e. after receiving high dependency and potentially intensive care, n = 132, 0.6%). A total of 21,038 babies remained in the analysis.

Summary characteristics of the included babies are provided in Table 1, which demonstrates that the population of babies discharged from neonatal care was broadly consistent throughout the years of the study. Each year, approximately 300,000 days of care were given to these babies, with the majority of this being special care.

A multistate Cox proportional hazards model, stratified for transition (the steps that can be taken through the care pathway), was fitted as depicted in Fig 1. The number of babies entering each state (the levels of care or death or discharge) is summarised in Table 2. Initially, a model

Table 2. Number of babies to enter and visit each state within the multistate model.

	To	**Intensive care**	**High dependency**	**Special care**	**Died**	**Discharged**
From						
Born		17,269	2,796	973	0	0
Intensive care			15,129	824	1316	0
High dependency care				17,665	260	0
Special care					186	19,276

doi:10.1371/journal.pone.0165202.t002

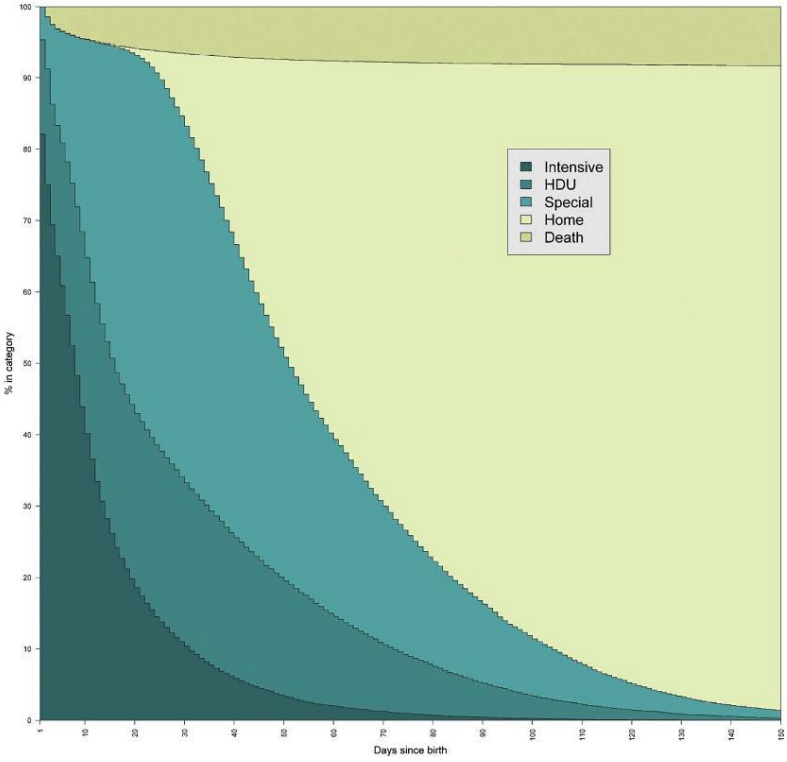


Fig 2. The percentage of babies receiving each level of care, or who have died or been discharged, over time.

doi:10.1371/journal.pone.0165202.g002

was fitted with no covariates, which is presented in Fig 2. This shows the percentage of babies in each state (y-axis), at any point in time following birth (x-axis). For example, ten days after birth 40% of babies were receiving intensive care, 25% were receiving high dependency and 30% were receiving special care. At this point, no babies had been discharged from neonatal care and 5% of babies had died. These are presented as percentages as the lack of adjustment in the model means that the results are the observed data.

Stacked plots were produced adjusted for each week of gestational age in order to show the probability of receiving any level of care, or of having died or being discharged over time (Fig 3). As week of gestation at birth increased the probability of mortality decreased and the need for higher levels of care also decreased. As expected, there was an inverse relationship between time of discharge home and gestational age. For example, for babies born at 24 weeks, at seven days of life the probability of receiving intensive care was 0.77; high dependency was 0.07; special care was 0.01 and the probability of having died was 0.15. Conversely, for babies born at 31 weeks, at seven days of life the probability of receiving intensive care was much lower at 0.20; high dependency was higher at 0.29; special care was 0.49. The probability of death for this

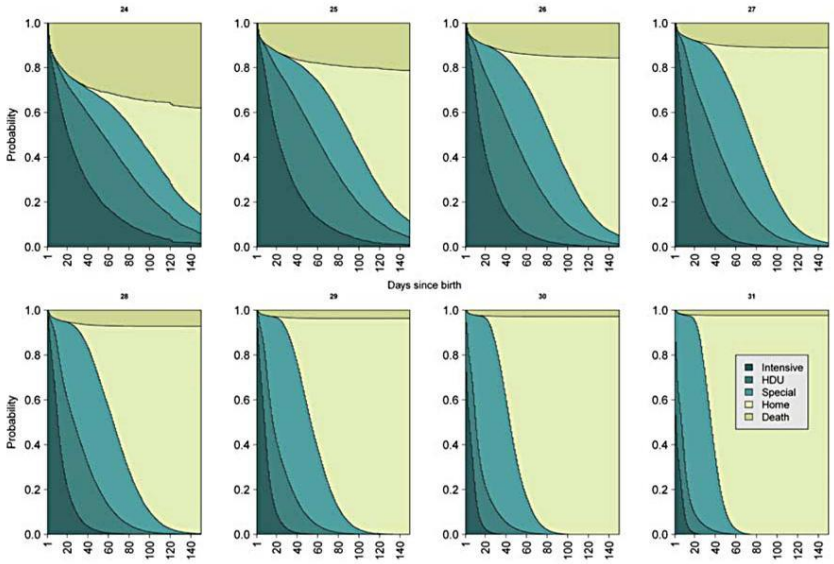


Fig 3. The probability of babies, adjusted for gestational age, receiving each level of care, or who have died or been discharged, over time.

doi:10.1371/journal.pone.0165202.g003

group was much lower at seven days: 0.02. Results related to the babies with the least data, i.e. those born at 24 and 25 weeks gestational age should be interpreted cautiously.

From Fig 3, the probabilities of receiving each level of care, or of having died or being discharged can be estimated. Tables 3 and 4 provide the probabilities at selected time points for babies born at 24 weeks and 31 weeks.

Table 3. Probabilities of 24 week babies receiving each level of care or who have died or been discharged estimated from the multistate model with an adjustment for gestational age.

Day following birth	Intensive care	High dependency	Special care	Discharge	Death
1	1.0	0.0	0.0	0.0	0.0
2	0.932	0.012	0.0	0.0	0.055
3	0.879	0.023	0.001	0.0	0.096
4	0.845	0.032	0.002	0.0	0.120
5	0.821	0.043	0.003	0.0	0.133
6	0.797	0.054	0.004	0.0	0.145
7	0.775	0.066	0.005	0.0	0.154
10	0.706	0.105	0.013	0.0	0.176
14	0.616	0.157	0.027	0.0	0.200
30	0.380	0.276	0.077	0.002	0.265
50	0.218	0.310	0.146	0.023	0.303
100	0.057	0.162	0.198	0.232	0.349
150	0.015	0.044	0.082	0.477	0.380

doi:10.1371/journal.pone.0165202.t003

Table 4. Probabilities of 31 week babies in each level of care or who have died or been discharged estimated from the multistate model with a categorical term for gestational age. Note that everyone has at least one day of care.

Day following birth	Intensive care	High dependency	Special care	Discharge	Death
1	0.533	0.324	0.144	0.0	0.0
2	0.449	0.330	0.216	0.0	0.006
3	0.383	0.300	0.306	0.001	0.011
4	0.334	0.296	0.356	0.001	0.014
5	0.288	0.298	0.398	0.002	0.015
6	0.244	0.296	0.443	0.002	0.016
7	0.200	0.290	0.493	0.002	0.017
10	0.089	0.236	0.654	0.003	0.019
14	0.027	0.151	0.797	0.005	0.021
30	0.0	0.038	0.671	0.269	0.023
50	0.0	0.005	0.094	0.878	0.024
100	0.0	0.0	0.0	0.977	0.024
150	0.0	0.0	0.0	0.977	0.024

doi:10.1371/journal.pone.0165202.t004

The expected time spent in each state can be calculated for all babies, irrespective of their outcome, born at each week of gestational age and these are provided in Table 5.

In addition to transition probabilities, it is also possible to estimate hazard ratios for each week of gestation as gestational age was modelled categorically (Table 6). Each hazard ratio can be interpreted as the hazard of experiencing that transition through the neonatal care pathway, for a given week of gestational age at birth, compared to the babies who make that transition who were born at 27 weeks. For example, the probability of death at any time point, i.e. the hazard, after receiving only intensive care is three times higher in the 24 weeks group compared to the 27 weeks (Hazard Ratio: 3.03, 95% CI: 2.51 to 3.65, $p<0.001$). Conversely, the hazard ratio of experiencing the same transition in the 31 weeks group compared to 27 weeks was 0.64 (95% CI: 0.49 to 0.84, $p = 0.001$), suggested a reduced hazard of death after receiving only intensive care in the 31 week babies compared to the 27 week babies.

Certain transitions should be interpreted with caution, as lack of data means hazard ratios are potentially poorly estimated. For example, very few babies receive intensive care and then special care, with no high dependency care, and this is seen in the hazard ratio of this transition for babies born at 31 weeks which is estimated as being 53.6 (95% CI: 31.4 to 91.6) with a wide confidence interval.

Two transitions were identified as potentially breaching the proportional hazards assumption: intensive care to high dependency and high dependency to special care. This is

Table 5. Expected time spent receiving each level of care, and total length of stay, by gestational age. Results are rounded up to the nearest day.

Gestational age	Intensive care (days)	High dependency (days)	Special care (days)	Total (days)
24	33	29	22	84
25	30	33	27	90
26	23	29	30	82
27	18	25	31	74
28	14	19	33	66
29	10	13	33	56
30	6	9	31	46
31	4	6	28	38

doi:10.1371/journal.pone.0165202.t005

Table 6. Hazard ratios of each transition by gestational age. The hazard ratio compares back to the baseline of 27 weeks gestational age for that transition.

Gestational age Transition	Hazard ratio	95% Confidence Interval	p-value
24 weeks			
IC -> HD	0.38	0.35 to 0.41	<0.01
IC -> SC	0.65	0.26 to 1.62	0.36
IC -> Death	3.03	2.51 to 3.65	<0.01
HD -> SC	0.51	0.47 to 0.56	<0.01
HD -> Death	1.85	1.21 to 2.82	0.01
SC -> Home	0.58	0.53 to 0.64	<0.01
SC -> Death	4.51	2.15 to 9.45	<0.01
25 weeks			
IC -> HD	0.52	0.48 to 0.56	<0.01
IC -> SC	0.25	0.07 to 0.88	0.03
IC -> Death	1.63	1.33 to 1.99	<0.01
HD -> SC	0.61	0.57 to 0.66	<0.01
HD -> Death	1.05	0.68 to 1.61	0.84
SC -> Home	0.69	0.64 to 0.75	<0.01
SC -> Death	3.40	1.73 to 6.71	<0.01
26 weeks			
IC -> HD	0.75	0.70 to 0.80	<0.01
IC -> SC	0.52	0.23 to 1.32	0.18
IC -> Death	1.25	1.02 to 1.53	0.03
HD -> SC	0.80	0.75 to 0.86	<0.01
HD -> Death	0.88	0.58 to 1.34	0.55
SC -> Home	0.80	0.75 to 0.86	<0.01
SC -> Death	2.96	1.68 to 5.21	<0.01
27 weeks			
IC -> HD	Reference	Reference	Reference
IC -> SC	Reference	Reference	Reference
IC -> Death	Reference	Reference	Reference
HD -> SC	Reference	Reference	Reference
HD -> Death	Reference	Reference	Reference
SC -> Home	Reference	Reference	Reference
SC -> Death	Reference	Reference	Reference
28 weeks			
IC -> HD	1.38	1.30 to 1.47	<0.01
IC -> SC	3.95	2.21 to 7.07	<0.01
IC -> Death	0.71	0.57 to 0.87	<0.01
HD -> SC	1.40	1.32 to 1.48	<0.01
HD -> Death	0.79	0.51 to 1.22	0.29
SC -> Home	1.25	1.18 to 1.33	<0.01
SC -> Death	0.59	0.31 to 1.04	0.07
29 weeks			
IC -> HD	2.04	1.92 to 2.17	<0.01
IC -> SC	9.93	5.71 to 17.3	<0.01
IC -> Death	0.44	0.34 to 0.57	<0.01
HD -> SC	1.97	1.85 to 2.09	<0.01
HD -> Death	0.49	0.28 to 0.86	0.01

(*Continued*)

Table 6. (*Continued*)

Gestational age Transition	Hazard ratio	95% Confidence Interval	p-value
SC -> Home	1.87	1.76 to 1.98	<0.01
SC -> Death	0.27	0.15 to 0.52	<0.01
30 weeks			
IC -> HD	2.86	2.69 to 3.04	<0.01
IC -> SC	26.7	15.6 to 45.9	<0.01
IC -> Death	0.55	0.42 to 0.72	<0.01
HD -> SC	2.44	2.30 to 2.60	<0.01
HD -> Death	0.26	0.12 to 0.55	<0.01
SC -> Home	3.17	2.99 to 3.36	<0.01
SC -> Death	0.16	0.08 to 0.30	<0.01
31 weeks			
IC -> HD	3.48	3.27 to 3.70	<0.01
IC -> SC	53.6	31.4 to 91.6	<0.01
IC -> Death	0.64	0.49 to 0.84	<0.01
HD -> SC	3.10	2.91 to 3.30	<0.01
HD -> Death	0.48	0.25 to 0.95	0.03
SC -> Home	6.26	5.90 to 6.63	<0.01
SC -> Death	0.10	0.05 to 0.20	<0.01

In this table the following acronyms are used: intensive care (IC); high dependency care (HD) and special care (SC)

doi:10.1371/journal.pone.0165202.t006

unsurprising as these transitions had the most data, and potentially needed to be modelled in a more complex manner, such as via using fractional polynomials[20] or time varying covariates. However, examination of the Schoenfeld residuals against a function of time indicated that the issues with these transitions were likely to be minimal, but future work will investigate this further. Further issues of proportional hazards may also exist in other transitions, most notably the transition into discharge.

Discussion

In this paper an alternative statistical approach has been proposed and illustrated in order to investigate length of neonatal in-patient stay, which allows the incorporation of babies that died in neonatal care, whilst also considering the levels of care that all babies require. Previous work investigating length of stay has either removed babies that die[2–4] or included them inappropriately, potentially biasing the results.[5] The modelling of survival and length of stay is becoming increasingly important in neonatal care, particularly as survival at early gestational ages improves and the care requirements of these babies increases.[21] It might seem appropriate to focus attention on survivors, as the resource use and workload they require is great. However, this does not portray an accurate picture of the total workload for neonatal care as for the most preterm babies mortality rates are high.[4,21]

Although previous work[7] has considered babies that die during their time in neonatal care, babies born very preterm can spend a long time in hospital and receive complex types of care.[4,7] Estimating total length of stay alone does not provide a clear picture of the experiences they have during their neonatal stay. Detailed information about neonatal care is important for counselling families and for commissioning services.

The probabilities of receiving each level of care, or of having died or being discharged presented here can be used in the counselling of parents and informing the conversations had with

clinicians. For example, nearly all 31 week babies were discharged or had died by 50 days after birth. Conversely, for babies born at 24 weeks, there was a small probability of remaining in hospital even 150 days after birth. Future work can provide detailed estimates for clinicians to use in the counselling of parents.

For a healthcare service increasingly focussed on costs, estimating detailed information about length of stay aids the informing of neonatal service provision, commissioning and funding. Here we estimated the expected time receiving each level of care for all babies, irrespective of outcome. These estimates could be used to aid the commissioning of resources by calculating the expected number of babies at each gestation and multiplying by the expected days of care to receive an estimate of the resource need.

The results presented here arise from use of multistate modelling[14,22] which is a relatively novel approach that has been used in other medical areas (e.g. cancer[23]). User-friendly software has now made implementation of this method more straightforward.[15,19] Multistate modelling allows consideration of multiple 'competing' outcomes (where the occurrence of one outcome prevents the occurrence of others), whilst also modelling intermediate steps (here levels of care). This allows for modelling of different levels of care, as well as death and discharge over time. Here the results are presented graphically, for ease of interpretation and in the future these plots or similar will provide a useful aide for clinicians giving information about the length of time spent at each level of care, whilst also considering the probability of surviving to discharge.

The data used here cover all admissions and discharges from English neonatal units, avoiding the bias that would be introduced by examining an individual hospital, and thus be subject to local policy and practice.[13] Whilst population differences will exist between individual units, our aim was to provide a population-based estimate, and future work could examine individual hospitals or Neonatal Networks to investigate variations within England.

Strengths and limitations

A strength of this work is that this is the first time that the entire neonatal care pathway has been described, using national data, whilst simultaneously considering the different levels of care and competing outcomes of death or discharge from neonatal care. However, many people, criticise the Cox model, which was used here, as having been overused when other methods would be more appropriate.[24] However, we did not want to impose distributional assumptions on the shape of the hazard function by using a parametric approach as examination of the observed hazards indicated this might be inappropriate.

There are issues of proportional hazards in at least two of the transitions but it is well known that the introduction of time dependent effects in Cox modelling is difficult and computationally intensive particularly with large datasets such as this one.[25] Future work will need to investigate alternative methods, including a flexible parametric approach[26] to allow further flexibility in all transitions. Additionally, the lack of data in the group of babies born at 24 and 25 weeks gestational age means the model here does not fit as well to these groups. This is particularly apparent when considering the time at which the probability of discharge occurs in the extremely preterm babies, and these results should be interpreted cautiously and investigated further.

Gestational age was modelled categorically and some would argue this may lead to a loss of statistical power.[27] However, there is an underlying clinical meaning to completed weeks of gestation which is often used in practice and allows for simple interpretation of the hazard ratios.[28] Other covariates including birthweight are likely to be important adjustments to make in this work[29] and future research will consider their inclusion, as here the aim was to

provide an introduction to this type of analysis. Similarly, this work has considered singletons, and multiples are known to comprise a large proportion of the preterm population. Future work will include multiples and consider appropriate stratification or adjustment of the methods presented here. Adjustments for covariates in a multistate analysis should be made with care to avoid overfitting, particularly where the number of events for a particular transition is small. Gestational age was adjusted for here on every transition, but adjustments can be made which are assumed to be the same for all transitions.

Only deaths which occurred in neonatal care are considered here, and deaths may have occurred elsewhere. However, it was not possible to retrieve this information. Future work will link the neonatal care with care received elsewhere. However, we believe the number of deaths in other locations is likely to be a small number.

Finally, care here was assumed to have occurred in a hierarchical manner, with all intensive care preceding high dependency, and all high dependency care preceding special care. Whilst this is unlikely to be the reality for some babies, particularly those that become more unwell during their time in hospital (e.g. contracting sepsis), this work still provides a useful resource for clinicians and the commissioning of services. Future work will investigate the order in which babies receive their differing levels of care.

Conclusion

Modelling of survival and length of stay has become increasingly important in neonatal care to provide parents with accurate and realistic information about their baby's care. It is also important for the commissioning of services and resources. Multistate modelling is a statistical approach which begins to describe the neonatal care pathway in a more complete way than has been previously undertaken. These methods allow for modelling of mortality and survival, whilst also considering all the potential levels of care which can occur along the care pathway. Estimates of the probability of being in hospital, receiving a given level of care, or of having died or been discharged can be produced and used to inform conversations between clinicians and parents. Future research should refine and improve these estimates particularly for the extremely preterm babies.

Acknowledgments

We would like to thank the reviewers for their comments which have improved this paper.

The authors would like to thank all the neonatal units that allowed their data to be used in this work. We would like to thank the Lead Clinicians of the UK Neonatal Collaborative: Dr Matthew Babirecki, Dr Liza Harry, Dr Oliver Rackham, Dr Tim Wickham, Dr Sanaa Hamdan, Dr Aashish Gupta, Dr Ruth Wigfield, Dr L M Wong, Dr Anita Mittal, Dr Julie Nycyk, Dr Phil Simmons, Dr Vishna Rasiah, Dr Sunita Seal, Dr Ahmed Hassan, Dr Karin Schwarz, Dr Mark Thomas, Dr Ainyne Foo, Dr Aravind Shastri, Dr Graham Whincup, Dr Stephen Brearey, Dr John Chang, Dr Khairy Gad, Dr Abdul Hasib, Dr Mehdi Garbash, Dr Nicci Maxwell, Dr David Gibson, Dr Pauline Adiotomre, Dr Jamal S Ahmed, Dr Abby Deketelaere, Dr Ramnik Mathur, Dr K Abdul Khader, Dr Ruth Shephard, Dr Abdus Mallik, Dr Belal Abuzgia, Dr Mukta Jain, Dr Simon Pirie, Dr Phil Simmons, Dr Stanley Zengeya, Dr Timothy Watts, Dr C Jampala, Dr Cath Seagrave, Dr Michele Cruwys, Dr Hilary Dixon, Dr Narendra Aladangady, Dr Hassan Gaili, Dr Matthew James, Dr M Lal, Dr Ambadkar, Dr Patti Rao, Dr Khalid Mannan, Dr Ann Hickey, Dr Dhaval Dave, Dr Nader Elgharably, Dr Meera Lama, Dr Lawrence Miall, Dr Jonathan Cusack, Dr Venkatesh Kairamkonda, Dr Jayachandran, Dr Kollipara, Dr J Kefas, Dr Bill Yoxall, Dr Jennifer Birch, Dr Gail Whitehead, Dr Bashir Jan Muhammad, Dr Aung Soe, Dr I Misra, Dr Tilly Pillay, Dr Imdad Ali, Dr Mark Dyke, Dr Michael Selter, Dr Nagesh Panasa, Dr

Lesley Alsford, Dr Alan Fenton, Dr Subodh Gupta, Dr Richard Nicholl, Dr Steven Wardle, Dr Tim McBride, Dr Naveen Shettihalli, Dr Eleri Adams, Dr Seif Babiker, Dr Margaret Crawford, Dr David Gibson, Dr Minesh Khashu, Dr Caitlin Toh, Dr M Hall, Dr P Amess, Dr Elizabeth Sleight, Dr Charlotte Groves, Dr Sunit Godambe, Dr Dennis Bosman, Dr Barbara Piel, Dr Banjoko, Dr N Kumar, Dr A Manzoor, Dr Wilson Lopez, Dr Angela D'Amore, Dr Shameel Mattara, Dr Christos Zipitis, Dr Peter De Halpert, Dr Paul Settle, Dr Paul Munyard, Dr Gitika Joshi, Dr David Bartle, Dr D Schapira, Dr Joanne Fedee, Dr Natasha Maddock, Dr Richa Gupta, Dr Deshpande, Dr Charles Godden, Dr P Amess, Dr Stephen Jones, Dr Alan Fenton, Dr Mahadevan, Dr Nick Brown, Dr Kirsten Mack, Dr Pauline Adiotomre, Dr Rob Bolton, Dr A Khan, Dr Paul Mannix, Dr Charlotte Huddy, Dr Salim Yasin, Dr Sian Butterworth, Dr Sunit Godambe, Dr Ngozi Edi-Osagie, Dr Bala Thyagarajan, Dr Peter Reynolds, Dr Nick Brennan, Dr Carrie Heal, Dr Sanjay Salgia, Dr Majd Abu-Harb, Dr Jacqeline Birch, Dr Chris Knight, Dr Simon Clark, Dr V Van Sommen, Dr Nandiran Ratnavel, Dr Mala Raman, Dr Hamudi Kisat, Dr Sara Watkin, Dr Kate Blake, Dr Jauro Kuna, Dr Mehdi Garbash, Dr Alison Moore, Dr Hari Kumar, Dr Gopi Vemuri, Dr Chris Rawlingson, Dr Delyth Webb, Dr Bird, Dr Sankara Narayanan, Dr Jason Gane, Dr Elizabeth Eyre, Dr Ian Evans, Dr Rekha Sanghavi, Dr Caroline Sullivan, Dr Laweh Amegavie, Dr Wynne Leith, Dr Vimal Vasu, Dr Andrew Gallagher, Dr Katia Vamvakiti, Dr Megan Eaton, Dr Guy Millman

Author Contributions

Conceptualization: SS BM ED.

Data curation: SS NM.

Formal analysis: SS BM KA.

Funding acquisition: SS.

Investigation: SS LB ED KA NM BM.

Methodology: SS KA BM.

Project administration: SS.

Resources: NM ED BM KA.

Software: SS.

Supervision: ED KA BM NM.

Validation: SS.

Visualization: SS BM KA.

Writing – original draft: SS.

Writing – review & editing: SS LB ED KA NM BM.

References

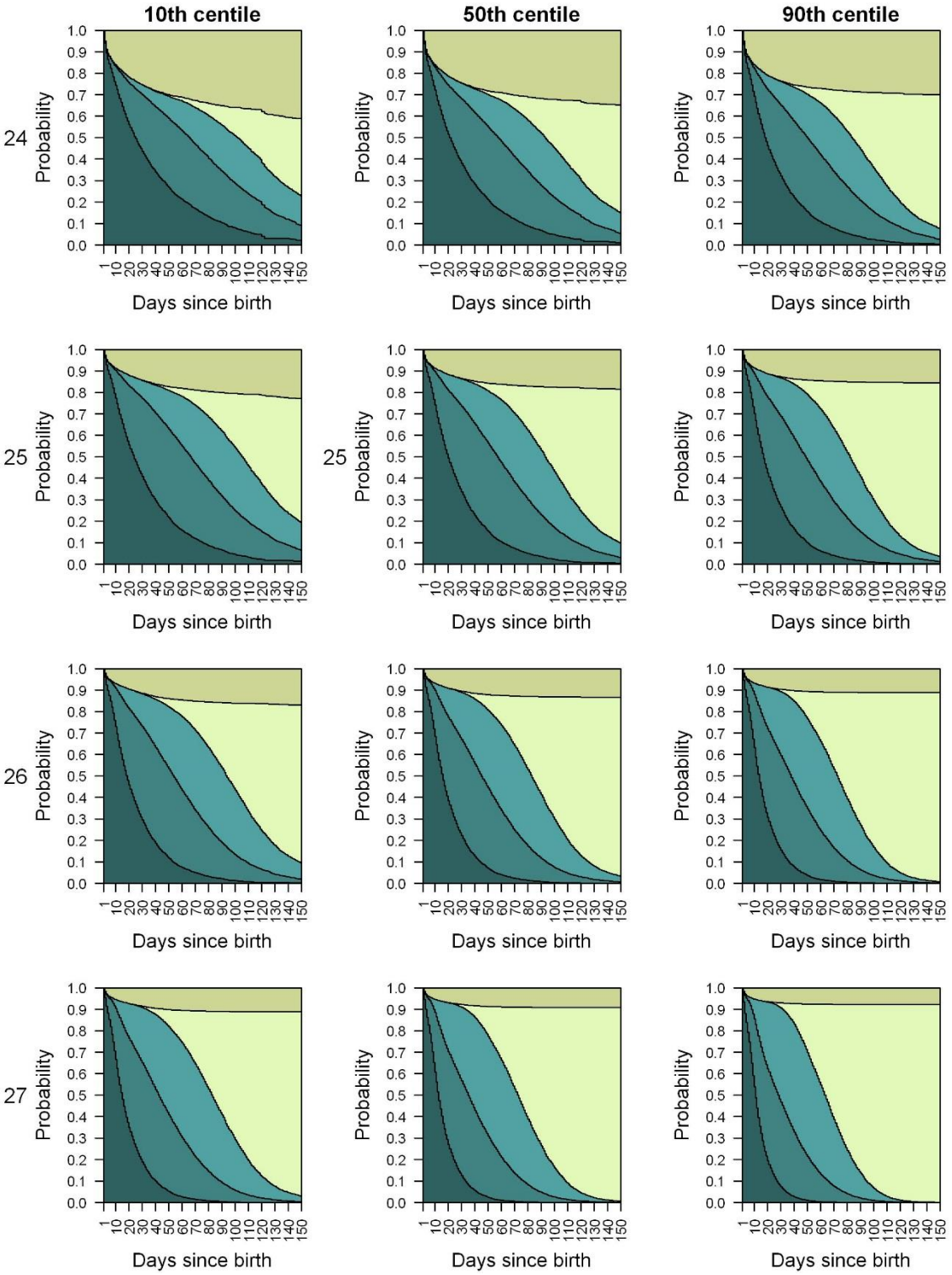
1. NNAP (2015) National Neonatal Audit Programme Annual Report 2015. London.
2. Altman M, Vanpee M, Cnattingius S, Norman M (2009) Moderately preterm infants and determinants of length of hospital stay. Archives of Disease in Childhood Fetal & Neonatal Edition 94: F414–418.
3. Lee HC, Bennett MV, Schulman J, Gould JB (2013) Accounting for variation in length of NICU stay for extremely low birth weight infants. J Perinatol 33: 872–876. doi: 10.1038/jp.2013.92 PMID: 23949836

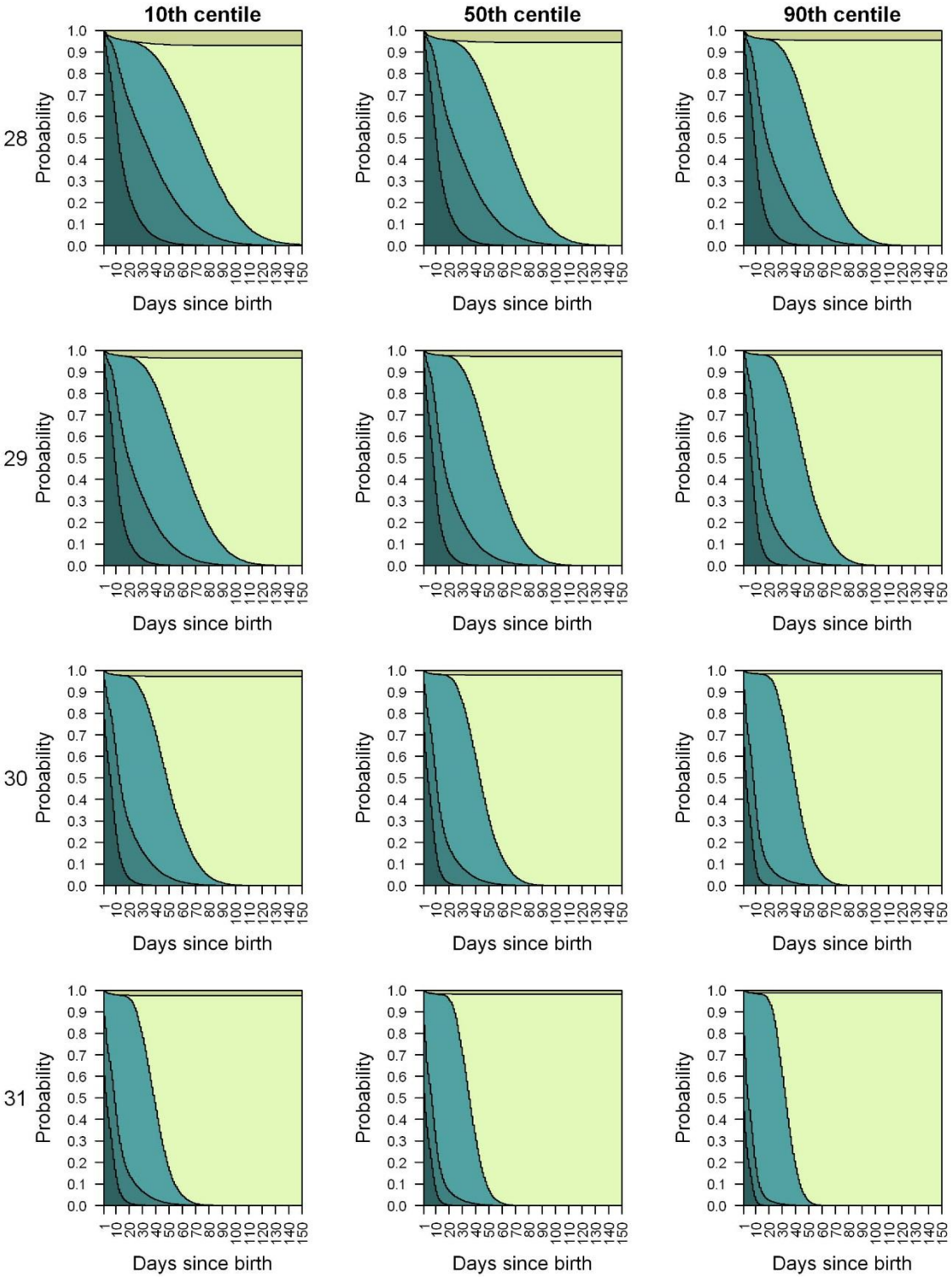
4. Manktelow B, Draper ES, Field C, Field D (2010) Estimates of length of neonatal stay for very premature babies in the UK. Archives of Disease in Childhood Fetal & Neonatal Edition 95: F288–292.
5. Bender GJ, Koestler D, Ombao H, McCourt M, Alskinis B, Rubin LP, et al. (2013) Neonatal intensive care unit: predictive models for length of stay. Journal of Perinatology 33: 147–153. doi: 10.1038/jp.2012.62 PMID: 22678140
6. Manktelow BN, Seaton SE, Field DJ, Draper ES (2013) Population-based estimates of in-unit survival for very preterm infants. Pediatrics 131: e425–432. doi: 10.1542/peds.2012-2189 PMID: 23319523
7. Hinchliffe SR, Seaton SE, Lambert PC, Draper ES, Field DJ, Manktelow BN (2013) Modelling time to death or discharge in neonatal care: an application of competing risks. Paediatric and Perinatal Epidemiology 27: 426–433. doi: 10.1111/ppe.12053 PMID: 23772944
8. British Association of Perinatal Medicine (2011) BAPM Categories of Care. Available online at: http://www.bapm.org/publications/.
9. Stark AR (2004) Levels of neonatal care. Pediatrics 114: 1341–1347. doi: 10.1542/peds.2004-1697 PMID: 15520119
10. Van Reempts P, Gortner L, Milligan D, Cuttini M, Petrou S, Agostino R, et al. (2007) Characteristics of neonatal units that care for very preterm infants in Europe: results from the MOSAIC study. Pediatrics 120: e815–825. doi: 10.1542/peds.2006-3122 PMID: 17908739
11. Hallsworth M, Farrands A, Oortwijn W, E H (2008) The provision of neonatal services: Data for international comparisons. Santa Monica, CA.
12. Davies A, On Behalf of Neonatal Cinical Reference Group (2014) Transitional care report, Published online at: http://www.wmscnsenate.nhs.uk/index.php/download_file/view/205/971/ [Last accessed 21/12/2015].
13. Smith L, Draper ES, Manktelow BN, Pritchard C, Field DJ (2013) Comparing regional infant death rates: the influence of preterm births <24 weeks of gestation. Arch Dis Child Fetal Neonatal Ed 98: F103–107. doi: 10.1136/fetalneonatal-2011-301359 PMID: 22684158
14. Putter H, Fiocco M, Geskus RB (2007) Tutorial in biostatistics: competing risks and multi-state models. Statistics in Medicine 26: 2389–2430. doi: 10.1002/sim.2712 PMID: 17031868
15. de Wreede LC, Fiocco M, Putter H (2010) The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. Computer Methods and Programs in Biomedicine 99: 261–274. doi: 10.1016/j.cmpb.2010.01.001 PMID: 20227129
16. Medlock S, Ravelli AC, Tamminga P, Mol BW, Abu-Hanna A (2011) Prediction of mortality in very premature infants: a systematic review of prediction models. PLoS One 6: e23441. doi: 10.1371/journal.pone.0023441 PMID: 21931598
17. Beyermann J, Putter H (2014) A note on computing average state occupation times. Demographic Research 30: 1681–1696.
18. Therneau T, Grambsch P (2000) Modelling Survival Data: Extending the Cox Model. Springer: New York.
19. de Wreede LC, Fiocco M, Putter H (2011) mstate: An R Package for the Analysis of Competing Risks and Multi-State Mod. Journal of Statistical Software 38.
20. Royston P, Ambler G, Sauerbrei W (1999) The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 28: 964–974. PMID: 10597998
21. Field DJ, Dorling JS, Manktelow BN, Draper ES (2008) Survival of extremely premature babies in a geographically defined population: prospective cohort study of 1994–9 compared with 2000–5. BMJ 336: 1221–1223. doi: 10.1136/bmj.39555.670718.BE PMID: 18469017
22. Andersen PK, Keiding N (2002) Multi-state models for event history analysis. Stat Methods Med Res 11: 91–115. PMID: 12040698
23. Putter H, van der Hage J, de Bock GH, Elgalta R, van de Velde CJ (2006) Estimation and prediction in a multi-state model for breast cancer. Biometrical Journal 48: 366–380. PMID: 16845902
24. Reid N (1994) A Conversation with Sir David Cox. Statistical Science: 439–455.
25. Sauerbrei W, Royston P, Look M (2007) A new proposal for multivariable modelling of time-varying effects in survival data based on fractional polynomial time-transformation. Biom J 49: 453–473. doi: 10.1002/bimj.200610328 PMID: 17623349
26. Lambert PC, Royston P (2009) Further development of flexible parametric models for survival analysis. Stata Journal 9: 265–290.
27. Altman DG, Royston P (2006) The cost of dichotomising continuous variables. BMJ 332: 1080. doi: 10.1136/bmj.332.7549.1080 PMID: 16675816

28. Baneshi MR, Nakhaee F, Law M (2013) On the Use of Fractional Polynomial Models to Assess Preventive Aspect of Variables: An Example in Prevention of Mortality Following HIV Infection. Int J Prev Med 4: 414–419. PMID: 23671772

29. Seaton SE, Barker L, Jenkins D, Draper ES, Abrams KR, Manktelow BN (2016) What factors predict length of stay in a neonatal unit, a systematic review (In Press). BMJ Open.

ADJUSTED STACKED PROBABILITY PLOTS FOR FEMALES




10th centile
50th centile
90th centile
28
29
30
31
Probability
Days since birth

THERNEAU-GRAMBSCH TEST FOR ADJUSTED MODEL[14]

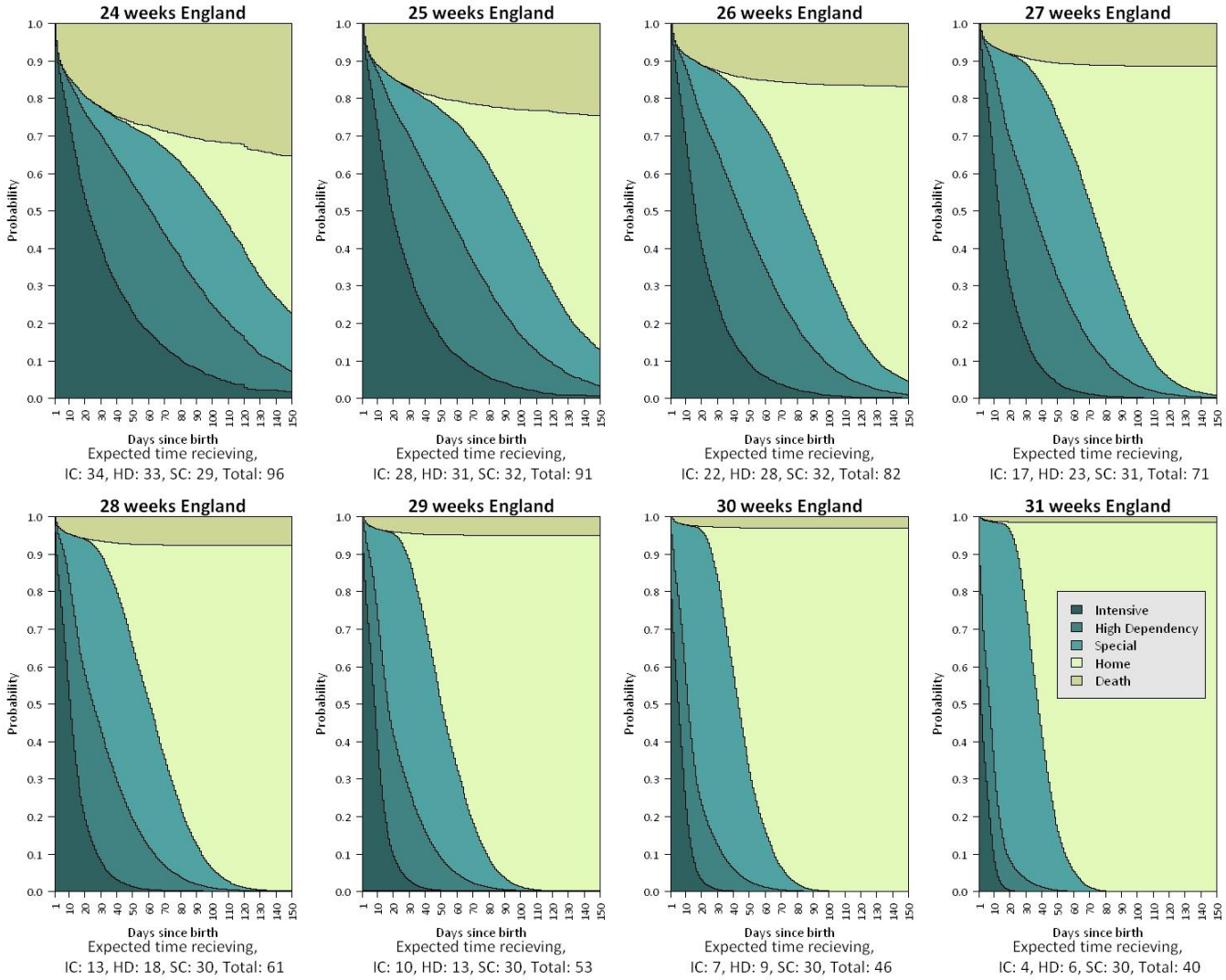
Gestational age (weeks) Transition	Chi-squared value	p-value
24 weeks		
IC -> HD	56.3	<0.001
IC -> SC	0.31	0.576
IC -> Death	4.47	0.035
HD -> SC	24.9	<0.001
HD -> Death	6.26	0.012
SC -> Home	1.32	0.251
SC -> Death	26.2	<0.001
25 weeks		
IC -> HD	25.8	<0.001
IC -> SC	1.59	0.207
IC -> Death	0.17	0.684
HD -> SC	24.9	<0.001
HD -> Death	2.99	0.084
SC -> Home	1.11	0.293
SC -> Death	11.6	<0.001
26 weeks		
IC -> HD	6.63	0.010
IC -> SC	0.20	0.657
IC -> Death	1.50	0.221
HD -> SC	14.3	<0.001
HD -> Death	1.02	0.314
SC -> Home	0.59	0.442
SC -> Death	4.44	0.035
27 weeks		
IC -> HD	Baseline	Baseline
IC -> SC	Baseline	Baseline
IC -> Death	Baseline	Baseline
HD -> SC	Baseline	Baseline
HD -> Death	Baseline	Baseline
SC -> Home	Baseline	Baseline
SC -> Death	Baseline	Baseline
28 weeks		
IC -> HD	11.0	<0.001
IC -> SC	3.74	0.053
IC -> Death	0.67	0.412
HD -> SC	14.1	<0.001
HD -> Death	2.00	0.158
SC -> Home	4.29	0.038
SC -> Death	1.29	0.256

[14] In this table the following acronyms apply: IC intensive care; HD high dependency; SC special care.

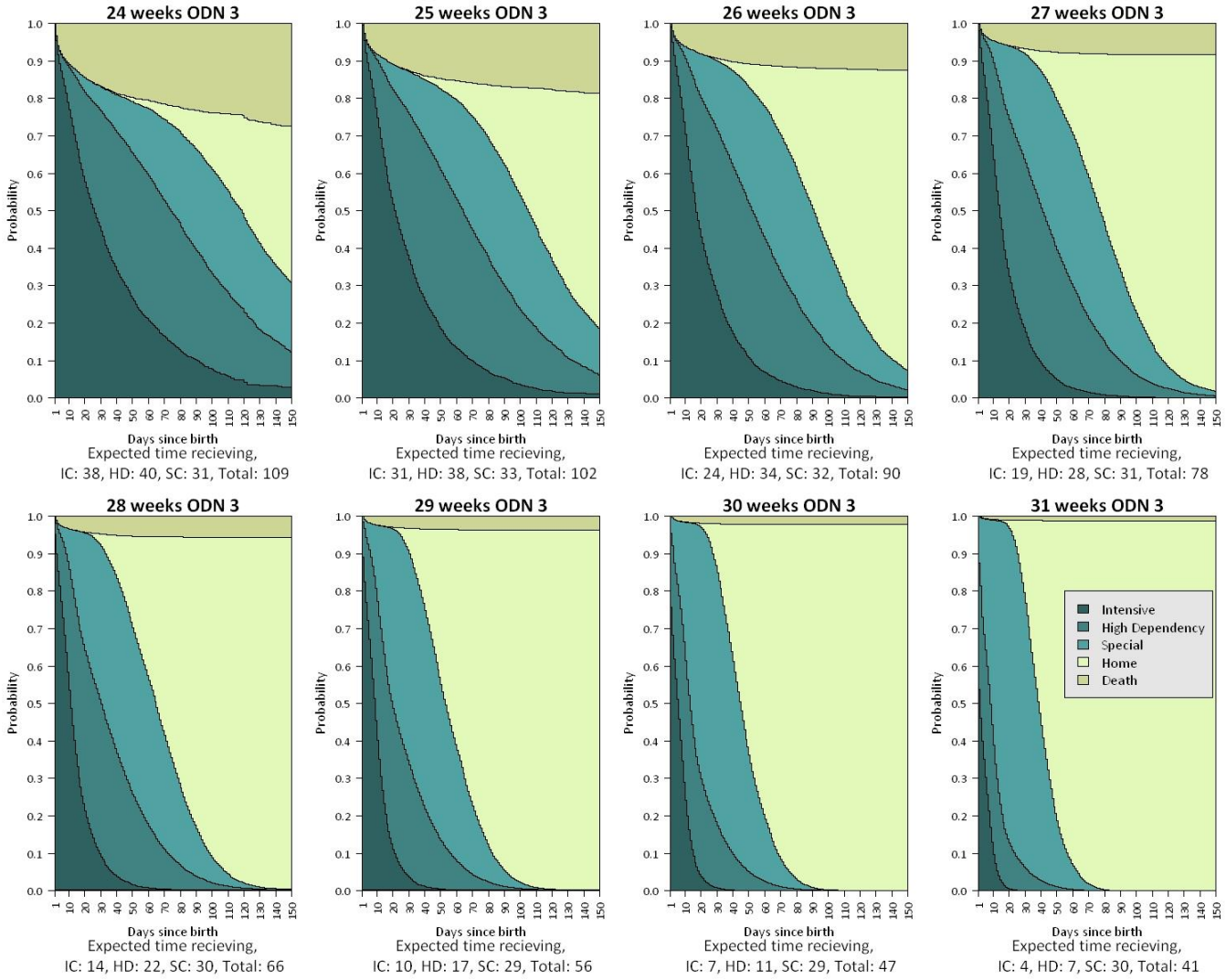
Gestational age (weeks) **Transition**	**Chi-squared value**	**p-value**
29 weeks		
IC -> HD	27.5	<0.001
IC -> SC	4.92	0.026
IC -> Death	1.18	0.278
HD -> SC	31.7	<0.001
HD -> Death	0.85	0.357
SC -> Home	22.0	<0.001
SC -> Death	2.09	0.149
30 weeks		
IC -> HD	69.0	<0.001
IC -> SC	2.89	0.015
IC -> Death	0.06	0.804
HD -> SC	57.6	<0.001
HD -> Death	0.76	0.385
SC -> Home	83.3	<0.001
SC -> Death	2.62	0.106
31 weeks		
IC -> HD	113.0	<0.001
IC -> SC	7.87	0.005
IC -> Death	0.37	0.545
HD -> SC	70.1	<0.001
HD -> Death	6.32	0.012
SC -> Home	186.0	<0.001
SC -> Death	4.70	0.03
Sex		
IC -> HD	0.09	0.761
IC -> SC	0.21	0.646
IC -> Death	0.00	0.992
HD -> SC	0.09	0.763
HD -> Death	0.61	0.434
SC -> Home	0.49	0.503
SC -> Death	0.044	0.833
Birthweight z-score		
IC -> HD	3.24	0.072
IC -> SC	0.17	0.680
IC -> Death	3.35	0.067
HD -> SC	11.5	<0.001
HD -> Death	0.42	0.516
SC -> Home	69.4	<0.001
SC -> Death	1.66	0.197

APPENDIX 7

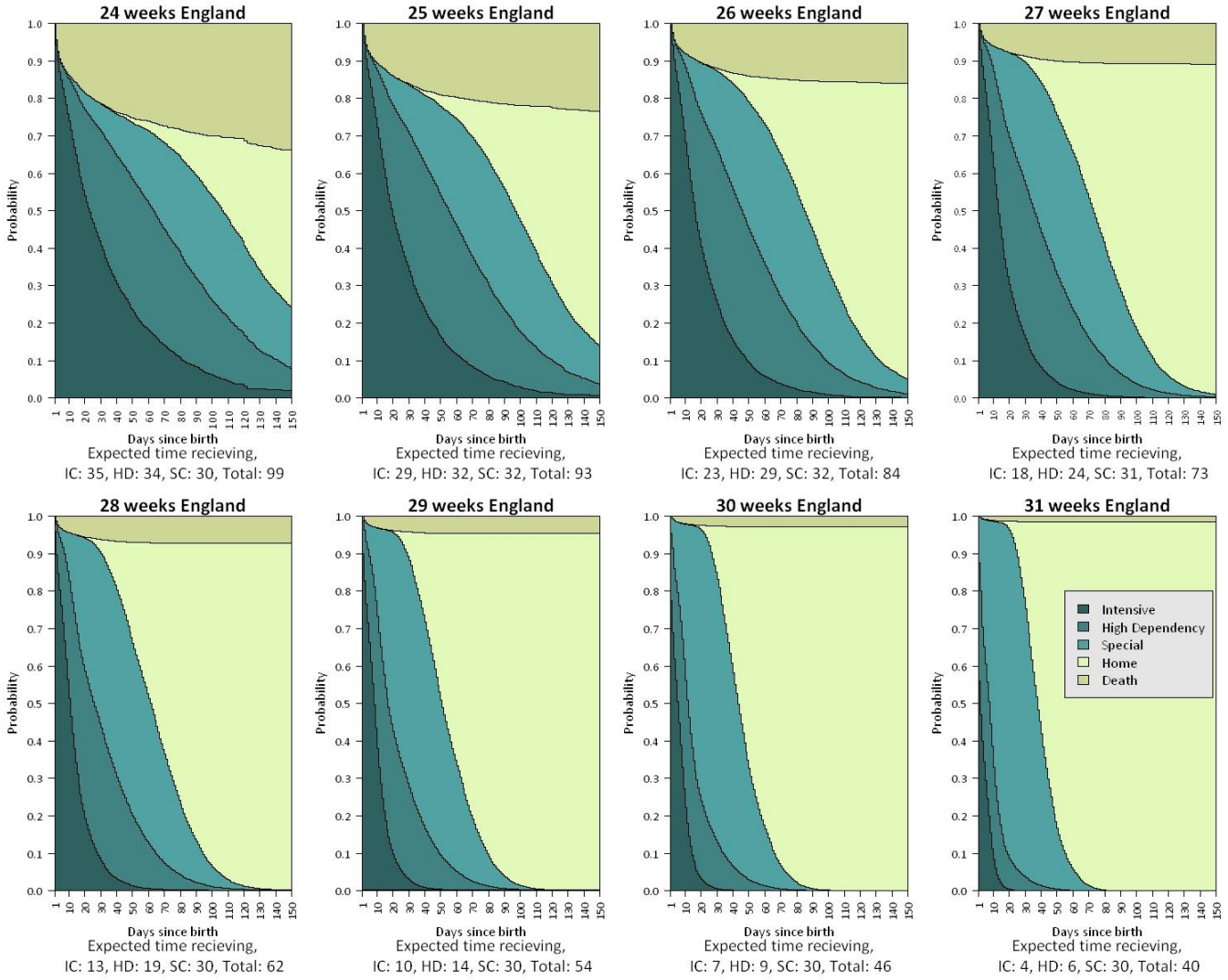
STACKED PROBABILITY PLOT FOR ENGLAND EXCLUDING ODN 3



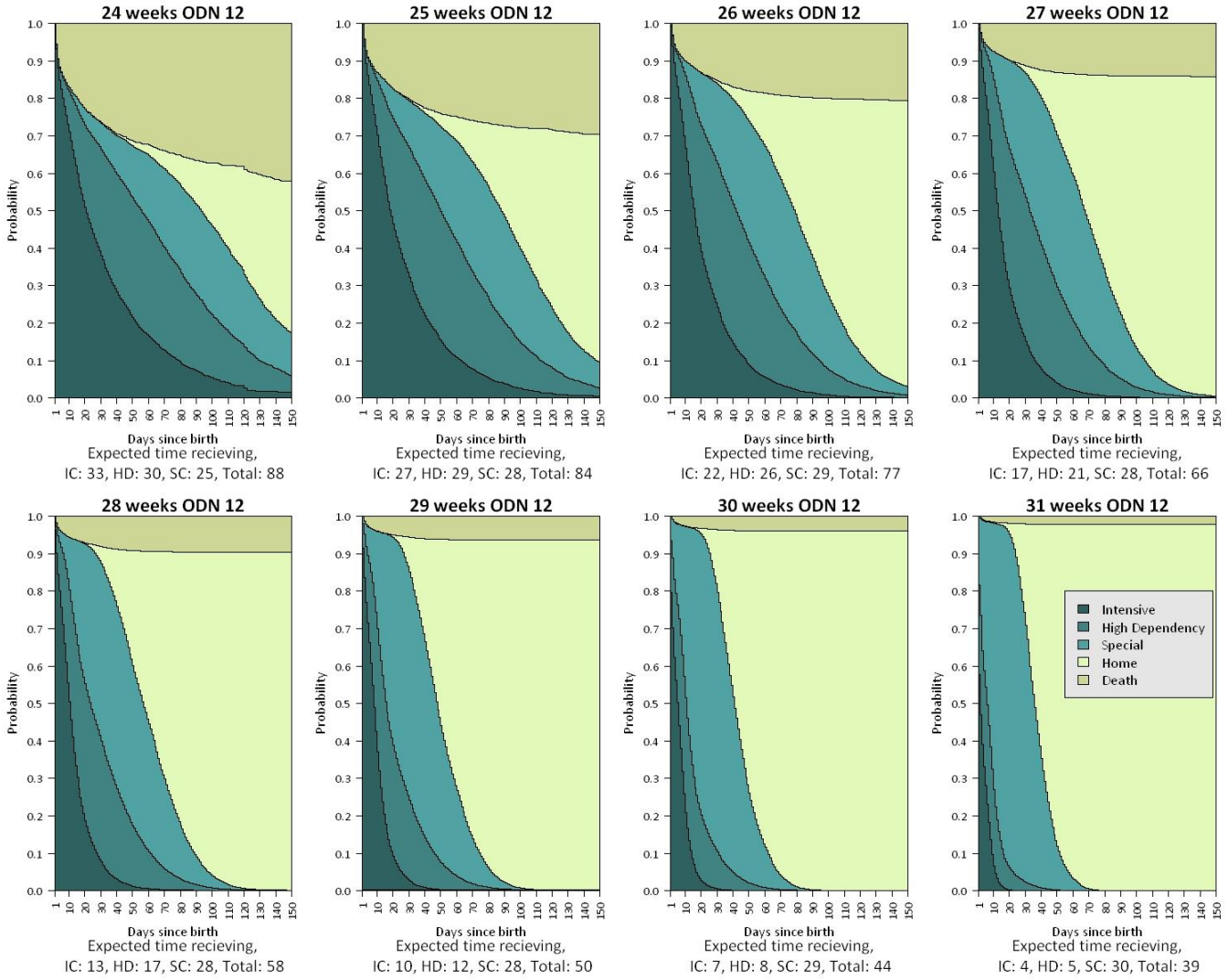
STACKED PROBABILITY PLOT FOR ODN 3



STACKED PROBABILITY PLOT FOR ENGLAND EXCLUDING ODN 12



STACKED PROBABILITY PLOTS FOR ODN 12



APPENDIX 8

DISSEMINATION AND IMPACT OF THIS THESIS

PEER REVIEWED PAPERS

Two academic outputs in peer-reviewed journals and one magazine article have produced from this thesis:

- **Seaton SE**, Barker L, Jenkins D, Draper ES, Abrams KR, Manktelow BN. What factors predict length of stay in a neonatal unit: a systematic review. BMJ Open 2016; 6:e010466.
- **Seaton SE**; Barker L. Modelling neonatal care pathways: investigating length of stay for preterm infants. Infant 2016, 12(3): 87-90.
- **Seaton SE**, Barker L, Draper ES, Abrams KR, Modi N, Manktelow BN. Modelling neonatal care pathways for babies born preterm: an application of multistate modelling. PLOS ONE 2016; 11(10):e0165202

PRESENTATION AT CONFERENCES

This work has been presented at the following national and international conferences:

- Prediction models for neonatal length of stay: a systematic review. Presented at the European Congress of Epidemiology, Maastricht (June 2015) as a poster presentation and part of a poster walk.
- *"He'll be home by his due date."* Multistate modelling to investigate neonatal length of stay. Presented at Epidemiology Congress of the Americas; Society for Perinatal and Paediatric Research (Miami, June 2016) as a poster presentation.
- When will my baby go home? Investigating neonatal care for preterm babies. Poster presentation (competitively selected) at the Postgraduate Festival of Research (University of Leicester, June 2016).
- Modelling neonatal care for preterm babies: an application of multistate modelling for neonatal length of stay. Oral presentation at Survival Analysis for Junior Researchers (University of Leicester, April 2017)

TALKS AND PRESENTATIONS

2014

- Modelling neonatal care pathways. Invited talk at Yorkshire and Humber Network meeting (Wakefield, November 2014)
- Modelling care pathways in neonatal care. Presentation at Health Sciences Departmental Conference (University of Leicester, November 2014)

2015

- Modelling neonatal care pathways. Probationary review presentation (University of Leicester, May 2015)

2016

- Modelling neonatal care pathways for preterm babies. Invited presentation to the National Neonatal Analysts Project (NNAP) Analysts meeting (London, February 2016)
- He'll be home by his due date: investigating neonatal care for preterm babies. Invited seminar at National Perinatal Epidemiology Unit (Oxford, February 2016)
- He'll be home by his due date: investigating neonatal care for preterm babies. Invited talk at Bliss (London, April 2016)
- Investigating care for preterm babies: an application of multistate modelling. Biostatistics seminar (University of Leicester, June 2016)
- Applying for an NIHR fellowship (NIHR@10 event, East Midlands, October 2016)

2017

- Investigating the care and length of stay of babies born preterm (University of Exeter, January 2017)
- When will my baby go home? Lunchtime lecture at Leicester Adult Education Centre (Leicester, May 2017)

PUBLICATIONS NOT RELATED TO THIS THESIS

Alongside this thesis I have also written the following papers in collaboration with my colleagues:

- Norris T, **Seaton SE**, Manktelow BN, Baker P, Kurinczuk JJ, Field DJ, Draper ES, Smith LK, on behalf of the MBRRACE-UK Collaboration. Updated birthweight centiles for England and Wales. Archives of Disease in Childhood: Fetal and Neonatal Edition (Online first)
- Manktelow BN; **Seaton SE**; Evans TA. Funnel plot limits to identify poorly performing healthcare providers when there is uncertainty in the value of the benchmark. Statistical Methods in Medical Research 2016; 25(6): 2670-2684.
- Boyle EM; Johnson S; Manktelow B; **Seaton SE**; Draper ES; Smith LK; Dorling J; Marlow N; Petrou S; Field DJ. Neonatal outcomes and delivery of care for infants born late preterm or moderately preterm: A prospective population-based study. Archives of Disease in Childhood Fetal & Neonatal Edition 2015; 100(6): F479-485.
- Guy A; **Seaton SE;** Boyle EM; Draper ES; Field DJ; Manktelow BN; Marlow N; Smith LK; Johnson S. Infants born late/moderately preterm are at increased risk for a positive autism screen at 2 years of age. Journal of Pediatrics 2015; 166(2): 269-275.
- Johnson S; Evans TA; Draper ES; Field DJ; Manktelow BN; Marlow N; Matthews R; Petrou S; **Seaton SE**; Smith LK; Boyle EM. Neurodevelopmental outcomes following late and moderate prematurity: Population-based cohort study. Archives of Disease in Childhood Fetal & Neonatal Edition 2016; 100(4): F301-308.
- Khan KA; Petrou S; Dritaki M; Johnson SJ; Manktelow B; Draper ES; Smith LK; **Seaton SE**; Marlow N; Dorling J; Field DJ; Boyle EM. Economic costs associated with moderate and late preterm birth: A prospective population-based study. BJOG: An International Journal Obstetrics and Gynaecology 2015; 122(11): 1495-505.
- Smith LK; Draper ES; Evans TA; Field DJ; Johnson SJ; Manktelow BN; **Seaton SE**; Marlow N; Petrou S; Boyle EM. Associations between late and moderately preterm birth and smoking, alcohol, drug use and diet: a population-based case-cohort.

Archives of Disease in Childhood: Fetal and Neonatal Edition 2015; 100(6): F486-491.

- Johnson S; **Seaton SE;** Manktelow BN; Smith LK; Field D; Draper ES; Marlow N; Boyle EM. Telephone interviews and online questionnaires can be used to improve neurodevelopmental follow-up rates. BMC Research Notes 2014; 7: 219.

MISCELLANEOUS

As a result of my work in this field I also undertook the following:

- Attended the report launch of *"Bliss baby report 2015: hanging in the balance"* as an invited guest of Bliss at the House of Common.
- Assisted the charity Together for Short Lives with the update of their guideline: *"A perinatal pathway for babies with palliative care needs"*

REFERENCES

1. National Audit Office. Caring for vulnerable babies: the reorganisation of neonatal services in England, 2007. Available online at: http://www.nao.org.uk/report/caring-for-vulnerable-babies-the-reorganisation-of-neonatal-services-in-england/ [Last accessed: 27/02/2017].
2. Mangham LJ, Petrou S, Doyle LW, Draper ES, Marlow N. The cost of preterm birth throughout childhood in England and Wales. Pediatrics. 2009;123(2):e312-27.
3. NNAP. National Neonatal Audit Programme Annual Report, 2016. Available online at: http://www.rcpch.ac.uk/improving-child-health/quality-improvement-and-clinical-audit/national-neonatal-audit-programme-nnap [Last accessed: 29/08/2017].
4. Lawn JE, Davidge R, Paul VK, Xylander S, de Graft Johnson J, Costello A, et al. Born Too Soon: Care for the preterm baby. Reproductive Health. 2013;10(1):S5.
5. Lee HC, Bennett MV, Schulman J, Gould JB, Profit J. Estimating length of stay by patient type in the neonatal intensive care unit. American Journal of Perinatology. 2016;33(8):751-7.
6. Hinchliffe SR, Seaton SE, Lambert PC, Draper ES, Field DJ, Manktelow BN. Modelling time to death or discharge in neonatal care: an application of competing risks. Paediatric and perinatal epidemiology. 2013;27(4):426-33.
7. Bliss. Bliss baby report: hanging in the balance, 2015. Available online at: http://www.bliss.org.uk/babyreport [Last accessed: 27/02/2017].
8. World Health Organisation. Preterm birth, 2015. Available online at: http://www.who.int/mediacentre/factsheets/fs363/en/ [Last accessed 13/09/2016].
9. Engle WA. Age terminology during the perinatal period. Pediatrics. 2004;114(5):1362-4.
10. Smith LK, Blondel B, Van Reempts P, Draper ES, Manktelow BN, Barros H, et al. Variability in the management and outcomes of extremely preterm births across five European countries: a population-based cohort study. Archives of Disease in Childhood - Fetal and Neonatal Edition. 2017 102:F400-F8.
11. Cummings J. Antenatal counseling regarding resuscitation and intensive care before 25 weeks of gestation. Pediatrics. 2015;136(3):588.
12. Smith LK, Draper ES, Manktelow BN, Pritchard C, Field DJ. Comparing regional infant death rates: the influence of preterm births <24 weeks of gestation. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2013;98(2):F103-7.
13. Nuffield Council on Bioethics. Critical care decisions in fetal and neoantal medicine: ethical issues, 2007. Available online at: https://nuffieldbioethics.org/wp-content/uploads/2014/07/CCD-web-version-22-June-07-updated.pdf [Last accessed: 20/07/2017].
14. Manktelow BN, Seaton SE, Field DJ, Draper ES. Population-based estimates of in-unit survival for very preterm infants. Pediatrics. 2013;131(2):e425-32.
15. Hintz SR, Poole WK, Wright LL, Fanaroff AA, Kendrick DE, Laptook AR, et al. Changes in mortality and morbidities among infants born at less than 25 weeks during the post-surfactant era. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2005;90(2):F128-33.
16. Department of Health Expert Working Group on Neonatal Intensive Care Services. Report of the neonatal intensive care services review group, 2003. Available online at:

http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4018744.pdf [Last accessed: 16/09/2016].
17. Hallsworth M, Farrands A, Oortwijn W, Hatziandreu E. The provision of neonatal services: data for international comparisons, 2007. Available online at: http://www.nao.org.uk/wp-content/uploads/2007/12/0708101_International_Comparisons.pdf [Last accessed: 16/09/2016].
18. Seaton SE, Barker L, Jenkins D, Draper ES, Abrams KR, Manktelow BN. What factors predict length of stay in a neonatal unit, a systematic review. BMJ Open. 2016;6(e010466).
19. Bender GJ, Koestler D, Ombao H, McCourt M, Alskinis B, Rubin LP, et al. Neonatal intensive care unit: predictive models for length of stay. Journal of Perinatology. 2013;33(2):147-53.
20. Costeloe KL, Hennessy EM, Haider S, Stacey F, Marlow N, Draper ES. Short term outcomes after extreme preterm birth in England: comparison of two birth cohorts in 1995 and 2006 (the EPICure studies). BMJ. 2012;345:e7976.
21. Moore T, Hennessy EM, Myles J, Johnson SJ, Draper ES, Costeloe KL, et al. Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies. BMJ. 2012;345:e7961.
22. British Association of Perinatal Medicine. BAPM categories of care, 2011. Available online at: http://www.bapm.org/publications/ [Last accessed: 20/02/2017].
23. British Association of Perinatal Medicine. Standards for hospitals providing neonatal intensive and high dependency care, 2001 and 2010 editions. Available online at: http://www.bapm.org/publications/ [Last accessed: 01/08/2017].
24. NHS England. Neonatal critical care - service specifications, 2015. Available online at: http://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/01/e08-serv-spec-neonatal-critical.pdf [Last accessed: 05/05/2015].
25. Davies A On Behalf of Neonatal Cinical Reference Group. Transitional care report, 2014. Available online at: http://www.wmscnsenate.nhs.uk/index.php/download_file/view/205/971/ [Last accessed: 21/12/2015].
26. Marlow N, Gill BA. Establishing neonatal networks: the reality. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2007;92(2):F137-42.
27. Seaton SE, Barker L. Modelling neonatal care pathways: investigating length of stay for preterm infants. Infant. 2016;12(3):87-90.
28. Seaton SE, Barker L, Draper ES, Abrams KR, Modi N, Manktelow BN. Modelling neonatal care pathways for babies born preterm: an application of multistate modelling. PLOS One. 2016;11(10):e0165202.
29. Hintz SR, Bann CM, Ambalavanan N, Cotten CM, Das A, Higgins RD, et al. Predicting time to hospital discharge for extremely preterm infants. Pediatrics. 2010;125(1):e146-54.
30. Richardson DK, Corcoran JD, Escobar GJ, Lee SK. SNAP-II and SNAPPE-II: Simplified newborn illness severity and mortality risk scores. Journal of Pediatrics. 2001;138(1):92-100.

31. Richardson DK, Gray JE, McCormick MC, Workman K, Goldmann DA. Score for Neonatal Acute Physiology: a physiologic severity index for neonatal intensive care. Pediatrics. 1993;91(3):617-23.
32. PROSPERO: International prospective register of systematic reviews. Available online at: http://www.crd.york.ac.uk/prospero/ [Last accessed: 08/09/2017].
33. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence based decisions. ACP Journal Club. 1995;123(3):12-3.
34. Organisation for Economic Co-operation and Development (OECD). OECD: better policies for better lives, 2014. Available online at: http://www.oecd.org [Last accessed: 27/02/2017].
35. University of York Centre for Reviews Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care: Centre for Reviews and Dissemination; 2009.
36. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Annals of Internal Medicine. 2013;158(4):280-6.
37. PRISMA. Preferred Reporting Items for Systematic Reviews and Meta-Analyses, 2015. Available online at: http://prisma-statement.org/ [Last accessed: 12/04/16].
38. Ambalavanan N, Carlo WA. Comparison of the prediction of extremely low birth weight neonatal mortality by regression analysis and by neural networks. Early Human Development. 2001;65(2):123-37.
39. Ambalavanan N, Carlo WA, Bobashev G, Mathias E, Liu B, Poole K, et al. Prediction of death for extremely low birth weight neonates. Pediatrics. 2005;116(6):1367-73.
40. Berry MA, Shah PS, Brouillette RT, Hellmann J. Predictors of mortality and length of stay for neonates admitted to children's hospital neonatal intensive care units. Journal of Perinatology. 2008;28(4):297-302.
41. Cole TJ, Hey E, Richmond S. The PREM score: a graphical tool for predicting survival in very preterm births. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2010;95(1):F14-9.
42. Evans N, Hutchinson J, Simpson JM, Donoghue D, Darlow B, Henderson-Smart D. Prenatal predictors of mortality in very preterm infants cared for in the Australian and New Zealand Neonatal Network. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2007;92(1):F34-F40.
43. Ge WJ, Mirea L, Yang J, Bassil KL, Lee SK, Shah PS. Prediction of neonatal outcomes in extremely preterm neonates. Pediatrics. 2013;132(4):e876-85.
44. Locatelli A, Roncaglia N, Andreotti C, Doria V, Doni D, Pezzullo JC, et al. Factors affecting survival in infants weighing 750g or less. European Journal of Obstetrics and Gynecology and Reproductive Biology. 2005;123(1):52-5.
45. Moro M, Figueras-Aloy J, Fernandez C, Domenech E, Jimenez R, Perez-Rodriguez J, et al. Mortality for newborns of birthweight less than 1500g in Spanish neonatal units (2002-2005). American Journal of Perinatology. 2007;24(10):593-601.
46. Parry G, Tucker J, Tarnow-Mordi W. CRIB II: an update of the clinical risk index for babies score. Lancet. 2003;361(9371):1789-91.
47. Shah PS, Ye XY, Synnes A, Rouvinez-Bouali N, Yee W, Lee SK, et al. Prediction of survival without morbidity for infants born at under 33 weeks gestational age: a user-friendly graphical tool. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2012;97(2):F110-5.

48. Tyson JE, Parikh NA, Langer J, Green C, Higgins RD. Intensive care for extreme prematurity - moving beyond gestational age. New England Journal of Medicine. 2008;358(16):1672-81.
49. Zernikow B, Holtmannspoetter K, Michel E, Pielemeier W, Hornschuh F, Westermann A, et al. Artificial neural network for risk assessment in preterm neonates. Archives of Disease in Childhood: Fetal and Neonatal Edition. 1998;79(2):F129-34.
50. King CP, da Silva O, Filler G, Lopes LM. Online calculator to improve counseling of short-term neonatal morbidity and mortality outcomes at extremely low gestational age (23-28 Weeks). American Journal of Perinatology. 2016;33(9):910-7.
51. Bolisetty S, Legge N, Bajuk B, Lui K. Preterm infant outcomes in New South Wales and the Australian Capital Territory. Journal of Paediatrics and Child Health. 2015;51(7):713-21.
52. Wilkinson DJ, Thiele P, Watkins A, De Crespigny L. Fatally flawed? A review and ethical analysis of lethal congenital malformations. BJOG. 2012;119(11):1302-8.
53. Claydon JE, Mitton C, Sankaran K, Lee SK. Ethnic differences in risk factors for neonatal mortality and morbidity in the neonatal intensive care unit. Journal of Perinatology. 2007;27(7):448-52.
54. Apgar V. A proposal for a new method of evaluation of the newborn infant. Curent Research in Anaethesia and Analgesia. 1953;32(4):260-7.
55. Wong D, Abdel-Latif M, Kent A. Antenatal steroid exposure and outcomes of very premature infants: a regional cohort study. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2014;99(1):F12-20.
56. Medlock S, Ravelli AC, Tamminga P, Mol BW, Abu-Hanna A. Prediction of mortality in very premature infants: a systematic review of prediction models. PLoS One. 2011;6(9):e23441.
57. Kaaresen PI, Dohlen G, Fundingsrud HP, Dahl LB. The use of CRIB (clinical risk index for babies) score in auditing the performance of one neonatal intensive care unit. Acta Paediatrica. 1998;87(2):195-200.
58. Altman M, Vanpee M, Cnattingius S, Norman M. Moderately preterm infants and determinants of length of hospital stay. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2009;94(6):F414-8.
59. Lee HC, Bennett MV, Schulman J, Gould JB. Accounting for variation in length of NICU stay for extremely low birth weight infants. Journal of Perinatology. 2013;33(11):872-6.
60. Manktelow BN, Draper ES, Field C, Field D. Estimates of length of neonatal stay for very premature babies in the UK. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2010;95(4):F288-92.
61. Zernikow B, Holtmannspotter K, Michel E, Hornschuh F, Groote K, Hennecke KH. Predicting length-of-stay in preterm neonates. European Journal of Pediatrics. 1999;158(1):59-62.
62. Freeman JV, Cole TJ, Chinn S, Jones PR, White EM, Preece MA. Cross sectional stature and weight reference curves for the UK, 1990. Archives of Disease in Childhood. 1995;73(1):17-24.
63. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080.

64. Einhorn HJ. Accepting error to make less error. Journal of Personality Assessment. 1986;50(3):387-95.
65. Pepler PT, Uys DW, Nel DG. Predicting mortality and length-of-stay for neonatal admissions to private hospital neonatal intensive care units: a Southern African retrospective study. African Health Sciences. 2012;12(2):166-73.
66. Shah S, Zemichael O, Meng HD. Factors associated with mortality and length of stay in hospitalised neonates in Eritrea, Africa: a cross-sectional study. BMJ Open. 2012;2(5).
67. Abo-Zaid G, Sauerbrei W, Riley RD. Individual participant data meta-analysis of prognostic factor studies: state of the art? BMC Medical Research Methodology. 2012;12:56.
68. Gale C, Morris I. The UK National Neonatal Research Database: using neonatal data for research, quality improvement and more. Archives of Disease in Childhood: Education and Practice Edition. 2016;101(4):216-18.
69. Imperial College London. Utilising the NNRD, 2017. Available online at: http://www.imperial.ac.uk/neonatal-data-analysis-unit/neonatal-data/utilising-the-nnrd/ [Last Accessed: 12/04/2017].
70. Phibbs CS, Schmitt SK. Estimates of the cost and length of stay changes that can be attributed to one-week increases in gestational age for premature infants. Early Human Development. 2006;82(2):85-95.
71. ACT: Valuing Short Lives. A Neonatal Pathway for Babies with Palliative Care Needs (First Edition, 2009). Available online at http://www.togetherforshortlives.org.uk/assets/0000/7095/Neonatal_Pathway_for_Babies_5.pdf [Last accessed: 9/9/2016].
72. North West Neonatal ODN. North West Perinatal/Neonatal Palliative Care Guideline, 2016. Available online at: http://www.neonatalnetwork.co.uk/nwnodn-default [Last accessed: 16/09/2016].
73. NHS Choices. Edwards' syndrome (Trisomy 18), 2014. Available online at: http://www.nhs.uk/conditions/edwards-syndrome/Pages/Introduction.aspx [Last accessed: 10/02/2017].
74. Joseph KS, Kramer MS, Allen AC, Mery LS, Platt RW, Wen SW. Implausible birth weight for gestational age. American Journal of Epidemiology. 2001;153(2):110-3.
75. Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. Annual Review of Public Health. 1997;18:83-104.
76. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Statistics in Medicine. 2007;26(11):2389-430.
77. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association. 1958;53(282):457-81.
78. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological). 1972;34(2):187-220.
79. Lambert PC, Royston P. Further development of flexible parametric models for survival analysis. Stata Journal. 2009;9(2):265-90.
80. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Statistics in Medicine. 2002;21(15):2175-97.

81. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. Springer: New York; 2003.
82. Breslow NE. Discussion following "Regression models and life tables" by David R Cox. Journal of Royal Statistical Society, Series B. 1972;34:187-220.
83. Kalbfleisch JD, RL P. Marginal Likelihoods Based on Cox's Regression and Life Model. Biometrika. 1973;60(2):267-78.
84. Breslow N. Covariance analysis of censored survival data. Biometrics. 1974;30(1):89-99.
85. Therneau T, Grambsch P. Modelling survival data: extending the Cox model. Springer, New York 2000.
86. Borucka J. Methods for handling tied events in the Cox proportional hazards model. Studia Oeconomica Posnaniensia. 2014;2(2):91-106.
87. Schoenfeld D. Partial residuals for the proportional hazards regression model Biometrika 1982;69(1):239-41.
88. Sauerbrei W, Royston P, Look M. A new proposal for multivariable modelling of time-varying effects in survival data based on fractional polynomial time-transformation. Biometrical Journal. 2007;49(3):453-73.
89. Royston P, Lambert PC. Flexible parametric survival analysis using Stata: beyond the Cox model. USA: Stata Press; 2011.
90. Durrleman S, Simon R. Flexible regression models with cubic splines. Statistics in Medicine. 1989;8(5):551-61.
91. Harrell FE. Regression modelling strategies: with applications to linear models, logistic regression and survival analysis. USA: Springer Series in Statistics; 2001.
92. Reid N. A Conversation with Sir David Cox. Statistical Science. 1994:439-55.
93. Altman DG, De Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. British Journal of Cancer. 1995;72(2):511-8.
94. Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994;81(3):515-26.
95. Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. Biometrics. 1978;34(4):541-54.
96. de Wreede LC, Fiocco M, Putter H. mstate: an R package for the analysis of competing risks and multi-state models. Journal of Statistical Software. 2011;38(7).
97. Bernoulli D. Essai d'une nouvelle analyse de la mortalite causee par la petite verole (English translation: 'An attempt at a new analysis of the mortality caused by smallpox'). Mémoires de Mathématiques et de Physique présentés à l'Académie Royale des Sciences. 1766.
98. Bernoulli D, Blower S. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. Reviews in Medical Virology. 2004;14(5):275-88.
99. Gail M. A review and critique of some models used in competing risk analysis. Biometrics. 1975;31(1):209-22.
100. Beyermann J, Allignol A, Schumacher M. Competing risks and multistate models with R. New York: Springer; 2013.
101. Hinchliffe SR. Advancing and appraising competing risks methodology for better communication of survival statistics: University of Leicester; 2013.

102. Fine JP, Gray R. A proportional hazards model for the subdistribution of a competing risk. Journal of the American Statistical Association. 1999;94(446):496-509.
103. Hinchliffe SR, Lambert PC. Flexible parametric modelling of cause-specific hazards to estimate cumulative incidence functions. BMC Medical Research Methodology. 2013;13(1):1-14.
104. Koller MT, Raatz H, Steyerberg EW, Wolbers M. Competing risks and the clinical community: irrelevance or ignorance? Statistics in Medicine. 2012;31(11-12):1089-97.
105. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974;19(6):716-23.
106. Cole TJ, Williams AF, Wright CM. Revised birth centiles for weight, length and head circumference in the UK-WHO growth charts. Annals of Human Biology. 2010;38(1):7-11.
107. Wolkewitz M, Vonberg RP, Grundmann H, Beyersmann J, Gastmeier P, Bärwolff S, et al. Risk factors for the development of nosocomial pneumonia and mortality on intensive care units: application of competing risks models. Critical Care. 2008;12(2):R44.
108. Brock GN, Barnes C, Ramirez JA, Myers J. How to handle mortality when investigating length of hospital stay and time to clinical stability. BMC Medical Research Methodology. 2011;11(1):144.
109. Scrucca L, Santucci A, Aversa F. Competing risk analysis using R: an easy guide for clinicians. Bone Marrow Transplant. 2007;40(4):381-7.
110. Hinchliffe SR, Lambert PC. Extending the flexible parametric survival model for competing risks. Stata Journal. 2013;13(2):344-55.
111. Petit N, Cammu H, Martens G, Papiernik E. Perinatal outcome of twins compared to singletons of the same gestational age: a case-control study. Twin Research and Human Genetics. 2011;14(1):88-93.
112. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. The International Journal of Epidemiology. 2012;41(3):861-70.
113. Gerds TA, Scheike TH, Andersen PK. Absolute risk regression for competing risks: interpretation, link functions, and prediction. Statistics in Medicine. 2012;31(29):3921-30.
114. Meira-Machado L, de Uña-Álvarez J, Cadarso-Suárez C, Andersen PK. Multi-state models for the analysis of time-to-event data. Statistical Methods in Medical Research. 2009;18(2):195-222.
115. Allignol A, Schumacher M, Beyersmann J. Estimating summary functionals in multistate models with an application to hospital infection data. Computational Statistics. 2011;26(2):181-97.
116. Iacobelli S, Carstensen B. Multiple time scales in multi-state models. Statistics in Medicine. 2013;32(30):5315-27.
117. Farewell VT, Cox DR. A note on multiple time scales in life testing. Journal of the Royal Statistical Society Series C (Applied Statistics). 1979;28(1):73-5.
118. Meira-Machado L, de Uña-Álvarez J, Cadarso-Suárez C. Nonparametric estimation of transition probabilities in a non-Markov illness–death model. Lifetime Data Analysis. 2006;12(3):325-44.
119. Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes. New York: Springer-Verlag; 1993.

120. Aalen OO, Johansen S. An empirical transition matrix for non-homogeneous markov chains based on censored observations. Scandinavian Journal of Statistics. 1978;5(3):141-50.
121. de Wreede LC, Fiocco M, Putter H. The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. Computer Methods and Programs in Biomedicine. 2010;99(3):261-74.
122. Beyermann J, Putter H. A note on computing average state occupation times. Demographic Research. 2014;30(62):1681-96.
123. Chernick MR. Bootstrap methods: a practitioner's guide1999.
124. Elliot R, Payne K. Essentials of Economic Evaluation in Healthcare. UK: Pharmaceutical Press; 2004.
125. Mani K, Lundkvist J, Holmberg L, Wanhainen A. Challenges in analysis and interpretation of cost data in vascular surgery. Journal of Vascular Surgery. 2010;51(1):148-54.
126. Battersby C, Michaelides S, Upton M, Rennie JM. Term admissions to neonatal units in England: a role for transitional care? A retrospective cohort study. BMJ Open. 2017;7(5):e016050.
127. Allen M, Spencer A, Gibson A, Matthews J, Allwood A, Prosser S, et al. Right cot, right place, right time: improving the design and organisation of neonatal care networks - a computer simulation study. Health Services and Delivery Final Report, 2015. Available online at: https://www.ncbi.nlm.nih.gov/books/NBK293953/ [Last accessed 01/08/2017]. 2015.
128. Marlow N, Bennett C, Draper ES, Hennessy EM, Morgan AS, Costeloe KL. 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. 2014;99(3):F181-8.
129. Bliss. Transfers of premature and sick babies, 2016. Available online at: http://www.bliss.org.uk/neonataltransport [Last accessed: 27/02/2017].
130. Developing Operational Delivery Networks: The Way Forward, 2012. Available online at: www.england.nhs.uk/2012/12/21/odn/ [Last accessed: 23/03/2015].
131. The Marmot Review. Fair society, healthy lives, 2010. Available online at: http://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review [Last accessed: 25/07/2017].
132. Seaton SE, King S, Manktelow BN, Draper ES, Field DJ. Babies born at the threshold of viability: changes in survival and workload over 20 years. Archives of Disease in Childhood - Fetal and Neonatal Edition. 2013;98(1):F15-20.
133. Manktelow BN, Smith LK, Seaton SE, Hyman-Taylor P, Kurinczuk JJ, Field DJ, et al. MBRRACE-UK Perinatal Mortality Surveillance Report: UK Perinatal Deaths for Births from January to December 2014. Leicester: The Infant Mortality and Morbidity Studies, Department of Health Sciences, University of Leicester; 2016.
134. Hougaard P. Frailty models for survival data. Lifetime Data Analysis. 1995;1(3):255-73.
135. Putter H, van Houwelingen HC. Frailties in multi-state models: Are they identifiable? Do we need them? Statistical Methods in Medical Research. 2015;24(6):675-92.
136. Clark DE, Ryan LM. Concurrent prediction of hospital mortality and length of stay from risk factors on admission. Health Services Research. 2002;37(3):631-45.

137. Taylor SL, Sen S, Greenhalgh DG, Lawless M, Curri T, Palmieri TL. A competing risk analysis for hospital length of stay in patients with burns. JAMA Surgery. 2015;150(5):450-6.
138. Mitchell BG, Ferguson JK, Anderson M, Sear J, Barnett A. Length of stay and mortality associated with healthcare-associated urinary tract infections: a multi-state model. Journal of Hospital Infection. 2016;93(1):92-9.
139. Clark DE, Ostrander KR, Cushing BM. A multistate model predicting mortality, length of stay, and readmission for surgical patients. Health Services Research. 2016;51(3):1074-94.
140. De Angelis G, Allignol A, Murthy A, Wolkewitz M, Beyersmann J, Safran E, et al. Multistate modelling to estimate the excess length of stay associated with meticillin-resistant Staphylococcus aureus colonisation and infection in surgical patients. Journal of Hospital Infection. 2011;78(2):86-91.
141. Cleves M, Gutierrez RG, Gould W, Marchenko YV. An introduction to survival analysis using stata. Stata Press, USA 2010.
142. Jackson C. flexsurv: a platform for parametric survival modeling in R. Journal of Statistical Software. 2016;70(8):33.
143. Neonatal Critical Care Minimum Data Set. Available online at: http://www.datadictionary.nhs.uk/data_dictionary/messages/supporting_data_sets/data_sets/neonatal_critical_care_minimum_data_set_fr.asp?shownav=1 [Last accessed: 04/05/2017].
144. Royal College of Obstetricians & Gynaecologists. Multiple Pregnancy: The Management of Twin and Triplet Pregnancies in the Antenatal Period. London: RCOG Press 2011.
145. Papiernik E, Zeitlin J, Delmas D, Blondel B, Kunzel W, Cuttini M, et al. Differences in outcome between twins and singletons born very preterm: results from a population-based European cohort. Human Reproduction. 2010;25(4):1035-43.
146. Titman AC. Model diagnostics in multi-state models of biological systems: University of Cambridge; 2007.
147. Bhutta ZA, Khan I, Salat S, Raza F, Ara H. Reducing length of stay in hospital for very low birthweight infants by involving mothers in a stepdown unit: an experience from Karachi (Pakistan). BMJ. 2004;329(7475):1151-5.
148. Scherf RF, Reid KW. Going home: what NICU nurses need to know about home care. Neonatal Network. 2006;25(6):421-5.
149. Melnyk BM, Feinstein NF, Alpert-Gillis L, Fairbanks E, Crean HF, Sinkin RA, et al. Reducing premature infants' length of stay and improving parents' mental health outcomes with the creating opportunities for parent empowerment (COPE) neonatal intensive care unit program: a randomized, controlled trial. Pediatrics. 2006;118(5):e1414-27.
150. Lee SK, O'Brien K. Parents as primary caregivers in the neonatal intensive care unit. CMAJ : Canadian Medical Association Journal. 2014;186(11):845-7.
151. O'Keeffe AG, Tom BDM, Farewell VT. Mixture distributions in multi-state modelling: some considerations in a study of psoriatic arthritis. Statistics in Medicine. 2013;32(4):600-19.
152. Ickowicz A, Sparks R. Modelling hospital length of stay using convolutive mixtures distributions. Statistics in Medicine. 2017;36(1):122-35.

153. Shetty S, Kennea N, Desai P, Giuliani S, Richards J. Length of stay and cost analysis of neonates undergoing surgery at a tertiary neonatal unit in England. Annals of the Royal College of Surgeons of England. 2016;98(1):56-60.
154. National Institute for Health and Clinical Excellence (NICE). Neonatal specialist care, 2010. Available online at: https://www.nice.org.uk/guidance/qs4/resources/neonatal-specialist-care-pdf-58296066757 [Last accessed 31/07/2017].
155. Walston F, Dixon V, May J, Harris S, Metayer L, Curley A. Bridging the gap: A survey of neonatal community care provision in England. Journal of Neonatal Nursing. 2011;17(2):69-78.
156. Langley D, Hollis S, Friede T, MacGregor D, Gatrell A. Impact of community neonatal services: a multicentre survey. Archives of Disease in Childhood - Fetal and Neonatal Edition. 2002;87(3):F204-8.
157. DeRienzo C, Kohler JA, Lada E, Meanor P, Tanaka D. Demonstrating the relationships of length of stay, cost and clinical outcomes in a simulated NICU. Journal of Perinatology. 2016;36(12):1128-31.
158. Brown S, Small R, Faber B, Krastev A, Davis P. Early postnatal discharge from hospital for healthy mothers and term infants. Cochrane Database of Systematic Reviews. 2002(3):Cd002958.
159. Metcalfe A, Mathai M, Liu S, Leon JA, Joseph KS. Proportion of neonatal readmission attributed to length of stay for childbirth: a population-based cohort study. BMJ Open. 2016;6(9):e012007.
160. Goyal N, Zubizarreta JR, Small DS, Lorch SA. Length of stay and readmission among late preterm infants: an instrumental variable approach. Hospital Pediatrics. 2013;3(1):7-15.
161. Tomashek KM, Shapiro-Mendoza CK, Weiss J, Kotelchuck M, Barfield W, Evans S, et al. Early discharge among late preterm and term newborns and risk of neonatal morbidity. Seminars in Perinatology. 2006;30(2):61-8.
162. Cedars-Sinai Medical Center. Hospital cuts length of stay for babies in the NICU by four days. Hospital Case Management. 2015;23(4):49-50.