The efficient use of the maternity workforce and the implications for safety and quality in maternity care: a population-based, cross-sectional study

Jane Sandall, Trevor Murrells, Miranda Dodwell, Rod Gibson, Susan Bewley, Kirstie Coxon, Debra Bick, Graham Cookson, Cathy Warwick and Diana Hamilton-Fairley

NHS

National Institute for Health Research

The efficient use of the maternity workforce and the implications for safety and quality in maternity care: a population-based, cross-sectional study

Jane Sandall,[1]* Trevor Murrells,[2] Miranda Dodwell,[1,3] Rod Gibson,[1,3] Susan Bewley,[1] Kirstie Coxon,[1] Debra Bick,[2] Graham Cookson,[4,5] Cathy Warwick[6] and Diana Hamilton-Fairley[7,8]

[1]Women's Health Academic Centre (King's Health Partners), Division of Women's Health, King's College London, London, UK
[2]Florence Nightingale Faculty of Nursing and Midwifery, King's College London, London, UK
[3]BirthChoiceUK, London, UK
[4]Department of Management, King's College London, London, UK
[5]Department of Health Care Management and Policy, University of Surrey, Guildford, Surrey, UK
[6]Royal College of Midwives, London, UK
[7]Guy's and St Thomas' NHS Foundation Trust, London, UK
[8]Health Education South London, London, UK

*Corresponding author

Declared competing interests of authors: none

This report should be referenced as follows:

Sandall J, Murrells T, Dodwell M, Gibson R, Bewley S, Coxon K, *et al*. The efficient use of the maternity workforce and the implications for safety and quality in maternity care: a population-based, cross-sectional study. *Health Serv Deliv Res* 2014;**2**(38).

Health Services and Delivery Research

Criteria for inclusion in the *Health Services and Delivery Research* journal

Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hsdr

This report

Abstract

The efficient use of the maternity workforce and the implications for safety and quality in maternity care: a population-based, cross-sectional study

Jane Sandall,[1]* Trevor Murrells,[2] Miranda Dodwell,[1,3] Rod Gibson,[1,3] Susan Bewley,[1] Kirstie Coxon,[1] Debra Bick,[2] Graham Cookson,[4,5] Cathy Warwick[6] and Diana Hamilton-Fairley[7,8]

[1]Women's Health Academic Centre (King's Health Partners), Division of Women's Health, King's College London, London, UK
[2]Florence Nightingale Faculty of Nursing and Midwifery, King's College London, London, UK
[3]BirthChoiceUK, London, UK
[4]Department of Management, King's College London, London, UK
[5]Department of Health Care Management and Policy, University of Surrey, Guildford, Surrey, UK
[6]Royal College of Midwives, London, UK
[7]Guy's and St Thomas' NHS Foundation Trust, London, UK
[8]Health Education South London, London, UK

*Corresponding author

Background: The performance of maternity services is seen as a touchstone of whether or not we are delivering high-quality NHS care. Staffing has been identified in numerous reports as being a critical component of safe, effective, user-centred care. There is little evidence regarding the impact of maternity workforce staffing and skill mix on the safety, quality and cost of maternity care in the UK.

Objectives: To understand the relationship between organisational factors, maternity workforce staffing and skill mix, cost and indicators of safe and high-quality care.

Design and methods: Data included Hospital Episode Statistics (HES) from 143 NHS trusts in England in 2010–11 (656,969 delivery records), NHS Workforce Statistics, England, 2010–11, Care Quality Commission Maternity Survey of women's experiences 2010 and NHS reference costs 2010/11. Ten indicators were derived from HES data. They included healthy mother and healthy baby outcomes and mode of birth. Adjustments were made for background characteristics and clinical risk. Data were analysed to examine the influence of organisational factors, staffing and costs using multilevel logistic regression models. A production function analysis examined the relationship between staffing, skill mix and output.

Results: Outcomes were largely determined by women's level of clinical risk [based on National Institute for Health and Care Excellence (NICE) guidance], parity and age. The effects of trust size and trust university status were small. Larger trust size reduced the chance of a healthy mother outcome and also reduced the likelihood of a healthy mother/healthy baby dyad outcome, and increased the chances of other childbirth interventions. Increased investment in staff did not necessarily have an effect on the outcome and experience measures chosen, although there was a higher rate of intact perineum and also of delivery with bodily integrity in trusts with greater levels of midwifery staffing. An analysis of the multiplicative effects of parity and clinical risk with the staffing variables was more revealing. Increasing the number of doctors had the greatest impact on outcomes in higher-risk women and increasing the number of midwives had the greatest impact on outcomes in lower-risk women. Although increased numbers of

support workers impacted on reducing childbirth interventions in lower-risk women, they also had a negative impact on the healthy mother/healthy baby dyad outcomes in all women. In terms of maximising the capacity of a trust to deliver babies, midwives and support workers were found to be substitutes for each other, as were consultants and other doctors. However, any substitution between staff groups could impact on the quality of care given. Economically speaking, midwives are best used in combination with consultants and other doctors.

Conclusions: Staffing levels have positive and negative effects on some outcomes, and deployment of doctors and midwives where they have most beneficial impact is important. Managers may wish to exercise caution in increasing the number of support workers who care for higher-risk women. There also appear to be limited opportunities for role substitution.

Future work: Wide variations in outcomes remain after adjustment for sociodemographic and clinical risk, and organisational factors. Further research is required on what may be influencing unexplained variation such as organisational climate and culture, use of NICE guidelines in practice, variation of models of care within trusts and women's choices.

Funding: The National Institute for Health Research Health Services and Delivery Research programme.

Contents

List of tables

List of figures

List of boxes

Glossary

Anti-D immunoglobulin Treatment administered in pregnancy or following birth to prevent creation of maternal antibodies to rhesus-positive (D) blood, in cases where there is a mismatch between maternal and infant rhesus-positive and -negative blood groups.

Apgar (score) Assessment of physical well-being conducted on baby immediately after birth, named after Dr Virginia Apgar.

Area under curve Area under the receiver operating curve. An area under curve of 0.5 is no better than tossing a coin, whereas an area under curve of 1 implies perfect prediction.

Augmentation Use of hormone infusion (oxytocin) [UK, Syntocinon®, Novartis Pharmaceuticals Ltd; US, Pitocin, Pfizer Ltd] to increase the number, strength and regularity of contractions, aiming to shorten the duration of labour.

β-coefficient Estimate of 'slope' derived from the regression model which measures the strength of the relationship between the predictor variable and the outcome.

Birthrate Plus Tool widely used to estimate midwifery and nursing staffing and skills mix requirements in maternity services.

Brachial plexus injury Damage to the newborn's brachial (arm) nerves, associated with interventions to resolve shoulder dystocia.

Breech presentation When fetus presents at the maternal pelvic outlet with buttocks or feet first, rather than the head.

Caesarean birth or caesarean Surgical delivery of baby through the mother's abdomen.

Case-mix adjustment Adjusting for differences in terms of clinical characteristics and other factors (e.g. age, gender and ethnicity).

Confidence interval An interval estimate of a population parameter, in this study typically regression coefficients. The 95% confidence interval (e.g. calculated using the mean proportion and its standard error) would contain the true value (e.g. population proportion) on 95 occasions (out of 100) if the same study were repeated 100 times on different randomly drawn samples [see Easton VJ and McColl JH. *Statistics Glossary v1.1*. University of Glasgow. URL: www.stats.gla.ac.uk/steps/glossary/confidence_intervals.html (accessed 4 July 2014)].

Cost function The function which represents the costs of production rather than the quantity as found in the production function.

Deep-vein thrombosis Formation of a clot in the venous system, normally in the veins of the legs.

Delivery with bodily integrity This term means that, following birth, the woman has not sustained any of the following: an abdominal wound (caesarean), an episiotomy (incision at the vaginal opening to facilitate birth), or a second-, third- or fourth-degree perineal tear. A first-degree tear is skin only, often does not require suturing and heals spontaneously; a second-degree tear involves injury to the perineum involving perineal muscles but not involving the anal sphincter; a third-degree tear involves partial or complete disruption of the anal sphincter muscles which may involve both the external and internal anal sphincter muscles; and a fourth-degree tear is where the anal sphincter muscles and anal mucosa have been disrupted.

Eclampsia Seizures in a pregnant or postnatal woman occurring as a complication of pre-eclampsia (gestational hypertension and/or proteinuria).

Econometric analysis The application of mathematics and statistics to apply economic models to data.

Effectiveness A measure of how good a unit is at producing a particular outcome.

Efficiency A measure of how efficient a unit of production (in this study a trust) is, where fully efficient would be a situation in which no more output can be produced without increasing inputs.

Elective caesarean Procedure planned prior to the onset of labour.

Emergency caesarean Procedure undertaken after labour has commenced.

Failure to rescue Refers to an episode during acute care where patient observations indicate deterioration over time but the service fails to respond in an appropriate or timely manner.

Fetal distress Clinical signs of hypoxia (oxygen deprivation) in the fetus prior to birth.

Funnel plots Graphical means of presenting dispersal of multiple data points.

Gestational age The age of the pregnancy measured in days and weeks since the start of the last menstrual period (or equivalent date 14 days before conception).

Hypoxic–ischaemic encephalopathy Brain injury caused by oxygen deprivation in the newborn.

Induction of labour Measures taken to artificially initiate onset of labour.

Instrumental birth Delivery assisted by ventouse or forceps.

Intrapartum Refers to care or events occurring during established labour.

Logistic regression/multiple logistic regression model Regression analysis when the dependent variable is binary (e.g. caesarean – yes/no).

Maternity support worker Clinical health-care assistant who is not a registered practitioner, but has received additional skills training and provides maternity care to women and babies under the supervision of a registered practitioner.

Monotonic relationship One which continues in the same direction (e.g. always increasing or always decreasing).

Multilevel The data are structured hierarchically (e.g. women within trusts).

Multilevel model The multilevel structure is incorporated into the regression model so that standard errors of the model parameters (see *β-coefficient*) are correctly estimated (i.e. clustering within trusts is accounted for).

Multiparous Term describing a pregnant woman who has previously given birth.

Nulliparous Term describing a pregnant woman who has not previously given birth.

Odds ratio The odds of an outcome occurring in one group (treatment) versus it occurring in another reference group (the control). In this study, the odds that an outcome (e.g. caesarean) happens to a woman who belongs to a particular group (e.g. those who have experienced one or more previous live births), compared with the odds of the outcome happening to women who belong to another group (e.g. those who have experienced no previous live births).

Operative delivery Caesarean or instrumental birth.

Other doctors Medical staff who are not appointed to consultant posts.

p-complementarity When a rise in the price of factor i leads to a fall in demand for factor j they are said to be p-complements.

p-value The probability that a calculated test statistic as large or larger occurred by chance alone. In regression analysis, the probability that the regression coefficient estimated (or one more extreme) would be observed if the population regression parameter were actually zero.

Parity Refers to the number of live births (> 24 weeks) a woman has had.

Post-term pregnancy Gestational age of > 42 weeks.

Pressure ulcer Ulcer (wound) on the skin, and sometimes involving tissues below the skin, caused by friction, shearing or pressure between bony prominence and bed or chair during periods of immobility.

Pre-term birth Birth occurring between 24 and 37 completed weeks of pregnancy.

Primary post-partum haemorrhage Blood loss exceeding 500 ml within 24 hours of birth.

Production function A function which relates input quantities to output quantities in production.

Pulmonary embolism See *Venous thromboembolism*.

q-complementarity When a rise in the quantity of factor i leads to a fall in demand for factor j they are said to be q-complements.

Random-effects meta-analysis A technique for combining the results of heterogeneous studies during a meta-analysis which uses a weighted average of the effect sizes based on study size.

Registrable birth If a baby is born alive, or stillborn, after 24 completed weeks, the birth may be registered.

Relative chi-squared An informal measure of model fit commonly used in health services research, which expresses the chi-squared statistic to the degrees of freedom. Commonly used cut-off points are 3, 4 and 5. The chi-squared for a multilevel model effect (e.g. mother's age) divided by its degrees of freedom (for categorical and ordinal variables the degrees of freedom are given by the number of groups minus 1).

Respiratory distress Clinical condition in which there is failure to acquire sufficient oxygen via the lungs.

Rhesus factor or disease An antigen found on some, but not all, red blood cells causes a condition in which antibodies in a rhesus-negative pregnant woman's blood attack her baby's blood cells, if the baby is rhesus positive, and this may cause jaundice and anaemia in the baby. It can be prevented with administration of anti-D immunoglobulin during pregnancy.

Risk adjustment A method for controlling for characteristics of the woman (e.g. age, parity, clinical risk, ethnicity) and/or the organisation (e.g. size, unit configuration, attached to a university) that may affect the probability of a particular outcome under study and are not controllable by the trust.

Sensitivity analyses Repeating the analysis in different ways to ascertain whether the results remain consistent or change.

Shoulder dystocia A complication of birth whereby the head is born but the baby's shoulders are stuck behind the bones of the pelvis.

Spontaneous vaginal delivery A birth that follows a labour of spontaneous onset (i.e. the labour is not induced), and occurs without ventouse, forceps or caesarean.

Stillbirth A baby born dead after 24 completed weeks of pregnancy.

Term pregnancy Gestational age between 37 and 42 completed weeks.

Thromboprophylaxis Measures taken to prevent the formation of venous thromboembolism.

Venous thromboembolism Blood clot formed in the veins, which may become detached and cause pulmonary embolism, or blockage of a vessel in the lungs, which is a life-threatening complication.

List of abbreviations

A&E	accident and emergency	NICE	National Institute for Health and Care Excellence
AHRQ	Agency for Healthcare Research and Quality	ONS	Office for National Statistics
AMU	alongside midwifery unit	OPCS-4	Office of Population Censuses and Surveys Classification of Interventions and Procedures version 4
AUC	area under the curve		
BR+	Birthrate Plus		
CC	complication or comorbidity	OR	odds ratio
CI	confidence interval	OU	obstetric unit
CQC	Care Quality Commission	PCT	primary care trust
CQUIN	Commissioning for Quality and Innovation	PPI	patient and public involvement
		PSI	patient-specific indicator
FMU	free-standing midwifery unit	QALY	quality-adjusted life-year
FTE	full-time equivalent	RCM	Royal College of Midwives
HES	Hospital Episode Statistics	RCOG	Royal College of Obstetricians and Gynaecologists
HRG	Health Reference Group		
ICD-10	*International Classification of Diseases* version 10	RCPCH	Royal College of Paediatrics and Child Health
IMD	Index of Multiple Deprivation	SANDS	Stillbirth and Neonatal Death Charity
IQI	Indicators for Quality Improvement		
MCWP	Maternity Care Working Party	SAS	Statistical Analysis System
NCT	National Childbirth Trust	SD	standard deviation
NHS IC	NHS Information Centre	SE	standard error
NHS QUEST	NHS Quality, Efficiency and Support Team	SHA	Strategic Health Authority
		VIF	variance inflation factor

Plain English summary

Maternity services are often seen as a test of whether or not we are delivering high-quality NHS care. Little is known, however, about the best way to organise obstetricians, midwives and support staff to get high-quality maternity care at the best cost. We used routinely collected data for the 650,000 women who gave birth in NHS hospitals in England in 2010/11 to answer this question.

We looked at the effects that different ways of staffing maternity units had on women and their babies, for example if both were healthy after the birth, if harm was avoided and the type of birth.

We found that the biggest influence on all of these outcomes was whether or not a woman had any conditions which might increase her risk of complications during the birth, her age and whether or not she had had a baby before. Hospital size had some effect. Better outcomes were achieved by increasing the number of obstetricians looking after women at higher risk, and increasing the number of midwives looking after lower-risk women. When more maternity support workers were employed, both women and their babies were less likely to have a healthy outcome. In trusts with higher levels of midwifery staffing, women were more likely to avoid bodily harm (surgery or tearing that required stitches) during birth. However, we found that, in general, there was no relationship between how much a trust spent on providing care for women giving birth and the outcomes for those women. We conclude that staffing levels have positive and negative effects on some outcomes, and employing obstetricians and midwives where they have most beneficial impact is important. Managers may wish to be cautious about employing more support workers in settings where many women are at higher risk, especially when both numbers and complexity of births are increasing.

We need to find out more about what else is influencing the large differences in outcomes in different hospitals when women who have similar risks are compared.

Scientific summary

Background

The performance of maternity services is seen as a touchstone of whether or not we are delivering high-quality NHS care. Staffing has been identified in numerous reports as being a critical component of safe, effective, user-centred care. Large variations in outcomes and of women's experiences exist by NHS trust and there is little empirical evidence regarding the impact of maternity workforce staffing and skill mix on the safety, quality and cost of maternity care in the UK.

Although maternity care is the commonest reason for hospital admission among women aged 15–49 years in the UK, there is lack of agreement on what measures should be used. Safety measures developed for general populations often do not include measures appropriate for pregnant or postnatal women. There are also methodological issues to address, for example population, context of care, data quality, variation in outcomes within and between units, and whether or not risk adjustment is used to address confounding factors.

Hospital Episode Statistics (HES) provides information on care provided by NHS hospitals and for NHS hospital patients treated elsewhere in England. Several studies have used HES data, and show that, despite data quality issues, analyses using judicious cleaning and case-mix adjustment can be useful in identifying variations in patterns of maternity care. All studies have found unexplained variation in a range of indicators across NHS trusts after adjusting for background characteristics and case mix and suggest further research to understand the impact of organisational and staffing factors, and the impact on cost.

Objectives

The aim of this project is to understand the relationship between organisational factors, maternity workforce staffing, skill mix, cost and indicators of safe and high-quality care. Therefore, this research aims to answer the following questions:

- How do organisational factors affect variability in maternal interventions and maternal and perinatal outcomes?
- What is the relationship between maternity staffing, skill mix and maternal and perinatal outcomes?
- What is the relationship between maternity skill mix, cost and outcomes?

Methods

Data
We acquired routinely gathered data for 2010–11 from NHS and other sources.

Data cleaning
The study restricted the records to NHS hospital deliveries resulting in a registrable birth. Duplicate delivery records were removed from the mothers' records. The babies' birth records also contained duplicate records. These duplicates were not removed, as the majority of the project's work concentrated on the mother's delivery record and the resources were not available to clean both. We consulted with the CQC regarding the Maternity Data Quality – Indicator specifications for maternity-related measures included within its surveillance programme. Where there were multiple records relating to the same

delivery episode, one record was chosen by using a scoring system that selected the record which contained information most relevant to the project.

We used the full census of women's deliveries in HES, which contains 656,969 delivery records, so there was no bias caused by non-response. Any biases would therefore be caused by missing data, poorly recorded data or omitted variables from the risk adjustment model. Sample sizes for some indicators were reduced by the choice of denominator to create the indicator.

Outcome measures

The selection of indicators was decided in consultation with the advisory group and informed by needing to have a balance of positive and negative indicators, the importance to women, costs, and the availability and quality of coding within the HES data set. Three indicators were derived to indicate a healthy mother and healthy baby, thus reflecting a concept of harm-free care, avoidance of longer-term morbidity, and a positive outcome. The mode of birth indicators were chosen to compare important processes and outcomes across trusts and with other studies. Ten final indicators (some were composites) that measured maternal and infant outcomes were developed. A decision was taken not to include any trust where fewer than 80% of women could be coded for a particular indicator. Combining this with missing postcode data meant that the percentage of delivery records used in the multilevel models was as follows: healthy mother and healthy baby indicators ($n = 431,391$; 66%), normal birth ($n = 467,022$; 71%), intact perineum ($n = 439,730$; 67%, or 89% of women who had a vaginal delivery), and delivery with bodily integrity, spontaneous vaginal delivery, elective caesarean, emergency caesarean and all caesareans ($n = 584,435$; 89%).

Independent variables

We included mothers' characteristics measured at the individual level known to affect the outcomes of interest. These included age, parity, ethnicity, area socioeconomic deprivation as measured by the Index of Multiple Deprivation, geographical location (urban/rural) and region.

Characteristics measured at trust level included size measured by number of deliveries, teaching status, maternity configuration (drawing on Birthplace in England typology of whether or not alongside and free-standing midwife units are part of trust provision) and staffing variables. These included staffing levels [full-time equivalent (FTE) obstetric medical staff, midwives and maternity support staff/100 maternities, FTE all staff/100 maternities] and skill mix (doctor/midwife and midwife/support worker ratios).

Level of clinical risk was measured according to the National Institute for Health and Care Excellence (NICE) intrapartum care guideline [NICE. *Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth*. NICE Clinical Guideline 55. 2007. URL: www.nice.org.uk/guidance/cg55/resources/guidance-intrapartum-care-pdf (accessed 10 september 2014)] by allocating women retrospectively to lower-risk, individual assessment and higher-risk status at the end of pregnancy. Conditions listed in the NICE intrapartum care guideline were matched by one of the project team to relevant *International Classification of Diseases* version 10 (ICD-10) codes, of which there are about 12,000. For certain conditions, other types of codes were matched, such as Office of Population Censuses and Surveys or HES Data Dictionary data items. Matching of codes was checked by another member of the project team, who is an obstetrician, and disagreements were resolved by consensus.

Analysis

Multilevel logistic regression models, where mothers (deliveries) were nested within trust, were fitted to the 10 maternity (dichotomous) indicators.

Three sensitivity analyses were performed (1) on trusts with a single obstetric unit, (2) on different levels of NICE risk assessment and (3) to test the effect of increasing/decreasing staffing levels on outcomes by parity and clinical risk group.

The economic analysis attempted to model maternity services from a production economics perspective to identify the relationship between inputs and outputs, and more specifically for this study the relationships between the different labour inputs. In particular, we wanted to establish the extent to which the different roles are substitutes or complements, that is competing or aiding inputs.

Results

Women's outcomes were largely determined by their clinical risk (based on NICE guidance), parity and age. Outcomes did vary by deprivation and ethnicity but these effects were lower by an order of magnitude. The effects of trust size and university status were small. Larger trust size reduces the chance of caesarean, but also reduces the chance of a healthy mother and healthy baby outcome(s) and increases the chances of other childbirth interventions.

Approximately 1–2% of the total variation in the outcome indicators was attributable to differences between trusts, whereas 98–99% of the variation was attributable to differences between mothers within trusts. The linear effects of the staffing variables were not statistically significant for eight indicators. The exceptions were as follows: increasing the number of midwives improved a woman's chance of delivering with bodily integrity and having an intact perineum. An analysis of the multiplicative effects of parity and clinical risk with the staffing variables was more revealing. Increasing the number of doctors has the greatest impact on outcomes in high-risk women and increasing the number of midwives has the greatest impact in low-risk women. However, caution needs to be exercised with increase in the number and deployment of support workers. Although increased numbers of support workers impacted on reducing childbirth interventions in low-risk women, it also had a negative impact on the healthy mother and healthy baby outcome(s) in all women.

We used trust-level data to investigate relationships between outcome measures, midwifery staffing levels and the cost of providing maternity services for NHS trusts in England. Higher midwifery staffing levels were associated with higher costs of each delivery, although the relationship was not strong. Only around 17% of the variation among trusts' delivery costs could be accounted for by variables included in this model. The remaining variation in the average cost of a delivery was not accounted for by maternal characteristics, size of trust, number of FTE registered midwives employed or antenatal spend and must be due to other factors not included in the analysis.

After adjusting for maternal characteristics and trust size, no relationship was found between the proportion of expenditure spent on antenatal care and operative delivery rates, or between higher operative delivery rates and higher delivery costs. Variations in costs between trusts were not related to the numbers of women having operative deliveries.

There was no association between cost per delivery and the normal birth rate, intact perineum rate, or any of the three healthy mother and healthy baby indicators, and women's experience of maternity care as measured by the average of the CQC scores. A relationship could not be found that explained postnatal costs in terms of variations in operative delivery rates once adjustments were made.

Having a higher proportion of women at increased clinical risk was associated with more expensive maternity care, as was the level of area deprivation, which approached statistical significance. These factors are currently reflected in the maternity pathway payment system. There appeared to be economies of scale across the total maternity episode (antenatal, intrapartum and postnatal care) by trust size, which were increased in trusts on a single site. However, larger trusts were associated with lower scores for women's experience, although no relationship was found between trust size and any clinical outcome.

From this study, the increased investment in staff did not necessarily have an effect on the outcome and experience measures chosen, apart from a higher intact perineum rate and higher levels of bodily integrity in trusts with higher levels of midwifery staffing.

The economic modelling analysis found midwives to be complements with both consultants and other doctors in the production of deliveries; that is, they should be used in combination. Consultants and other doctors were found to be substitutes for each other. Midwives and support workers were also found to be substitutes for each other. A major limitation is that we were not able to analyse impact of medical staffing, because of problems in combined data for obstetricians and gynaecologists.

Conclusions

There is some indication that staffing levels have positive and negative effects on some outcomes, and deployment of doctors and midwives where they have the impact that is most beneficial is important. Managers may wish to exercise caution in increasing more support workers in high-risk settings. There appears to be little opportunity for role substitution.

There are wide variations in a range of outcomes that remain after adjustment for sociodemographic and background risk. Further research is required on what may be influencing unexplained variation, such as organisational climate and culture, use of NICE guidelines in practice and variation of models of care within trusts.

Organisational factors, such as trust configuration, size, models of care, staffing levels, skill mix, staff deployment and safety culture, and women's choices remain unknown factors in the understanding of important influences of the quality and safety of maternity care. Better staffing data, especially in relation to consultant activity, would enable a more definitive analysis.

Funding

Chapter 1 Background and research objectives

Introduction

The chapter commences with an overview of drivers for high-quality maternity care, followed by a discussion of evidence relevant to defining and measuring quality and safety in maternity care, use of routine data, maternity health-care workforce, quality and safety indicators and health-care workforce and efficiency, and other literature to inform study aims and objectives.

Several bibliographic databases, including PubMed, The Cochrane Library, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), EconLit and that of the National Institute for Health and Care Excellence (NICE), were searched for relevant primary and secondary studies and guidelines. Search terms included 'maternity care', 'safety', 'quality', 'outcomes', 'midwifery care', 'obstetric care', 'acute hospital settings', 'maternity workforce', 'cost effectiveness', 'efficiency', 'production function' and 'stochastic frontier'. No date limits were used and only studies published in English were considered. Relevant policy reports were identified from searches of relevant websites of government and professional organisations.

As a systematic review was not conducted, a formal search strategy was not developed; however, priority was given to evidence relevant to UK-based maternity care, including studies undertaken in high-income countries with similar maternity systems (i.e. where midwives and obstetricians were responsible for providing care) which considered quality and safety of maternity care. Evidence relevant to health-care workforce issues that may or may not have been conducted in UK maternity settings was also considered.

High-quality maternity care

All NHS providers have a mandate to enhance the quality of patient care, and the performance of maternity services has been viewed as a window into whether or not quality health services in general are being delivered.[1] The Department of Health's *National Service Framework for Children, Young People and Maternity Services*,[2] with its 10-year time frame for implementation, and *Maternity Matters*[3] were consistent in the commitment to deliver a choice of safe, accessible, high-quality maternity care which was women focused and family centred. Underpinning principles included the view that pregnancy and birth are normal life events, maximising the opportunity for all women, regardless of risk profile, to have as physiological and positive a birth experience as possible. The coalition government has not developed or published a cohesive maternity policy but has published indications of its commitment to women's choice of maternity care,[4,5] continuity, and improved outcomes.[6]

The professional commitment to improving the quality of maternity care, and sustaining the workforce to achieve this, has been led by the relevant Royal Colleges, including the Royal College of Midwives (RCM) and Royal College of Obstetricians and Gynaecologists (RCOG), often with joint College publications. In 2007, a joint report was published by the RCOG, the RCM, the Royal College of Anaesthetists and the Royal College of Paediatrics and Child Health (RCPCH) to provide guidance to develop equitable, high-quality standards for UK maternity care.[7] Standard 30 (out of 30) focused on provision of a high-quality workforce and promotion of appropriate leadership, skill mix and competencies in midwifery, obstetrics, anaesthetics and paediatrics. A 2009 position statement from the RCM on staffing standards in midwifery services to assist commissioners and providers, endorsed by the RCOG and RCPCH, recommended a minimum ratio of 1 midwife per 28 births per year. Falling below this ratio would be a strong indication that a service should undertake a thorough workforce review.[8] The RCOG report *High Quality Women's Health Care*[9] emphasised how high-quality care could promote health over the life course for women and their infants. It proposed changes that focused 'on the needs of the woman and her baby by providing the right care, at the right

time, in the right place, provided by the right person and which enhances the woman's experience' (Foreword, page iv), and the fundamental role that midwives have in delivering high-quality health care also continues to be recognised.[10]

Defining quality

One of the key issues to address in any health system is how to derive robust, appropriate and usable measures of the quality and safety of care and measures of outcomes of care which resulted in harm. The ideal measure should be easy to define and observe, should reflect priority outcomes for patients, clinicians and service providers and should identify those areas where quality of care/outcome could be improved.

The Institute of Medicine[11] identified six dimensions of quality: namely that health care should be safe, effective, patient-centred, timely, efficient and equitable. Some of these have been widely adopted within the UK.[12–14] The current UK government has taken the approach that process indicators or targets are unnecessarily bureaucratic and distract from the important objectives of improving safety, reducing morbidity and improving patient experience more broadly, and has therefore focused quality measurement on evidence-based clinical outcomes.[15] The resulting NHS Outcomes Framework provides a national overview of the quality of NHS care, provides accountability and acts as a catalyst for driving quality improvement involving five domains:[16]

1. preventing people from dying prematurely
2. enhancing quality of life for people with long-term conditions
3. helping people to recover from episodes of ill health or following injury
4. ensuring that people have a positive experience of care
5. treating and caring for people in a safe environment and protecting them from avoidable harm.

The NHS Outcomes Framework for 2013/14 introduced new outcome measurements.[6] Improvement areas in maternity care and their relevant indicators include:

- reducing deaths in mothers (or at least maintaining the low level) (domain 1)
- reducing deaths in babies – neonatal deaths and stillbirths (domain 1)
- helping women to recover from ill health or injury following birth (domain 3)
- improving women and their families' experience of maternity services – women's experience of maternity services (domain 4)
- treating and caring for people in a safe environment and protecting them from avoidable harm (domain 5).

The maternity indicators identified by the Outcome Framework, whilst important in determining quality of care, cover only a small aspect of maternity services, and other bodies have also outlined key indicators. For example, the UK-wide Midwifery 2020 recommended in its report *Delivering Expectations*[13] a number of additional indicators to measure the quality of midwifery care. These included reducing perineal trauma; uninterrupted skin-to-skin contact between mother and baby following the birth; continuity of midwife care; and increasing the normal birth rate, using the definition of normal birth published by the Maternity Care Working Party in its consensus statement.[17] The Commissioning for Quality and Innovation (CQUIN) framework was introduced in 2009 to enable commissioners of health services to reward excellence by giving financial incentives to local health-care providers to deliver nationally agreed quality improvements and better outcomes for patients. There are four National CQUIN Scheme goals in 2013/14.[18] These are the Family and Friends Test (introduced for maternity in October 2013), dementia care, venous thromboembolism and the NHS Safety Thermometer,[19] of which there is a maternity version currently in development by the NHS Quality, Efficiency and Support Team (NHS QUEST). Local CQUIN schemes agreed with local commissioners should be in place in late 2013, and collaboration is encouraged where contracts are made with several

commissioners. Commissioners are encouraged to use appropriate existing indicators such as the Indicators for Quality Improvement (IQI),[20] Advancing Quality Alliance,[21] the NHS Atlas of Variation and NICE Quality Standards.[22] Maternity specific indicators used for CQUIN include IQI indicators for smoking cessation during pregnancy, prevalence of breastfeeding at 6–8 weeks and access to maternity services by 12 weeks + 6 days.[20] Indicators of quality, therefore, should measure a balance of aspects of harm or adverse outcomes, care which promotes health and patient-derived measures of the experience of care.

Measuring quality

Measuring quality of health care is not a new concept, and Donabedian described his approach in a paper published in the *Journal of the American Medical Association*.[23] It was based on three components: structure, process and outcome, which were viewed as inter-related. Structure referred to the conditions under which care was provided, including staffing levels and mix, facilities and equipment, and organisational characteristics such as supervision and performance review. Process referred to activities carried out and care provided, such as diagnosis, treatment and education. These activities are often carried out by staff. Outcomes referred to changes – good or bad – in individuals which could be attributed to the health care received. Donabedian differentiated these *measures* from *actual* aspects of quality, instead considering them to be alternative types of information which could be used to *infer* or *indicate* good quality. He highlighted the necessity of using these only when there was a relationship of cause and effect between the three components.

The King's Fund report *Getting the Measure of Quality*[24] still considered this approach to be useful in developing quality indicators, although the RCOG has pointed out that some obstetric indicators are hard to classify in this way[25] – for example, a caesarean may be a process initiated by a clinician or an outcome following another intervention and may have a good or bad effect (or both) on the health of two individuals, mother and baby.

Structural characteristics – the way a health-care system is organised – may have an important impact on the quality of care provided. The Francis Report[26] has highlighted how failings in management and governance caused serious failings in the processes and outcomes of care, but the direct relationship between these can be difficult to assess.

Where it is known that particular processes, such as continuity of care, are signifiers of quality of care, for example because they impact on some aspect of quality such as safety or patient experience, the measurement of the extent of that process can be used as an indicator of quality.[23] Process measures have the advantage of directly measuring care that is received by patients and potentially increasing the detection of poor care. They are capable of being measured contemporaneously, giving a more immediate assessment of quality. However processes can also be hard to measure and data may not be available. They can be subject to manipulation or 'gaming', particularly when performance is being assessed externally or financial incentives are at stake.

Clinical outcome measures have advantages over process measures in that they can assess the health outcomes (favourable or unfavourable) of patients who have received care. Outcomes are often routinely collected, making the data more readily available for analysis. They can be less subject to manipulation, but have the disadvantage of being affected by factors other than care, such as coding accuracy, disease severity, comorbidities and other independent characteristics or demographics. As it is not always possible to directly attribute outcomes to specific processes of care, in reality it may be more difficult than assumed to use outcome indicators to improve the quality of care in this way.

Measuring quality, safety and harm in maternity care

Although maternity care is the commonest reason for hospital admission among women aged 15–59 years in the UK,[27] there is lack of agreement on what measures should be used. Safety measures developed for general populations often do not include measures appropriate for pregnant or postnatal women.[28] There are also methodological issues to address, for example population, context of care, data quality, variation in outcomes within and between units, and whether or not risk adjustment was used to address confounding factors.

It is increasingly recognised that measures traditionally used in maternity care such as maternal and neonatal mortality are essential, but also that other measures of quality of care are needed.[29] Studies to date which have reported the development of a quality measure for maternity care have tended to focus on patient-specific indicators (PSIs), such as primary caesarean delivery rate,[30] or obstetric trauma in caesarean, instrumental and unassisted deliveries [used by the Agency for Healthcare Research and Quality (AHRQ)[31]] or aspects of patient satisfaction.[32]

Several studies have considered the risk of interventions in maternity care (e.g. caesarean birth), which could inform their use as a potential measure of quality. Paranjothy *et al.*[33] from the NICE Collaborating Centre for Women and Children examined the variation in caesarean rates between maternity units looking at case-mix differences in a national prospective cross-sectional study. Data were collated from 216 maternity units in England and Wales on women who gave birth between May and July 2000. The relationship between case-mix characteristics and odds of a caesarean birth before or during labour was investigated using logistic regression models. Overall caesarean rates standardised for case mix were then calculated for each maternity unit. Heterogeneity between units was examined using random-effects meta-analysis. Adjustment for case-mix differences explained 34% of the variance in caesarean rates. The odds of having a caesarean birth before and during labour increased with maternal age. Women from ethic minority groups had lower odds of caesarean birth before labour and increased odds in labour. Women who had a previous vaginal birth had lower odds of caesarean, although the magnitude of this for caesarean before and during labour was markedly different. Findings showed that the variation in organisation of services, women's preferences for mode of birth, staffing levels and clinician attitudes were all important factors to consider when quality and appropriateness of maternity care are evaluated.

Using routine data

Hospital Episode Statistics (HES) provides information on care provided by NHS hospitals and for NHS hospital patients treated elsewhere in England. It is the data source for a wide range of health-care analysis for the NHS, government and many other organisations such as the private benchmarking services CASPE Healthcare Knowledge Systems and Dr Foster.

Every episode of hospital inpatient care generates a patient record which includes demographic information along with details of the episode of care, diagnoses using *International Classification of Diseases* version 10 (ICD-10) codes[34] and procedures such as operations using Office of Population Censuses and Surveys Classification of Interventions and Procedures (OPCS-4) codes.[35] Where a woman's record additionally includes the delivery of a baby, limited further information on the birth is collected in a 'maternity tail' to the mother's record. Individual HES records for the same person can be linked across time and providers to enable long-term patterns of care to be studied. The individual records of mother and baby are not linked.

Some researchers have compared outcomes of the US AHRQ PSIs, including obstetric trauma indicators, with routinely collected HES inpatient data, in an attempt to generate quality indicators for England. Raleigh *et al.*[36] compared UK and US data for nine PSIs using a case–control analysis of HES data for 2003–4, 2004–5 and 2005–6 for all English trusts. Length of stay and mortality between cases

(patients experiencing the particular safety event measured by an indicator) and controls were matched for age, sex, health resource group (standard groupings of clinically similar treatments that use similar levels of health-care resource), main specialty and trust. They found some consistency in national rates for the nine indicators and, for all but one indicator, hospital stay and mortality were longer. The authors concluded that internationally comparative indicators could be derived from English data, although further validation was needed, and recording needed to be improved.

Bottle and Aylin[37] also compared outcomes for nine AHRQ indicators (10 were originally selected but one – iatrogenic pneumothorax – was dropped for lack of equivalent ICD-10 code) for use in English routine data in relation to established measures of negative outcome including mortality. Using case-mix adjustment they found wide variations between trusts which were potentially due to inadequate adjustment, differences in coding definitions between trusts, or poor quality of coding. They concluded that the derivation of patient safety indicators from HES data were potentially useful for prospective evaluation of data quality.

Studies using HES data show that despite data quality issues, analyses using judicious cleaning and case-mix adjustment can be useful in identifying variations in patterns of maternity care. Some of the results showing the importance of case-mix adjustment using HES data have been validated using data from the Millennium Cohort Study.[38] This study showed that, for women having their first baby, operative birth rose with increasing maternal age, and that for all women mode of birth differed significantly by ethnicity.

Obstetric outcomes derived from HES inpatient maternity data from 146 English NHS trusts were analysed by Bragg *et al.*[39] in a cross-sectional analysis to ascertain if variation in unadjusted rates of caesarean births could be explained by maternal characteristics and clinical risk factors. The main outcome measure was rate of caesarean birth per 100 births (live or stillborn). The population included women aged 15–44 years who had a singleton birth between 1 January and 31 December 2008. The likelihood of women having a caesarean birth given their age, ethnicity, parity, socioeconomic status, deprivation status and clinical risk factors (including previous caesarean birth, breech presentation and fetal distress) were entered into a multiple logistic regression model.

A total of 147,726 (23.8%) women had a caesarean birth, which was more likely if they had a previous caesarean or a breech presentation. Elective and emergency caesarean rates by NHS trust were adjusted for maternal characteristics and clinical risk factors, using funnel plots to show significant variation between trusts. Maternal age, ethnicity and parity (particularly in multiparous women with a previous caesarean) were all significant factors in determining the likelihood of a caesarean birth. A number of clinical risk factors, including diabetes, hypertension and placental problems, also predisposed women to having a caesarean. However, adjusting for maternal characteristics and clinical risk factors did not greatly reduce the variation between individual trusts, with the observed variation in caesarean rates being 14.9–32.1%. This variation was largely due to emergency caesarean rates, rather than elective caesarean rates, which showed much less variation. The authors concluded that case-mix adjustment was necessary in order to compare caesarean rates between trusts, and that the remaining variation may be due to differences in clinical practice regarding emergency caesarean between trusts.

This work was furthered by the RCOG in its Clinical Indicators Project,[25] which identified 11 potential performance indicators derived from HES data, the basis for selection being validity (reflecting quality of care), fairness, sufficient statistical power and ability to technically code the outcome adequately. In addition, the suite of indicators had to cover various dimensions of care to give a balanced picture of the service. The RCOG identified key issues with HES maternity data quality, particularly duplicate records, records not relating to deliveries and incomplete or inconsistent recording of data items. Despite extensive data cleaning to remove duplicates and records which did not relate to a delivery episode, identifying units with inconsistent or missing data and adjusting for case mix, this limited the use of some potential indicators. Results were stratified between nulliparous and multiparous women and adjusted for maternal

age, ethnicity, social deprivation and a number of clinical risk factors such as previous caesarean, diabetes, hypertension, gestational age and birthweight. The results were shown in funnel plots and demonstrated a large variation in intrapartum processes and outcomes which could not be explained by random fluctuations. For example, among women giving birth for the first time, there was a twofold difference between hospitals with the highest and lowest rates of emergency caesarean after induction of labour (20% compared with 40%) and of instrumental birth (16% compared with 32%). The report concluded that further understanding of the unexplained variation is important in order to compare performance across trusts.

Organisational factors, such as trust configuration, size, models of care, staffing levels, skill mix, staff deployment and safety culture, remain unknown factors in the understanding of important influences of the quality and safety of maternity care, some of which are considered in the next section.

The maternity health-care workforce and quality and safety indicators

'The performance of maternity services is a touchstone of whether we are delivering quality based on patient safety, effectiveness of care and patient experience.'[40] This statement from the NHS Chief Executive, David Nicholson, followed the review of maternity services carried out by the Healthcare Commission [now the Care Quality Commission (CQC)] in 2007. This review raised key concerns that, in some trusts, levels of staffing were well below average and may have been inadequate. It recognised that staffing was a contentious issue, as it underpins the quality of the service but at the same time is the most costly element of providing that service.[41]

In 2008 the report of the independent inquiry commissioned by The King's Fund, *Safe Births: Everybody's Business*,[42] and the Healthcare Commission's review of maternity services, *Towards Better Births*,[41] identified areas in need of improvement, including staffing, training and communication. Staffing has been identified in numerous reports as being a critical component of safe, effective, patient-centred care. Staffing levels contributed to 3.5% of all reported safety incidents across the NHS[43] with workforce factors likely to contribute to a far higher total proportion. This has resulted in the dilemma of maintaining, and ideally improving, the quality and safety of care in a maternity service facing greater demand and increasing complexity in the health of childbearing women.

Research from a number of sources, including studies of other health-care professionals, points towards better-quality care, improved outcomes or fewer adverse events being associated with higher levels of registered nurse staffing.[44–49] However, while reduced complication levels may be associated with reduced length of inpatient stay (improved productivity), the association between higher registered nurse staffing and reduction in stay is not universally supported. Kane's systematic review found evidence to support such a relationship in intensive care and outcomes for surgical patients but not for medical patients, and further highlights complexity and the challenges of attributing cause.[50]

There is also a body of literature on the optimal level of staffing for doctors from the USA[51,52] and some evidence of a medical staffing outcome relationship from both the USA and the UK.[44,53] Other evidence that considers medical staffing suggests complex inter-relationships between workload, efficiency and quality.[54–58] However, this literature is more limited in extent than that on nurses, and there are also significant concerns with drawing causal inference from the extant literature to the UK maternity workforce.

The complementary and substitutability of nurses/midwives and doctors is even less well documented in large studies exploring routine practice (as opposed to experimental implementations). Outcomes may be sensitive to ratios between nurses and medical staff. For example, in the UK, a higher total of clinically qualified staffing (doctors + nurses) per bed and a higher number of doctors relative to the number of nurses were both associated with lower mortality-based failure to rescue in the fully adjusted analysis.[57]

In maternity care, a survey of health-care professionals showed that many believed that low staffing levels have a direct impact on safety of maternity services as a result of increased error rates, burnout, tiredness and less direct care.[58] Respondents were of the opinion that higher midwifery staffing levels would allow all women to have one-to-one care in labour, reduce intervention rates, reduce postnatal hospital stays and release money to reinvest in services.

However, few studies have investigated the link between obstetric and midwifery staffing and outcomes.[42] Joyce et al.[59] drew on cross-sectional data from all 65 maternity units in the Thames region between 1994 and 1996, covering a total of 540,834 live births and stillbirths. After adjustment for birthweight, perinatal units with a more 'interventionist' approach (defined by higher rates of caesareans, epidurals and instrumental births) and higher levels of consultant obstetric staff were found to be associated with lower stillbirth rates; and this effect persisted after adjustment for other possible predictive and confounding factors. An analysis using HES 2008 data matched with staffing variables from the Maternity Matters Benchmarking Dataset found a relationship between higher levels of full-time equivalent (FTE) midwifery staffing and a lower chance of readmission at 28 days; however, risk adjustment was limited.[60] However, observational studies have limited capacity to identify causal pathways.

The NHS Operating Framework 2010/11 identified the need to help local managers to identify optimum skill mix for quality and productivity.[1] Birthrate Plus (BR+) is widely used in the UK to calculate the number of midwives required in a NHS maternity unit.[61] Despite the widespread use and recommended use of BR+, it is not known whether ratios or staffing establishment numbers reflect 'the ideal' or 'what is current' and how these are related to providing a high-quality and safe maternity service.

Strategic approaches to maternity support worker development are under way at a national level in Scotland, Wales and Northern Ireland. However, there is limited and inconclusive evidence that changing workforce skill mix or substitution of roles in maternity care and other acute or primary care settings is associated with improved health outcomes or a reduction in costs. Few, if any, studies have considered the potential trade-offs between staff groups to optimise quality and efficiency, nor have they attempted to explore differential effects on different outcomes simultaneously.

Health-care workforce and efficiency

The majority of the literature on the relationships between the health-care workforce and outcomes including efficiency and effectiveness is based within acute secondary care. Very little relates specifically to maternity services, although there may be lessons to learn. Work examined so far points to a relatively simple gradient of improving outcomes with more registered nurses, and improvements in both outcomes and cost-effectiveness with richer skill mix.

Moving beyond the nursing workforce, economic evaluations of nurse for doctor substitution (which could be construed, in part, as involving a dilution of skill mix) also suggest that such substitution can be cost-effective or lead to a net cost reduction.[62] The optimal use of scarce and expensive labour resources will depend upon whether or not they are complements or substitutes. There are two common approaches to this question used in the production economics literature: p-complementarity and q-complementarity.[63,64] Traditionally, p-complementarity is evaluated from a cost function, but in health-care applications cost data are not often available for all inputs. However, q-complementarity can be investigated via the production function, but is not often addressed. A rare example of this approach to health care is by Thurston and Libby,[65] who estimated the staffing relationships for primary physician services in the USA. They found that nurses are q-complements for physicians, while technicians and unqualified nurse aides are q-substitutes for nurses, in the production of primary care visits. Economically, very little is known about the complementarity or substitutability of staff groups (skill mix) within the NHS, despite there being critical changes in the composition of the workforce over recent years. We have not found examples which address this important question from within any acute care settings.

Some of the economic models above also point to two conceptually distinct 'outcomes' for a given health-care team: quality (represented primarily by patient safety in the existing literature) and productivity (represented by volume of cases treated or length of stay). This is also embodied in the current NHS Quality, Innovation, Productivity and Prevention (QIPP) programme, which seeks to improve quality and productivity simultaneously, although it is not clear whether improvements in both are separate, linked or traded off. While there is considerable evidence on the benefits of investment in improved patient safety, very little is known about the impact on a health-care provider's efficiency and output of diverting resources to this cause.[66]

Cost-effectiveness and effective use of fixed resources involving alteration in the composition of the clinical team is clearly dependent upon wage differentials. Replication and extension of US findings in other health economies is clearly warranted. It also seems clear from the existing evidence that there is unlikely to be a general relationship between skill mix and quality/productivity that generalises across care settings. Furthermore, all the above-cited economic models are limited because the staffing variation observed in cross-sectional observational studies is assumed to be causing the differences that are observed. The effect of variation associated with nurse staffing is assumed to be accurately determined by parameters derived from regression equations, even though it is clear that neither costs nor outcomes are the result of a deterministic process.

In relation to economic evaluations of skill mix change and outcomes, this research in general is limited.[62] Jones *et al.*[57] noted that, while there are a few hospitals that have relatively low staffing levels but appear to produce good outcomes, there are hospitals with high staffing levels that appear to produce poor outcomes. This suggests that high staffing levels may be merely indicative of aspects of care, and existing economic models have simply presumed that the relationships observed are causal. However, it is unknown if reductions in staffing levels and mix would produce a corresponding reduction in outcome. Improving outcomes through staffing changes is not costless for health-care providers and standard microeconomic theory would suggest that they are subject to diminishing marginal returns. This notwithstanding, little is really known about the impact of variations of workforce and skills mix either positively or negatively in relation to health-care providers' operational efficiency or the potential to substitute one grade of staff for another (e.g. nurses for doctors, health-care support workers for nurses, clinicians for managers, let alone midwives for obstetricians) and its impact upon outputs.

Some studies have considered the costs of providing maternity care and how costs vary between hospitals. Laudicella *et al.*[67] undertook an analysis using patient level data comparing obstetric departments between hospitals. They examined the effect of patient characteristics on costs and considered factors to explain differences in costs between hospitals. Using HES record data, they mapped costs to individual patients and found that costs were driven by women's characteristics to a greater extent than was explained by the type of birth they had. Costs were higher for women who lived in an area of greater social deprivation or had a number of obstetric risk factors. Even after adjusting for maternal characteristics and the type of birth as identified by the Health Reference Group (HRG) code, they found large variations in costs of obstetric care. They proposed that these might arise from differences in coding practice, differences in how costs were apportioned within accounting systems or differences in efficiency.

Further work done at the Centre for Health Economics, York,[68] considered cost and length of stay for women having babies. They found that older women, those with more risk factors, those from poorer areas and those having more complex births with interventions had a longer stay and higher costs. These factors have been rising for several years and continue to increase.

Task-shifting offers another possible route to cost savings. There is no robust evidence about the cost-effectiveness of maternity support workers. Several studies have sought to compare the costs of midwife-led care with consultant/medically led care. The studies use a variety of methods in their costing calculations and some include elements of ante- and postnatal care in addition to the intrapartum period. This makes it difficult to draw conclusions.

A comparative analysis of normal hospital birth in nine European countries confirmed the importance of labour costs and skill mix as determinants of total delivery costs.[69] While medical tests and drugs accounted for only 1–10% of these costs for all countries, staffing accounted for as much as 74% of total costs in Germany and 63% in Spain, although the equivalent figures were only 25% in Italy, 28% in Denmark, 34% in France and 42% in England. Denmark, France and England are identified as examples of countries that primarily use midwives to provide support before, during and after birth, while Germany and Spain almost always have an obstetrician present during birth, which accounts for their additional staff costs. The researchers conclude that higher nurse-to-physician ratios reduce costs because midwives and nurses are able to take on many medical tasks that would otherwise be performed by doctors.

Five studies in a Cochrane review that compared continuity of midwife-led with shared or medical-led care in 13 trials involving 16,242 women at low and mixed risk included cost data, using different economic evaluation methods. All found savings associated with midwife-led intrapartum care. Although the studies were inconsistent in their approach to estimating maternity care costs, it seems there is potential for cost-saving with midwife-led care.[70] Based on scant existing evidence, there appears to be a trend towards a cost-saving effect for midwife-led continuity care compared with other care models.[71] The estimated mean cost saving for each eligible maternity episode is £12.38. This translates to an aggregate saving of £1.16M per year, if half of all eligible women avail themselves of midwife-led care at booking. This equates to an aggregate gain of 37.5 quality-adjusted life-years (QALYs) when expressed in terms of health gain using a NICE cost-effectiveness threshold of £30,000 per QALY. The uptake of midwife-led maternity services affects results on two levels: first by its role in determining caseload per midwife and thus mean cost per maternity episode; second at the aggregate level by determining the total number of women who start in midwife-led services nationally.[72]

Other cost drivers

Staffing is not the only driver of costs in maternity services. Other factors, such as equipment use, also play a part. In addition, factors such as the mode and place of birth have implications not just for costs but also for staffing requirements. The delivery setting has clear implications for staffing levels and skill mix. The cost-effectiveness of alternative planned places of birth was assessed with individual-level data from the Birthplace national prospective cohort study in 147 trusts in England between 2008 and 2010 involving 64,538 women at low risk of complications before the onset of labour. Incremental cost per adverse perinatal outcome avoided, adverse maternal morbidity avoided and additional normal birth were costed. The total unadjusted mean costs were £1066, £1435, £1461 and £1631 for births planned at home, in free-standing midwifery units (FMUs), in alongside midwifery units (AMUs) and in obstetric units (OUs) respectively. Much of the cost saving was attributed to lower caesarean rates in non-OU settings. For multiparous women at low risk of complications, planned birth at home was the most cost-effective option. For nulliparous low-risk women, planned birth at home is still likely to be the most cost-effective option but is associated with an increase in adverse perinatal outcomes.[73]

The staffing costs of intrapartum care delivery are difficult to identify because of the complexity of disentangling not just the intrapartum element from ante- and postnatal care, but also the staffing component from associated costs, such as birth setting, mode of delivery and length of stay. A further difficulty in interpreting the evidence is that the available data come from different national systems of maternity care, which makes direct cost comparison difficult. Given this, the evidence of the financial implications of different staffing models is limited. The available evidence, however, suggests that midwife-led models of care and out-of-hospital midwife-led settings could provide a safe and, in many cases, cost-effective alternative to medically led intrapartum care.

The evidence presented in this chapter has highlighted limited empirical evidence regarding the impact of maternity workforce staffing and skill mix on the safety, quality and cost of maternity care in the UK, variations in outcomes and women's experiences of care, despite a number of policy drivers and recommendations. The aim of this project was to understand the relationships between maternity

workforce staffing, skill mix, cost and a range of outcomes including patient safety and quality indicators, and efficiency. The research aimed to answer the following questions:

- How do organisational factors affect variability in maternal interventions and maternal and perinatal outcomes?
- What is the relationship between maternity staffing, skill mix and maternal and perinatal outcomes?
- What is the relationship between maternity staffing, cost and outcomes?

Chapter 2 Design and methods

Research questions

In order to fulfil the aims and objectives, our research question asked to what degree are maternal quality and safety outcomes explained by characteristics of NHS trusts, staffing levels and skill mix, after adjusting for mothers' characteristics including clinical risk and sociodemographic factors. We wished to know: How important are staffing levels and skill mix in determining outcomes for women and babies? What is the relationship between maternity workforce staffing levels, quality and safety outcomes and health-care output? To what extent is there an optimal staffing mix? What is the implication of efficiency savings for key quality and safety indicators?

Study design

A cross-sectional analysis was undertaken using multilevel logistic regression to investigate the relationship between a number of quality and safety outcomes and the workforce configuration adjusting for confounding characteristics. A cost analysis using a framework where both outcomes and cost are taken into account when measuring efficiency was also completed.

Patient and public involvement

In order to ensure that service user views were included in the proposed research, a trained childbirth educator, with experience of facilitating learning and discussion with pregnant women and their partners (MD), was involved in the development of the proposal and the design of the research, and named as a joint applicant on the grant proposal. She is also an experienced maternity service user representative, having represented the views of women on a number of maternity projects and committees, including Midwifery 2020, the Maternity Care Working Party and the RCOG Women's Network. She has been concerned with the experience of NHS patients generally and has served as a lay member on the NICE Guidance Development Group for Patient Experience in Adult NHS Services. Her interest in maternity research is reflected in her role as voluntary research networker for the National Childbirth Trust (NCT).

In addition, she has an interest in how maternity data can be used to inform women making decisions and choices about childbirth. Together with the data analyst on this project (RG), she represented BirthChoiceUK, a voluntary organisation which has, for the last 14 years, helped women to make choices about their maternity care through information provided by the website www.BirthChoiceUK.com.[74] It provided information and maternity statistics for each maternity unit in the UK in an accessible format for parents to help them know what questions to ask locally and to decide where to plan to have their baby.

As joint applicant, and a user researcher, MD contributed to the development of the research proposal. This led to the consideration of wider quality outcomes for maternity services rather than a focus primarily on safety, and a more woman-focused approach to the concepts of productivity and efficiency.

It was originally intended that she would only attend and contribute to co-investigator meetings. However, it became clear that her knowledge of HES data and coding, quality metrics and women's experience of maternity care – all derived from previous service user representative activities and driven from a woman-centred point of view – would be essential and valuable to the project and she therefore joined the project team. In doing so, she was able to have a hands-on role in the project, contributing on a regular basis to the direction of the study, including the choice of indicators to represent quality of

maternity care and undertaking part of the research herself, developing a way of identifying women who had clinical risk factors relevant to birth.

She has already had, and will continue to have, a role in dissemination, having given both oral and poster presentations on the project and its interim findings at a number of conferences. Findings from this project will also inform future developments of BirthChoiceUK in its aim to provide high-quality information for parents. For example, HES data, cleaned and stratified by clinical risk and parity using techniques developed in this project, are displayed on the Which? Birth Choice website (www.which.co.uk/birth-choice[75]), created in partnership with BirthChoiceUK. These data will help women understand differences in outcomes between maternity units in England and contribute to their decision about where to give birth.

Further patient and public involvement (PPI) was contributed by the members of the advisory group (see *Appendix 1*). The purpose of the advisory group was specifically to advise the co-investigator group on the study questions, analysis and outputs. The advisory group included representatives from the user organisations NCT, the Stillbirth and Neonatal Death Charity (SANDS) and the National Maternity Support Foundation (Jake's Charity), which were able to comment on the progress and findings of the study at our two advisory group meetings. Their help will be sought for dissemination of the findings.

Data sources

We used the following data sets:

- HES from 143 NHS trusts in England, admitted patients (including the 'maternity and baby tail') for the NHS year 2010–11. Specifically, the population is women who delivered in an obstetric or maternity unit based in a NHS trust in England 2010–11.
- NHS Workforce Statistics, England: 2010–11 from the Health and Social Care Information Centre, which includes staff in post including bank staff.
- CQC Maternity Survey of Maternity Provider Trusts 2007 and survey of women's experiences, 2007 and 2010, from the UK Data Archive.
- Office for National Statistics (ONS) Number of Maternities by Establishment, 2001/02 to 2010/11.
- The BirthChoiceUK database, held by Rod Gibson Associates Ltd and containing information on NHS trusts and maternity units by location and type (e.g. OU, FMU, AMU).
- Reference cost data by NHS trust – NHS reference costs 2010/11 were available under Open Government Licence v2.0 (URL: www.nationalarchives.gov.uk/doc/open-government-licence/version/2/).[76]

Data storage, governance and ethics

The data were stored in a MySQL database [MySQL version 5.5. This is an open source program (www.mysql.com) that is overseen by Oracle Corporation, Redwood Shores, CA, USA]. The flat HES data files were reorganised into a relational database to facilitate faster processing. To further speed processing, the mothers' delivery records were separated from the general inpatient records as were mothers' non-delivery inpatient records. A similar split was carried out for babies' birth records. The loading, organisation and cleaning of data as well as the calculation of indicators were all performed by routines written in Python, thus ensuring reproducibility. The number of deliveries by communal establishment code for 2010/11 was obtained from the ONS under Open Government Licence v1.0 (URL: www. nationalarchives.gov.uk/doc/open-government-licence/). These were matched with communal establishment place using a list obtained from NHS Connecting for Health. Communal establishment place was matched to NHS trust using the BirthChoiceUK database. Advice regarding ethical approval was sought from King's College Research Ethics Committee, which advised that ethics approval was not required because the team planned to undertake secondary analysis of existing anonymised data.

Careful measures were taken to safeguard the source data, in line with College research governance policies and procedures.

The HES data set available to the study consisted of:

- HES Inpatient Records 2000/01 to 2010/11
- HES Outpatient records 2003/04 to 2010/11
- HES accident and emergency (A&E) records 2007/08 to 2010/11
- ONS-HES Linked Mortality Statistics (including neonatal deaths) 2000/01 to 2010/11.

In summary, a mother has a single HES delivery record with fields available to record a limited amount of information on nine babies born in a single episode. Each baby has its own HES birth record, but this is not linked to the mother's delivery record. Only the year 2010/11 was of direct relevance to the project. However, using anonymised unique patient identifiers in the HES records, women delivering babies could be linked to previous inpatient and previous delivery records for years back to 2000/01. This allowed a more complete picture of a woman's obstetric history to be built; for example events such as stillbirth or a caesarean in a previous pregnancy which were considered risk factors for the delivery recorded in 2010/11.

Data cleaning and quality

The quality of HES data is improving. Murray *et al.*[77] examined the range and completeness of birth information recorded in HES and tested an approach for minimising the effect of hospital-level variations by selecting hospitals with high completeness of recording (90%) for key fields. They found that the proportion of missing data in key birth record fields has been decreasing annually, such as gestational age and birthweight (from 46.2% and 43.9% in 2005/06 to 18.1% and 16.9% in 2009/10 respectively). They compared the important characteristics such as size and access to specialist neonatal care between 71 high-coding and 85 low-coding hospitals and found no significant differences, suggesting hospitals with high birth record completeness may be generalisable and representative of all hospitals. Knight *et al.*[78] found that analyses using HES data were affected by the completeness and consistency of data. They found that different analysis rules had a small effect on the statistics at a national level but the effect could be substantial for individual NHS trusts.

Thus, we have drawn upon the guidelines developed by Sinha *et al.*[79] for reporting on data cleaning and quality. *Appendix 2* provides an account of full data cleaning conducted for the study, prior to data analysis. The study restricted the records to NHS hospital deliveries resulting in a registrable birth. Duplicate delivery records were removed from the mother's records. The babies' birth records also contained duplicate records. These duplicates were not removed, as the majority of the project's work concentrated on the mother's delivery record and the resources were not available to clean both. We consulted the CQC regarding its Maternity Data Quality – Indicator specifications for maternity-related measures included within its surveillance programme, and developed a scoring system that was used to select records with the largest amount of most useful and relevant data to the project.

Study size and bias

We used the full census of 656,969 women's deliveries in HES, so there was no bias caused by non-response. Any biases would therefore be caused by missing data, poorly recorded data or omitted variables from the risk adjustment model. Sample sizes for some indicators were reduced by the choice of denominator to create the indicator.

For example, healthy mother and healthy baby indicators ($n = 518,698$; 79%) were limited to records for which there was known gestational age, birth status, birthweight and postnatal duration. Normal birth

($n = 548,272$; 83%) was limited to records for which there was known onset of labour, mode of delivery and anaesthetic, and intact perineum ($n = 493,449$; 75%) was limited to women who did not have a caesarean. The final five indicators (delivery with bodily integrity, spontaneous vaginal delivery, elective caesarean, emergency caesarean and all caesareans) all had high levels of completeness ($n = 656,135$; 99.9%). Missing postcode information meant that a further 5462 delivery records (1%) were lost from the analysis, as were 68,482 records (10%) with mother's age missing.

A pre-specified decision was taken not to include any trust where fewer than 80% of women could be coded for a particular indicator. Combining this with missing postcode data meant that the percentages of delivery records used in the multilevel models were as follows: healthy mother and healthy baby indicators ($n = 431,391$; 66%), normal birth ($n = 467,022$; 71%), intact perineum ($n = 439,730$; 67%, or 89% of women who had a vaginal delivery), and delivery with bodily integrity, spontaneous vaginal delivery, elective caesarean, emergency caesarean and all caesareans ($n = 584,435$; 89%).

Derivation of process and outcome indicators

The Institute of Medicine's *Crossing the Quality Chasm* report[11] identified large gaps between health care that is, and that could and should be delivered in the United States. The report identified six aims for quality improvement that have been widely adopted internationally; thus health care should be safe, effective, patient centred, timely, efficient and equitable. Some outcomes are also influenced by factors outside health care, for example by social and structural determinants of health. Some neonatal outcomes are influenced by quality of care in the neonatal sector. *Boxes 1* and *2* detail the result of a review of quality indicators used in maternity care. Sources include AHRQ,[30] UK Policy and guidelines,[6,80] CQC,[81]

BOX 1 Process indicators

Women who have seen a midwife or a maternity health-care professional, for health and social care assessment of needs, risks and choices by 12 weeks and 6 days of pregnancy.

Mothers, who delivered their babies between 24 and 34 weeks' gestation, given any dose of antenatal steroids.

High-risk women undergoing caesarean who receive thromboprophylaxis.

Skin-to-skin contact at birth.

Women who receive blood transfusion.

Low-risk women who give birth vaginally and who receive a blood transfusion in same admission.

Percentage of women who receive an appropriate prophylactic antibiotic at the time of caesarean.

Augmentation of labour.

Induction.

Induction of labour or elective caesarean before 37 weeks.

Epidural rates.

BOX 1 Process indicators (*continued*)

Instrumental births, including ventouse.

Elective caesarean.

Emergency caesarean.

Vaginal birth after caesarean.

Admission to neonatal intensive care unit.

Percentage of term babies transferred/admitted to a neonatal intensive care nursery or special care nursery for reasons other than congenital abnormality.

Percentage of D (Rhesus) negative, unsensitised patients, regardless of age, who gave birth during a 12-month period who received anti-D immunoglobulin at 26–30 weeks' gestation.

Readmission (mother, baby).

Length of stay (mother, baby).

Admission of mother high-dependency unit/intensive care unit.

Transfer to another unit.

BOX 2 Outcome indicators

Women's experience of maternity services CQC questions.

Normal birth without interventions.

Normal birth excluding epidural.

Third- and fourth-degree tear.

Caesarean hysterectomy.

Incidence of primary postpartum haemorrhage > 1000 ml and > 1500 ml.

Pressure ulcer.

Postoperative pulmonary embolism or deep-vein thrombosis.

Percentage of obstetric patients receiving epidural/spinal analgesia who experience a post-dural puncture headache.

Foreign body left during procedure.

BOX 2 Outcome indicators (*continued*)

Maternal sepsis.

Postoperative sepsis.

Maternal mortality.

Perinatal mortality.

Respiratory distress/hypoxia/hypoxic ischaemic encephalopathy.

Birth trauma injury to neonate.

Brachial plexus injury.

Antepartum and intrapartum stillbirth.

Term.

Preterm.

Apgar score of < 7 at 5 minutes.

Sepsis in baby.

Percentage of obstetric patients receiving epidural/spinal analgesia who experience a post-dural puncture headache.

Postoperative wound dehiscence.

Gestational age baby.

Birthweight for gestational age.

CQUIN,[18,19] ONS,[82] NHS Operating Framework 2012/13,[83] RCOG,[25] Midwifery 2020,[13] *Safer Childbirth: Minimum Standards for the Organisation and Delivery of Care in Labour*,[7] Information Centre Indicators for quality improvement 2013[20] and the Australian Council on Healthcare.[84]

Indicators used in analysis

The selection of indicators was decided in consultation with the advisory group and informed by needing to have a balance of positive and negative indicators, the importance to women, costs, and the availability and quality of coding within the HES data set. Three indicators were derived to indicate a healthy mother and healthy baby, thus reflecting a concept of harm-free care, avoidance of longer-term morbidity and a positive outcome. The mode of birth indicators were chosen to compare important processes and outcomes across trusts and with other studies. Ten final indicators (some were composites) that measured maternal and infant outcomes were developed and are outlined in *Table 1*.

The subcomponents for each composite indicator are shown in *Tables 2–5*. A full description of each composite outcome is also provided; please see *Chapter 3, Outcome indicators*.

TABLE 1 Outcome indicators used in analysis: numbers achieving the indicator and percentages of eligible total

Indicator	Number	%
Healthy mother and healthy baby		
Healthy mother	143,349	27.6
Healthy baby	441,357	85.1
Healthy mother/healthy baby dyad	127,106	24.5
Mode of birth indicators		
Delivery with bodily integrity	211,303	32.2
Normal birth	220,720	40.3
Spontaneous vaginal delivery	409,579	62.4
Intact perineum	213,199	43.2
Caesareans		
Elective	65,807	10.0
Emergency	96,849	14.8
All caesareans	162,656	24.8

TABLE 2 Delivery with bodily integrity: composite indicator components

Component	Number	%
Uterine damage		
No[a]	649,089	98.9
Yes	7046	1.1
Second-/third-/fourth-degree tear		
No[a]	492,785	75.1
Yes	163,350	24.9
Sutures		
No[a]	471,070	71.8
Yes	185,065	28.2
Episiotomy		
No[a]	558,864	85.2
Yes	97,271	14.8
Caesarean		
No[a]	493,479	75.2
Yes	162,656	24.8
Total	656,135	100.0
Total: delivery with bodily integrity	211,303	32.2

a Component included in composite.

TABLE 3 Normal birth: composite indicator components

Component	Number	%
Method of delivery		
Spontaneous vertex[a]	344,342	62.8
Spontaneous other cephalic[a]	2373	0.4
Breech including partial breech extraction[a]	1058	0.3
Low forceps, not breech	22,470	4.1
Other forceps, not breech	11,973	2.2
Ventouse, vacuum extraction	33,675	6.1
Breech extraction (not otherwise specified)	142	0.0
Elective caesarean	52,985	9.7
Emergency caesarean	78,794	14.4
Other	10	0.0
Episiotomy		
No[a]	466,307	85.1
Yes	81,965	14.9
Induction		
No[a]	432,213	78.8
Yes	116,059	21.2
General anaesthetic		
No[a]	535,614	97.7
Yes	12,658	2.3
Regional anaesthetic		
No[a]	441,867	80.6
Yes	106,405	19.4
Total	548,272	100.0
Total: normal birth	220,720	40.3

a Component included in composite.

TABLE 4 Healthy mother: composite indicator components

Component	Number	%
Delivery with bodily integrity		
Yes[a]	168,606	32.5
No	350,092	67.5
Instrumental delivery		
No[a]	454,323	87.6
Yes	64,375	12.4
Maternal sepsis		
No[a]	517,040	99.7
Yes	1658	0.3
Anaesthetic complication		
No[a]	517,760	99.8
Yes	938	0.2
Mother returns home ≤ 2 days		
Yes[a]	420,917	81.1
No	97,781	18.9
Mother readmitted within 28 days		
No[a]	491,779	94.8
Yes	26,919	5.2
Total	518,698	100.0
Total: healthy mother	143,349	27.6

a Component included in composite.

TABLE 5 Healthy baby: composite indicator components

Component	Number	%
Baby's weight		
< 2.5 kg	32,772	6.3
2.5–4.5 kg[a]	477,181	92.0
> 4.5 kg	8745	1.7
Gestational age		
< 37 weeks	33,665	6.5
37–42 weeks[a]	462,181	89.1
> 42 weeks	22,852	4.4
Live baby		
Yes[a]	516,591	99.6
No	2107	0.4
Total	518,698	100.0
Total: healthy baby	441,357	85.1

a Component included in composite.

Independent variables

We included mother's characteristics measured at the individual level that are known to affect the outcomes of interest. These included age, parity, clinical risk at the end of pregnancy as measured by the NICE guideline for intrapartum care,[80] ethnicity, area socioeconomic deprivation as measured by the Index of Multiple Deprivation (IMD),[85] geographical location (urban/rural) and region.

Trust-level characteristics included size measured by number of deliveries, teaching status, maternity configuration drawing on the Birthplace in England[86] typology of whether or not AMUs and FMUs are part of trust provision and staffing variables. These included staffing levels (FTE obstetric medical staff, midwives and maternity support staff per 100 maternities, FTE all staff per 100 maternities) and skill mix (doctor/midwife and midwife/support worker ratio).

Allocating women to risk categories using data contained within Hospital Episode Statistics records

Although no pregnancy can be considered entirely risk-free, by the same token, none are entirely 'risky'. Women with recognised medical conditions and complications face greater risk of adverse outcomes and morbidities than those without. Each condition may have a spectrum of risk, and different associated adverse outcomes. Clinically, risks and plans are tailored to the individual. For the purposes of this research, some women are regarded as having 'higher-risk' pregnancies because of pre-existing medical conditions, a complicated previous obstetric history or conditions that develop during pregnancy. These women and their babies may have different outcomes from women regarded as at 'lower risk'. Trusts may have differing proportions of higher-risk women and this could affect the outcomes of maternity services. A system of allocating women retrospectively to lower- and higher-risk status at the end of pregnancy, using information contained in their HES records, allowed us to make necessary adjustments to take this into account.

Methodology

The criteria for determining a higher-risk pregnancy

The Birthplace in England[86] study compared safety of birth in different settings for women judged to be at low risk of complications at labour onset. For this study, women were classified as low risk if, immediately prior to the onset of labour, they were not known to have any of the medical conditions or situations listed in the NICE intrapartum care guideline[80] that result in increased risk for the woman or baby during or shortly after labour, where care in an OU would be expected to reduce this risk. A few additional conditions not listed in the guidelines which might also be expected to confer increased risk were also used to determine the risk status.

The NICE guidelines themselves have two categories of higher risk: those women who have factors which indicate an increased risk of complications (e.g. sickle cell disease), and those women who require individual assessment when choosing place of birth (e.g. sickle cell trait). In this study, the NICE intrapartum care guideline[80] was used to determine the risk status of women, which produced three categories of NICE risk status: low risk, requires individual assessment and at increased risk.

Matching higher-risk conditions with data items contained in Hospital Episode Statistics records

Conditions listed in both the NICE intrapartum care guidelines were matched by one of the project team (MD) to relevant four-alphanumeric digit ICD-10 codes, of which there are about 12,000. Initial matching largely used the existing grouping of ICD-10 codes by chapter and block (e.g. O30 multiple gestation, I10–I15 hypertensive diseases). This included a number of codes for which there were few or no women diagnosed with a particular condition. For certain conditions, other types of codes were matched, such as OPCS-4 or HES Data Dictionary data items, for example to identify breech presentation or multiple pregnancy.

Matching of codes was checked by another member of the project team who is an obstetrician (SB) and disagreements were resolved by consensus. In order to verify that all significant increased risk codes had been included, the numbers of women with a HES delivery record in 2010/11 with each of the four-digit ICD-10 codes were identified. Any codes remaining unallocated, with more than 10 women diagnosed with the condition, were individually checked by both MD and SB to see if they:

(a) had been missed, and should have been included as being on one of the lists categorising women as being at increased risk of complications
(b) had not specifically been included on the NICE list of conditions but nonetheless were considered to constitute a diagnosis which would place a women at increased risk of complications sufficient to recommend hospital delivery ('additional codes').

Each ICD-10 code and other relevant codes were allocated a number or letter determining the level of risk that the diagnosis would confer according to the NICE guideline, as follows in *Table 6*.

For the NICE guideline list, the condition could either be pre-existing or have arisen during the current pregnancy so that it was present by the end of pregnancy or onset of labour (i.e. in time to advise a woman on obstetric referral or place of birth). Some conditions related to events arising in a previous pregnancy or delivery. For these, previous linked records were searched to determine if a woman had a factor increasing her risk of complications. For example, a woman who had had eclampsia following a previous birth could be identified using code O152 (eclampsia in the puerperium) in a previous delivery record. Pre-existing conditions not related to pregnancy (e.g. cardiac disease) were also searched for in any previous linked inpatient record.

It was decided to exclude induction of labour from the list of conditions that put women at increased risk according to NICE, as this is a procedure rather than a diagnosis. While a woman having an induction would have to give birth in an OU, her risk status at the end of pregnancy should be determined by the diagnosis of a condition rather than the decision to induce. This also enabled the distinction between documented medical and 'social' indications for induction.

Determining the risk status of women

Each woman's HES records for the past 10 years were searched for codes that were tagged to be 4, 43, 2, 23.

If any of a woman's records contained a 4 or 43, she was retrospectively categorised as 'NICE increased risk'.

Any woman with no records containing a 4 or 43 type code, but containing a 2 or 23 code was categorised as needing 'NICE individual assessment'.

TABLE 6 Risk classification

Status	Status number	Number of ICD-10 codes allocated to nearest 50 (current record)
NICE low risk	0	1650
NICE individual assessment	2	50
NICE increased risk	4	1050
Not considered (as no women diagnosed with condition)	1	7600
Not considered (as number of women diagnosed between 1 and 10)	1	1900
Additional individual assessment code	23	50
Additional increased risk code	43	50

If none of a woman's records contained any codes labelled 4, 43, 2 or 23, she was categorised at 'NICE low risk'.

Methodological limitations

Because of the large number of ICD-10 codes, not every code was allocated to a NICE risk category. However, all codes with a count of over 10 women were allocated, as were 22% of codes with a count of under 10 (but at least one) women. Because of this, the proportion of higher-risk women may have been underestimated. However, as conditions were coded initially using chapters and blocks of ICD-10 codes, it was unlikely that unallocated codes would relate to conditions on the NICE guideline list.

There was not always a direct match for conditions, and in these circumstances a clinical judgement was made. In some cases it was possible to obtain further information using other codes. For example, a large number of women were diagnosed with 'unspecified' asthma (J459, $n = 36,000$) and it was not possible to determine which of these were mild or at severe increased risk of complications. Those with verified severe asthma were identified by looking for records of women who had previously been hospitalised with asthma as their main diagnosis according to previous inpatient records ($n = 4000$); they were classified as 'increased risk' whereas the rest were classified as being at low risk of complications in the absence of any other risk factors. Nonetheless, this was still not a perfect match, as it was not possible to determine how many women had 'required an increase in treatment' as specified by the NICE guideline.

In some cases it was not possible to determine whether the diagnosis was antenatal, intrapartum or postpartum and a judgement had to be made, always erring on the side of caution. In some circumstances other codes could be used to inform that judgement. For example, where a woman was diagnosed with anaemia, this could be antepartum, putting her at increased risk of complications during the birth, or postpartum, which would not be relevant for allocation to risk category. Women diagnosed as anaemic who had had an intrapartum or postpartum haemorrhage were excluded from the increased risk category in the absence of any other risk factors.

Incomplete or inaccurate coding also affected the ability to assign a risk status to women. Trusts which were poor at coding would tend to show more lower-risk women than trusts which were good at coding.

Overall, there were a number of limitations to this method. However, initial testing suggested that it was a useful tool which distinguished well between the groups of women. However, because of unallocated codes which may have been clinically significant on an individual basis, it was considered to be of use only in retrospective analysis rather than in a clinical context prospectively categorising women.

Use in regression analyses

For the regression analyses, women categorised as 'NICE increased risk' or 'NICE individual assessment' were subsequently combined into a single category, 'higher risk'. Women categorised here as 'NICE low risk' were subsequently referred to as 'lower risk'.

Tables 7–10 are reproduced from the NICE intrapartum care guideline[80] and show medical conditions and other factors that should be taken into consideration when planning place of birth.

TABLE 7 Medical conditions indicating increased risk suggesting planned birth at an OU

Disease area	Medical condition
Cardiovascular	Confirmed cardiac disease
	Hypertensive disorders
Respiratory	Asthma requiring an increase in treatment or hospital treatment
	Cystic fibrosis
Haematological	Haemoglobinopathies – sickle cell disease, beta-thalassaemia major
	History of thromboembolic disorders
	Immune thrombocytopenia purpura or other platelet disorder or platelet count below 100,000
	Von Willebrand's disease
	Bleeding disorder in the woman or unborn baby
	Atypical antibodies which carry a risk of haemolytic disease of the newborn
Infective	Risk factors associated with group B streptococcus whereby antibiotics in labour would be recommended
	Hepatitis B/C with abnormal liver function tests
	Carrier of/infected with HIV
	Toxoplasmosis – women receiving treatment
	Current active infection of chickenpox/rubella/genital herpes in the woman or baby
	Tuberculosis under treatment
Immune	Systemic lupus erythematosus
	Scleroderma
Endocrine	Hyperthyroidism
	Diabetes
Renal	Abnormal renal function
	Renal disease requiring supervision by a renal specialist
Neurological	Epilepsy
	Myasthenia gravis
	Previous cerebrovascular accident
Gastrointestinal	Liver disease associated with current abnormal liver function tests
Psychiatric	Psychiatric disorder requiring current inpatient care

HIV, human immunodeficiency virus.
Source: Reproduced with permission from NICE Clinical Guideline 55,[80] *Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth*, pp. 11–12. © NICE, 2007.

TABLE 8 Other factors indicating increased risk suggesting planned birth at an OU

Factor	Additional information
Previous complications	Unexplained stillbirth/neonatal death or previous death related to intrapartum difficulty
	Previous baby with neonatal encephalopathy
	Pre-eclampsia requiring preterm birth
	Placental abruption with adverse outcome
	Eclampsia
	Uterine rupture
	Primary postpartum haemorrhage requiring additional treatment or blood transfusion
	Retained placenta requiring manual removal in theatre
	Caesarean section
	Shoulder dystocia
Current pregnancy	Multiple birth
	Placenta praevia
	Pre-eclampsia or pregnancy-induced hypertension
	Preterm labour or preterm prelabour rupture of membranes
	Placental abruption
	Anaemia – haemoglobin less than 8.5 g/dl at onset of labour
	Confirmed intrauterine death
	Induction of labour
	Substance misuse
	Alcohol dependency requiring assessment or treatment
	Onset of gestational diabetes
	Malpresentation – breech or transverse lie
	Body mass index at booking of greater than 35 kg/m^2
	Recurrent antepartum haemorrhage
Fetal indications	Small for gestational age in this pregnancy (less than fifth centile or reduced growth velocity on ultrasound)
	Abnormal fetal heart rate/Doppler studies
	Ultrasound diagnosis of oligo-/polyhydramnios
Previous gynaecological history	Myomectomy
	Hysterotomy

Source: Reproduced with permission from NICE Clinical Guideline 55,[80] *Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth*, pp. 12–13. © NICE 2007.

TABLE 9 Medical conditions indicating individual assessment when planning place of birth

Disease area	Medical condition
Cardiovascular	Cardiac disease without intrapartum implications
Haematological	Atypical antibodies not putting the baby at risk of haemolytic disease
	Sickle cell trait
	Thalassaemia trait
	Anaemia – haemoglobin 8.5–10.5 g/dl at onset of labour
Infective	Hepatitis B/C with normal liver function tests
Immune	Immune non-specific connective tissue disorders
Endocrine	Unstable hypothyroidism such that a change in treatment is required
Skeletal/neurological	Spinal abnormalities
	Previous fractured pelvis
	Neurological deficits
Gastrointestinal	Liver disease without current abnormal liver function
	Crohn's disease
	Ulcerative colitis

Source: Reproduced with permission from NICE Clinical Guideline 55,[80] *Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth*, pp. 13–14. © NICE 2007.

TABLE 10 Other factors indicating individual assessment when planning place of birth

Factor	Additional information
Previous complications	Stillbirth/neonatal death with a known non-recurrent cause
	Pre-eclampsia developing at term
	Placental abruption with good outcome
	History of a previous baby more than 4.5 kg
	Extensive vaginal, cervical or third- or fourth-degree perineal trauma
	Previous term baby with jaundice requiring exchange transfusion
Current pregnancy	Antepartum bleeding of unknown origin (single episode after 24 weeks of gestation)
	Body mass index at booking of 30–34 kg/m^2
	Blood pressure of 140 mmHg systolic or 90 mmHg diastolic on two occasions
	Clinical or ultrasound suspicion of macrosomia
	Para 6 or more
	Recreational drug use
	Under current outpatient psychiatric care
	Age over 40 years at booking
Fetal indications	Fetal abnormality
Previous gynaecological history	Major gynaecological surgery
	Cone biopsy or large loop excision of transformation zone
	Fibroids

Source: Reproduced with permission from NICE Clinical Guideline 55,[80] *Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth*, pp. 14–15. © NICE 2007.

Statistical methods

Multilevel logistic regression models, where mothers (deliveries) were nested within trust, were fitted to the 10 dichotomous maternity indicators using the Statistical Analysis System (SAS) version 9.3 (SAS Institute Inc., Cary, NC, USA) Generalised Mixed Models procedure (GLIMMIX). An unstructured variance–covariance matrix parameterised via the matrix's Cholesky root was used. The root is squared to give the residual variance (σ^2). We present the residual σ in findings. This parameterisation has good computational and statistical properties (it guarantees that the variance–covariance matrix is at least positive semidefinite), and is often preferred to an unstructured variance–covariance matrix. Other simper forms of the variance–covariance matrix are available but they require certain assumptions to be made, so a decision was taken not to impose a predetermined structure. The independent variables used in the model are shown in *Table 11*.

TABLE 11 Independent variables used in statistical models

Variable	Categories/definition
Mother's characteristics	
Mother's age (years)[a]	≤ 19, 20–24, 25–29, 30–34, 35–39, 40–44, ≥ 45
Mother's parity[a]	0, 1, 2, 3, 4 or more
Clinical risk[b]	Lower, higher (includes individual assessment)
Ethnicity[a]	Not given/not known/not stated
	English/Welsh/Scottish/Northern Irish/British (white)
	Irish (white)
	Gypsy or Irish traveller
	Any other white background
	White and black Caribbean (mixed)
	White and black African (mixed)
	White and Asian (mixed)
	Any other mixed/multiple ethnic background
	Indian (Asian or Asian British)
	Pakistani (Asian or Asian British)
	Bangladeshi (Asian or Asian British)
	Chinese
	Any other Asian background
	African (black or black British)
	Caribbean (black or black British)
	Any other black/African/Caribbean background
	Arab
	Any other ethnic group, please describe
Postcode-linked data	
IMD[a]	1 = most deprived to 5 = least deprived
Rural/urban classification[a]	No information/other postcode
	Urban ≥ 10,000 – sparse
	Urban ≥ 10,000 – less sparse

TABLE 11 Independent variables used in statistical models (*continued*)

Variable	Categories/definition
	Town and fringe – sparse
	Town and fringe – less sparse
	Village – sparse
	Village – less sparse
	Hamlet and isolated dwelling – sparse
	Hamlet and isolated dwelling – less sparse
Strategic Health Authority[a]	North East
	North West
	Yorkshire and Humber
	East Midlands
	West Midlands
	East of England
	London
	South East Coast
	South Central
	South West
Trust-level data	
Trust size[c]	ONS maternities (in thousands)
University trust[d]	Attached to a university – yes/no
Configuration[e]	OU, OU/AMU, OU/FMU, OU/AMU/FMU
Staffing variables[f]	
Set 1	FTE doctors per 100 maternities
	FTE midwives per 100 maternities
	FTE support workers per 100 maternities
Set 2	FTE all staff per 100 maternities
	Doctor-to-midwife ratio
	Support worker-to-midwife ratio

a Categories and definitions for these variables were constructed with reference to the HES Data Dictionary.[87]
b Clinical risk variables were derived from the NICE Clinical Guideline 55 for intrapartum care.[80]
c Number of maternities by establishment were requested from the ONS Demographic Analysis Unit. Julie Jeffries, Demographic Analysis Unit, ONS, 2012, personal communication. Demographic Analysis Unit, ONS. *An Estimate of the Number of Maternities by Establishment for the Financial Years 2001/2002 to 2010/2011*. Unpublished; 2012. © Crown copyright 2012, provided to the team under Open Government License v2.0 (URL: www.nationalarchives.gov.uk/doc/ open-government-licence/version/2/), and then used in conjunction with the BirthChoiceUK database, used with permission from Rod Gibson Associates Ltd, London, UK, to determine trust size (Rod Gisbson, Rod Gibson Associates Ltd, 2012, personal communication).
d University trust information was derived from the Association of University Hospital Trusts.[88]
e BirthChoiceUK database used with permission from Rod Gibson Associates Ltd, London, UK, (Rod Gibson, Rod Gibson Associates Ltd, 2012, personal communication).
f Staffing variables information requested from Health and Social Care Information Centre. Health and Social Care Information Centre, 2012, personal communication. Health and Social Care Information Centre (2010–11). *Medical and Non-Medical Workforce Census. Provisional NHS Hospital and Community Health Services Monthly Workforce Statistics*. Unpublished; 2012.

Two models were fitted to the data using the mothers' characteristics, sociodemographics and trust-level variables with either set 1 or set 2 of the staffing variables. Note we use the abbreviation FTE for full-time equivalent staff. Some of the categories do not appear in the statistical models because they were treated as missing data. Often these are related to the fact that the woman's postcode was missing (e.g. for rural/urban classification, if a woman's postcode was missing then she would not have a code for IMD).

The results from the 10 multilevel models were summarised in three tables (see *Chapter 3*, *Tables 23–25*).

Specifically for this study, symbols have been used to indicate the magnitude of each model effect using the relative chi-squared (chi-squared/degrees of freedom) value in a heuristic way to help summarise the results and provide a sense of which variables were important in predicting particular outcomes. The relative chi-squared values for all statistically significant effects were categorised as follows: minor (1–99), moderate (100–999), strong (1000–9999) or dominant ($\geq 10,000$).

The ↑ has been used to indicate a positive monotonic relationship (i.e. rises upwards in a linear or non-linear way) between a continuous predictor and the outcome, ↓ to indicate a negative monotonic relationship, * for non-monotonic relationships (e.g. increases with parity and then declines from parity 2 onwards) and ↔ for relationships involving categorical variables [e.g. ethnicity, Strategic Health Authority (SHA)]. So, for positive continuous relationships the symbols ↑, ↑↑, ↑↑↑, ↑↑↑↑ were used for minor, moderate, strong and dominant effects respectively. Where a variable was close to statistical significance (i.e. $p > 0.05$ but ≤ 0.10), this has been placed in parentheses in the report table (see *Tables 23–25*). Note the relative chi-squared value for each of the variables is presented in a separate table in the results section (see *Table 22*).

Sometimes outcomes were achieved more often as deprivation increased. This was represented by an upwards arrow ↑. Conversely, when they happened less often this was indicated by a downwards arrow ↓.

Independent variables were added in blocks starting with mother-level variables, then sociodemographics, trust-level and finally staffing variables. There were two different staffing variable models: the first used FTE per 100 maternities for the three staffing groups (doctors, midwives and support staff) and the second model used FTE (all) staff per 100 maternities alongside two ratio variables – doctor-to-midwife ratio and support worker-to-midwife ratio – to test the effect of substitution of one staff group with another.

The intercept was the only parameter allowed to vary between trusts, to ensure that clustering of mothers and babies within trusts was properly accounted for in the estimation of the parameter estimate standard errors (SEs). Funnel plots of the unadjusted and adjusted proportions (*y*-axis) against the number of known deliveries for each indicator (*x*-axis) were plotted. In *Appendix 3*, the funnel plots for the intercept-only model (unadjusted proportions) and the model that used set 1 of the staffing variables (adjusted proportions) are presented, noting that the plots for the two staffing models were very similar.

We did not attempt to adjust the funnel plot limits for overdispersion,[89] which would happen if more trusts fell outside the control limits than would normally be expected. The main purpose of the plots was to judge what impact the model independent variables had upon the variation in the indicator rates across trusts rather than to attempt to identify outlier trusts, which is best achieved using adjusted funnel plots. Typically, model independent variables reduce the amount of variation between units of analysis (e.g. trusts) although it is possible for the variation to increase: for example, after the addition of the variables representing mothers' characteristics into the model the variation increased for 8 of the 10 indicators.

Three sensitivity analyses were performed by fitting the statistical models to (1) trusts which only had a single OU ($n = 50$), (2) replacing the dichotomous clinical risk grouping with a three-category variable (NICE lower risk, individual assessment, NICE higher risk) and (3) testing the following interactions (multiplicative effects):

- parity × FTE doctors per 100 maternities
- parity × FTE midwives per 100 maternities
- parity × FTE support workers per 100 maternities
- clinical risk × FTE doctors per 100 maternities
- clinical risk × FTE midwives per 100 maternities
- clinical risk × FTE support workers per 100 maternities.

This was performed in order to find out if the effect of increasing/decreasing staffing levels upon outcomes was the same across different parities and clinical risk groups.

Costing analysis

The cost analysis attempted to model maternity services to identify the relationship between inputs and outputs, and more specifically for this study the relationships between the different labour inputs. In particular, we wanted to establish the extent that the different roles are substitutes or complements, i.e. are competing or aiding inputs.

Data sources

The unit of analysis was the trust. The output variable (Q) was the number of deliveries. The staffing (input) data were taken from the National Workforce Dataset, a reference data set of standardised definitions used consistently across the NHS and governed by Information Standards. Before 2010 the staffing data were provided to the NHS Information Centre (NHS IC) by trusts using an annual census survey. From 2010 onwards the staffing data were extracted directly from the Electronic Staff Records system into a central Data Warehouse and updated quarterly. Full data quality information is available on the NHS IC website, but it employs a Workforce Validation Engine to validate data submitted by trusts and to report back errors and quality scores to trusts. The continual updating of the Electronic Staff Records data and records and the time between our census points and the data extraction date should ensure good-quality data.

Descriptive statistics for the staffing data are reported in *Tables 12* and *13*. The medical staff were separated into two groups of 'consultant' and 'other doctors', with the latter group consisting of the remaining medical staff. It was not possible to determine the split between obstetric and gynaecological work. The non-medical staff were grouped into three distinct groups: 'managers', 'midwives' and 'support staff'.

Reference cost data by NHS trust

NHS reference costs 2010/11 were available under Open Government Licence.[76] The following data were extracted for each NHS trust and primary care trust (PCT) providing maternity care:

- number of deliveries, based on activity, less excess bed-days
- activity being undertaken, given by currency code and description
- antenatal costs, given by relevant currency codes
- delivery costs, given by delivery currency codes and excluding excess bed-days
- postnatal costs, given by relevant currency codes, including excess bed-days for delivery codes
- total maternity costs – sum of above costs.

Further details of these are given in *Appendix 2*.

TABLE 12 Obstetric/gynaecological medical staffing in England 2010

Grade	FTE	Head count	Average FTE
Associate specialist	134	163	0.84
Consultant	1725	1820	10.85
Foundation year 2	296	297	1.86
House officer	100	100	0.63
Hospital practitioner and clinical assistant	13	75	0.08
Other	5	12	0.03
Registrar group	2819	2906	17.73
Senior house officer	54	56	0.34
Specialty doctor	183	219	1.15
Staff grade	54	64	0.34
Grand total	5383	5712	33.85

TABLE 13 Non-medical staffing in England 2010

Role	FTE	Average FTE
01_Nurse Consultant	64.1	0.4
02_Modern Matron	409.4	2.6
02_Nursery nurse	474.3	3.0
03_Manager	399.4	2.5
03_Nursing assistant/auxiliary	2287.1	14.4
04_Healthcare assistant	2746.1	17.3
04_Registered nurse – Children	763.0	4.8
05_Registered midwife	20,095.8	126.4
05_Support worker	960.9	6.0
10_Other 1st level	4806.7	30.2
11_Other 2nd level	115.7	0.7
Grand total	33,122.5	208.3

Methodology to explore midwifery staffing levels, outcomes and the cost of providing maternity services for NHS trusts

Data on medical staffing were not detailed enough to ascertain the split between obstetric and gynaecological responsibilities in a trust. Therefore the costing analysis investigated the relationships between midwifery staffing levels, midwifery-related outcome measures and the cost of providing maternity services for NHS trusts in England and draws on an analysis for the RCM. For each NHS trust and PCT providing maternity services, the following were obtained:

- the average maternal age at time of delivery, determined using the HES data field MATAGE
- the percentage of nulliparous women delivering in the trust, determined using the data field NUMPREG and linkage with previous delivery records
- the average ranking of the overall IMD, determined using the data field IMD04RK
- the percentage of women who were deemed to be at increased risk of complications at onset of labour, determined using the risk allocation methodology described earlier.

These were used to create a demographic profile for each NHS trust.

Modelling techniques
The analysis was done using data aggregated at the NHS trust level, rather than at the level of the women's delivery record. Although a multilevel model working down to the level of individual women's outcomes would be possible, it was the intention of this investigation to keep the analysis simple and capture the essence of any relationships using trust-level data.

All the data sources were amalgamated in an Access database (Microsoft Access Professional Plus 2010, Microsoft Corporation, Redmond, WA, USA). The statistical package R [R version 3.01 (lm and glm packages), The R Foundation for Statistical Computing, Vienna, Austria] drew the relevant data for the analysis from this database. The standard R regression routines were used to perform the analysis. In order to establish relationships between midwifery staffing, expenditure and maternity outcomes, a number of modelling techniques were explored, including transformation of variables and the use of generalised linear models. These rarely provided much advantage over ordinary linear regression analysis, so whenever possible the simpler approach was preferred.

Reference costs
Throughout this analysis all the costs were converted to costs per delivery. To account for geographical variation in labour and capital costs between trusts, the costs were divided by the Market Forces Factor.

As costs were always positive, a log transformation was considered. However, in practice it was found that the costs were reasonably normally distributed and problems did not arise with predicted negative costs. Therefore, in the interests of simplicity the regression was performed on the non-transformed data.

Midwives
The number of FTE registered midwives per delivery was used as a variable. This is the inverse of the more usually quoted midwife ratio (deliveries per midwife) but was used in preference as it was more in accord with linear regression analysis.

Operative delivery rates, normal birth rates
These outcomes were constrained to be in the range 0% to 100%, so a logistic transformation was used to ensure that the predicted values fell into this range. Standard linear regression techniques were used on the transformed values.

Care Quality Commission scores
These scores (see *Appendix 2* for more detail) were constrained to take values between 0 and 10 so a logistic transformation was used to force the constraint on the regression model.

Written complaints
These data had a Poisson distribution with many zero entries. The most appropriate model was a Poisson Generalised Linear model.

Exclusion of data for quality and bias

A number of trusts were excluded from the analysis:

- Two PCTs and two NHS trusts which provided only midwifery services. These have different patterns of care and costs from trusts providing obstetric services.
- Seven NHS trusts which had atypical or inconsistent data, or data which were overly influential in regression analyses. Inclusion of these trusts could lead to spurious conclusions about relationships between variables.

As a result, analyses were performed using data from 134 English NHS trusts and 2 PCTs.

Analysis
Initial analyses were undertaken to establish (1) the validity of the number of deliveries included in the reference cost data and (2) the relationship between trust profiles and outcomes. A number of hypotheses were then tested to determine the relationship between midwifery staffing, costs of maternity care and outcomes.

Economic modelling methodology

In economics, a production function describes the mechanism for converting a vector of inputs (x) into output (y). After selecting the appropriate functional form, econometric estimation of the function's parameters allows the output elasticities to be calculated and returns to scale to be found. In the absence of data on input prices at the maternity services level of analysis, we adopted a production (i.e. quantity) function approach. Many health-care studies using production functions (as opposed to cost functions) have adopted Reinhardt's[90] specification of the production function, which was the first to include multiple labour inputs (registered nurses, technicians, administrative staff and doctors). However, this function assumes all inputs to be substitutes (solely because of the absence of cross-products). The advance in production function analysis of the 1970s gave rise to two flexible econometric specifications that allow us to relax this overly strict assumption. Berndt and Christensen[91] introduced the transcendental-logarithmic (translog) production function and Diewert[92] introduced the generalised linear production function (also known as the Allen, McFadden and Samuelson production function).

Using either of these functions would have allowed us to estimate the relationship between the labour inputs because the regression coefficient on the cross-products (interaction effects) can be simply used to calculate the Hicks[64] elasticity of complementarity (see Sato and Koizumi[93] or Syrquin and Hollender[94] for an explanation). However, an advantage of the Diewert[92] specification is that it allows zero quantities for some inputs, which may be a more realistic assumption when labour inputs are disaggregated, as they are in our study. This modelling enabled us to examine the output contribution of the different staff inputs (output elasticities) and their influence upon the productivity of other staff inputs (i.e. whether they are complements or substitutes). With these results available, we were able to investigate the input substitution possibilities available to hospitals under different scenarios.

Following Diewert[92] we adopted a generalised linear production function defined as:

$$Y = F(X) = F(X_1, \ldots, X_K) = \sum_{i=1}^{K} \sum_{j=1}^{K} \alpha_{ij} \sqrt{X_i} \sqrt{X_j}$$

where in our study $K = 4$, $X = \{$consultants, other doctors, midwives and support staff$\}$ and $Y = Q$, corresponding to the number of deliveries. This model was estimated using ordinary least squares in Stata (Stata 12, StataCorp LP, College Station).

To examine the q-complementarity (and therefore to answer the question relating to skill mix), we calculated the Hicks elasticity of complementarity,[64] η^H defined for any two staffing inputs i, j ($i, \neq j$):

$$\eta_{ij}^{H} = \frac{ff_{ij}}{f_i f_j} \forall i \neq j$$

where

$$f_{ij} = \partial^2 f / \partial x_i \partial x_j$$

The elasticities were computed at the means and the SEs using the delta method.

Chapter 3 Findings

The over-riding aim of this research was to understand the relationships between maternity workforce size, skill mix and quality outcomes including patient safety and quality, effectiveness and unit-level efficiency. The following sections first report the unadjusted variations between trusts on the range of selected outcomes. In our adjusted models, we then investigate the relationship between maternity workforce staffing and outcomes, and to what extent there may be an optimal staffing level (objective 1). We then explore the relationship between maternity skill mix and outcomes, and to what extent there may be an optimal staffing mix (objective 2). In both the above analyses we take into account how organisational factors may affect variability in outcomes. Finally, we explore the relationships between maternity workforce size, skill mix and quality outcomes including patient safety and quality, effectiveness and cost (objective 3).

Profile of women who gave birth in 2011

Two-thirds of women who gave birth in 2011 were aged between 20 and 34 years (67%); most were nulliparous (43%) or had one previous live birth (32%) (*Table 14*). Women were more likely to be classified as higher risk than lower risk (55% vs. 45%) according to the definition of women at increased risk of complications based on the NICE intrapartum care guidelines.[80] Included in the 55% were 4% of women who required individual assessment to determine if they were at increased risk of complications. About two-thirds of women were categorised as white British, a further 9% were from another white background, 4% were Pakistani, 3% were African, 3% were Indian and 2% were from other Asian backgrounds. A higher proportion of women were living in a deprived area based on the IMD;[78] 28% versus 15% from the least deprived quintile. Most women lived in denser urban areas (86%).

TABLE 14 Women's demographic and sociodemographic profile

Individual-level variable	Number	%
Mother's age (years)		
≤ 19	32,654	5.0
20–24	112,035	17.1
25–29	162,397	24.7
30–34	165,734	25.2
35–39	93,330	14.2
40–44	21,091	3.2
≥ 45	1246	0.2
Unknown	68,482	10.4
Parity		
0	281,789	42.9
1	208,858	31.8
2	91,168	13.9
3	39,588	6.0
4 or more	35,566	5.4
Clinical risk		
Lower	297,774	45.3
Higher	359,195	54.7

TABLE 14 Women's demographic and sociodemographic profile (*continued*)

Individual-level variable	Number	%
Ethnicity (ONS definition)		
Not given/not known/not stated	29,570	4.5
English/Welsh/Scottish/Northern Irish/British	428,979	65.3
Irish	3040	0.5
Any other white background	56,785	8.6
White and black Caribbean	3105	0.5
White and black African	1997	0.3
White and Asian	1622	0.3
Any other mixed/multiple ethnic background	3562	0.5
Indian	20,398	3.1
Pakistani	26,482	4.0
Bangladeshi	8938	1.4
Chinese	4523	0.7
Any other Asian background	13,520	2.1
African	22,873	3.5
Caribbean	6784	1.0
Any other black/African/Caribbean background	6396	1.0
Any other ethnic group	18,395	2.8
IMD 2010 quintiles		
1 = most deprived	181,928	27.7
2	146,512	22.3
3	120,393	18.3
4	103,684	15.8
5 = least deprived	98,990	15.1
Unknown	5462	0.8
Rural/urban classification		
No information/other postcodes	4123	0.6
Urban ≥ 10,000 – sparse	1140	0.2
Town and fringe – sparse	2251	0.3
Village – sparse	2003	0.3
Hamlet and isolated dwelling – sparse	1150	0.2
Urban ≥ 10,000 – less sparse	561,968	85.5
Town and fringe – less sparse	43,957	6.7
Village – less sparse	28,879	4.4
Hamlet and isolated dwelling – less sparse	11,498	1.8
Total	656,969	100.0

Profile of NHS trusts in 2011

Tables 15 and *16* show that most trusts operated their maternity service either through one or more OUs only (42%) or through an OU with an AMU (OU/AMU, 32%). Nearly 30% of trusts were attached to a university and London SHA had the greatest number of trusts (24, 17%). The average number of births per trust in 2011 was 4620 (range 1214 to 10,678). There were 4.80 FTE staff for every 100 births, of which

TABLE 15 Trust profile

Trust variable	Number	%
SHA		
North East	8	5.6
North West	21	14.7
Yorkshire and Humber	14	9.8
East Midlands	8	5.6
West Midlands	15	10.5
East of England	17	11.9
London	24	16.8
South East Coast	11	7.7
South Central	10	7.0
South West	15	10.5
University trust		
No	101	70.6
Yes	42	29.4
Configuration as of September 2010		
OU	60	42.0
OU/AMU	45	31.5
OU/AMU/FMU	20	14.0
OU/FMU	18	12.6
Total	143	100.0

TABLE 16 Trust staffing variables

Trust variable	Mean	SD	Range
ONS maternities (thousands)	4.62	1.99	1.21–10.68
Staffing variables			
FTE doctors per 100 maternities	0.82	0.22	0.21–1.65
FTE midwives per 100 maternities	3.08	0.50	1.11–4.71
FTE support workers per 100 maternities	0.90	0.35	0.05–2.88
FTE all staff per 100 maternities	4.80	0.77	2.43–8.66
Doctor-to-midwife ratio	0.27	0.07	0.07–0.50
Support worker-to-midwife ratio	0.30	0.11	0.02–0.85

SD, standard deviation.

0.82 FTE were doctors (0.21 to 1.65) (one trust had a particularly low FTE, the next lowest had 0.44 FTE and all other trusts had 0.50 FTE and above), 3.08 FTE were midwives (1.11 to 4.71) and 0.90 FTE were support workers (0.05 to 2.88). The ratio of doctors to midwives averaged 0.27 (0.07 to 0.50) and support workers to midwives 0.30 (0.02 to 0.85).

Outcome indicators

The analysis that follows focuses on 10 indicators described in *Chapter 2*, *Table 1* and listed in *Table 17*. Five of these were composites of three or more other indicators. The 10 indicators have been placed into three groups. The healthy mother and healthy baby indicators form a natural group, as do the mode of birth and caesarean indicators. Healthy outcomes comprised (1) healthy mother (delivery with bodily integrity, mother returned home within 2 days, not readmitted within 28 days and without instrumental delivery, maternal sepsis or anaesthetic complication), (2) healthy baby (baby's weight 2.5–4.5 kg, gestational age 37–42 weeks, live baby) and (3) healthy mother/healthy baby dyad (cases in which both the mother and the baby were healthy). Mode of birth comprised (4) delivery with bodily integrity (without caesarean, uterine damage, second-/third-/fourth-degree tear, sutures and episiotomy), (5) normal birth (without induction, instrumental and caesarean birth, episiotomy, general and/or regional anaesthetic), (6) spontaneous vaginal delivery and (7) intact perineum. Caesarean indicators comprised (8) elective caesarean, (9) emergency caesarean and (10) all caesareans. The full reports of the models informing the results are in *Appendix 4*.

How each indicator varies across trusts is shown in *Table 17* and variation by the variables listed in *Tables 15* and *16* is shown graphically in *Appendix 3*. Intact perineum was the indicator that had the greatest variation between trusts and elective caesarean the least.

Unadjusted variation in outcomes by mothers' characteristics, sociodemographic, trust-level and staffing variables
Those mothers whose age and IMD were not known were excluded from this and subsequent analyses. In the unadjusted analysis, a large proportion of the variation observed at the level of the individual is attributable to mothers' age, parity and level of clinical risk (*Table 18*).

TABLE 17 Ten indicators: variation by trust

Indicator	Number of trusts	Mean (%)	SD	Range
Healthy mother and healthy baby				
Healthy mother	113	28.0	5.4	13.6–48.5
Healthy baby	113	85.2	3.0	76.9–90.3
Healthy mother/healthy baby dyad	113	24.9	4.9	12.0–44.7
Mode of birth				
Delivery with bodily integrity	143	32.6	5.9	19.6–52.8
Normal birth	119	40.1	4.8	26.0–51.1
Spontaneous vaginal delivery	143	62.7	4.5	46.6–73.8
Intact perineum	143	43.6	6.8	25.7–66.2
Caesarean				
Elective caesarean	143	10.0	1.7	5.6–16.6
Emergency caesarean	143	14.6	2.5	9.0–24.9
All caesareans	143	24.6	3.5	15.2–36.1

SD, standard deviation.

TABLE 18 Variation in outcomes by mother's characteristics: unadjusted analysis (%)

Variable	Healthy mother and healthy baby					Mode of birth				Caesarean			
	Mother (n = 431,391)	Baby (n = 431,391)	Dyad (n = 431,391)	DwBI (n = 584,435)	NB (n = 467,022)	SVD (n = 584,435)	IP (n = 439,730)	Elective (n = 584,435)	Emergency (n = 584,435)	All (n = 584,435)
Mother's age (years)													
≤ 19	31.4	83.9	27.4	37.7	43.6	72.7	44.3	2.4	11.3	13.7
20–24	33.2	84.9	29.3	38.6	44.7	70.1	47.4	5.2	12.5	17.7
25–29	28.7	85.5	25.6	33.4	41.7	64.8	43.2	8.1	14.1	22.1
30–34	24.4	85.4	21.7	28.8	38.1	59.4	39.8	11.5	15.6	27.1
35–39	23.5	84.6	20.7	28.0	35.1	54.7	42.4	16.5	17.0	33.5
40–44	22.5	84.2	20.0	27.5	29.4	49.3	46.3	20.4	19.8	40.1
≥ 45	20.6	78.3	17.8	25.2	20.0	37.3	54.8	30.3	23.5	53.7
Parity (grouped)													
0	14.7	83.4	12.7	18.8	29.4	53.6	25.5	4.8	20.2	25.0
1	30.5	87.2	27.4	34.8	45.9	67.3	46.7	13.7	11.3	25.0
2	41.2	86.2	36.8	47.0	48.9	70.5	62.7	15.0	9.6	24.5
3	47.4	85.0	42.2	54.5	50.6	72.3	71.6	13.7	9.9	23.6
4 or more	51.9	82.4	45.1	60.6	50.2	72.9	79.6	12.7	10.8	23.5
Clinical risk													
Lower risk	35.0	95.4	33.6	39.1	59.0	77.4	43.4	0.9	8.2	9.0
Higher risk	21.2	76.3	16.5	26.5	24.0	50.1	43.0	17.6	20.2	37.8
Ethnicity ONS													
Not known	22.0	86.5	19.8	25.0	39.0	60.4	32.8	7.2	15.8	23.0
British	29.0	84.7	25.6	33.9	39.5	63.1	45.0	10.3	13.8	24.1
Irish	20.7	84.4	18.3	25.8	31.7	54.5	37.5	14.1	16.4	30.5

TABLE 18 Variation in outcomes by mother's characteristics: unadjusted analysis (%) (*continued*)

| Variable | Healthy mother and healthy baby | | | Mode of birth | | | | Caesarean | | |
	Mother (*n* = 431,391)	Baby (*n* = 431,391)	Dyad (*n* = 431,391)	DwBI (*n* = 584,435)	NB (*n* = 467,022)	SVD (*n* = 584,435)	IP (*n* = 439,730)	Elective (*n* = 584,435)	Emergency (*n* = 584,435)	All (*n* = 584,435)
Any other white	23.2	86.9	20.8	27.2	39.0	60.4	36.4	9.6	15.0	24.6
White and black Caribbean	40.5	83.8	34.9	46.1	47.5	70.8	59.0	6.4	14.5	20.9
White and black African	29.7	86.7	27.1	35.3	44.7	63.6	48.5	11.4	15.6	27.0
White and Asian	22.7	84.5	20.4	28.5	38.6	58.3	39.4	11.3	15.6	26.9
Other mixed/multiple	27.4	85.6	24.7	33.3	40.4	62.5	45.5	11.1	14.9	26.0
Indian	15.5	84.3	13.6	18.5	34.6	55.3	26.3	10.1	19.1	29.2
Pakistani	31.0	84.2	27.4	36.6	47.5	67.6	47.2	8.2	13.9	22.2
Bangladeshi	27.5	84.2	24.2	32.9	47.3	66.6	43.6	8.2	15.9	24.1
Chinese	17.1	89.4	15.9	19.6	44.6	62.4	25.6	8.0	14.4	22.5
Other Asian	20.8	86.3	18.3	24.3	38.9	58.9	34.0	10.3	17.8	28.1
African	27.8	84.6	24.9	33.6	40.8	60.1	50.9	11.7	21.4	33.1
Caribbean	37.7	84.0	33.7	45.9	45.9	66.0	63.4	8.9	18.2	27.0
Other black, Asian or Chinese	30.2	82.7	26.2	38.2	42.5	63.3	54.4	9.4	20.1	29.5
Other ethnic group	23.4	86.5	21.0	28.7	40.4	60.1	39.6	10.1	16.7	26.8

DwBI, delivery with bodily integrity; IP, intact perineum; NB, normal birth; SVD, spontaneous vaginal delivery.

TABLE 19 Variation in outcomes by sociodemographic variables: unadjusted analysis (%)

Variable	Healthy mother and healthy baby			Mode of birth				Caesarean		All (n = 584,435)
	Mother (n = 431,391)	Baby (n = 431,391)	Dyad (n = 431,391)	DwBI (n = 584,435)	NB (n = 467,022)	SVD (n = 584,435)	IP (n = 439,730)	Elective (n = 584,435)	Emergency (n = 584,435)	
IMD 2010 quintiles										
1 = most deprived	33.3	84.0	29.3	39.2	43.8	66.5	51.3	8.5	14.4	22.9
2	28.6	84.8	25.3	33.5	40.5	63.1	44.7	9.4	15.1	24.5
3	25.6	85.4	22.8	30.1	38.3	61.2	40.5	10.1	14.9	25.0
4	23.6	85.8	21.0	27.7	37.2	59.9	37.8	11.3	14.8	26.1
5 = least deprived	20.9	86.1	18.8	24.8	36.2	58.6	34.3	12.1	14.8	26.9
Rural/urban classification										
Urban ≥ 10,000 – sparse	32.1	82.0	28.0	40.7	43.2	66.1	52.4	9.9	12.4	22.3
Town and fringe – sparse	29.9	84.3	26.9	37.5	44.7	66.1	48.9	10.2	12.6	22.8
Village – sparse	30.9	84.7	27.2	37.7	45.3	67.1	49.4	10.1	12.5	22.6
Hamlet and isolated dwelling – sparse	27.0	82.2	23.5	32.7	39.6	60.8	44.3	11.9	13.9	25.8
Urban ≥10,000 – less sparse	27.5	85.0	24.4	32.4	39.9	62.5	43.4	9.8	14.9	24.7
Town and fringe – less sparse	28.0	85.5	24.9	31.9	39.5	62.9	42.6	10.6	13.7	24.3
Village – less sparse	26.4	86.0	23.7	30.5	38.8	61.5	41.3	11.7	13.8	25.5
Hamlet and isolated dwelling – less sparse	24.6	84.8	21.8	28.8	38.2	60.0	39.4	12.0	14.2	26.2

TABLE 19 Variation in outcomes by sociodemographic variables: unadjusted analysis (%) (*continued*)

| Variable | Healthy mother and healthy baby | | | | Mode of birth | | | | | Caesarean | | |
	Mother (*n* = 431,391)	Baby (*n* = 431,391)	Dyad (*n* = 431,391)	DwBI (*n* = 584,435)	NB (*n* = 467,022)	SVD (*n* = 584,435)	IP (*n* = 439,730)		Elective (*n* = 584,435)	Emergency (*n* = 584,435)	All (*n* = 584,435)
SHA											
North East	34.2	84.0	30.4	39.2	41.5	64.6	51.1		9.1	13.2	22.4
North West	29.2	85.4	26.0	33.9	40.5	64.7	44.4		9.6	13.3	22.9
Yorkshire and Humber	32.0	84.3	28.3	37.6	43.0	66.2	48.7		8.8	13.6	22.4
East Midlands	29.6	85.2	26.6	34.4	41.8	63.9	44.5		9.2	13.4	22.6
West Midlands	29.8	83.8	25.9	34.5	40.2	63.7	46.4		9.9	15.0	24.9
East of England	27.6	85.3	24.4	32.1	40.0	62.5	43.5		10.5	15.1	25.6
London	23.1	86.2	20.5	27.4	37.5	58.5	38.6		10.7	17.3	28.1
South East Coast	23.9	85.4	21.1	28.4	37.3	60.8	38.6		11.0	14.5	25.5
South Central	23.7	84.5	21.1	29.3	39.9	61.2	39.4		9.8	15.1	24.9
South West	25.9	84.6	22.9	32.1	38.1	62.8	42.6		10.3	13.7	24.0

DwBI, delivery with bodily integrity; IP, intact perineum; NB, normal birth; SVD, spontaneous vaginal delivery.
Note: totals may not add to 100% because of rounding.

The variation was much lower for sociodemographic, trust-level and staffing variables than for mothers' characteristics (see *Tables 18–21*). Most indicators varied by mothers' age, with the prime examples being normal birth [20% (aged ≥ 45 years) to 44% (aged ≤ 19 years)], spontaneous vaginal delivery [37% (aged ≥ 45 years) to 73% (aged ≤ 19 years)] and elective caesarean [2% (aged ≤ 19 years) to 30% (aged ≥ 45 years)]. Parity had a strong influence on certain outcomes (e.g. healthy mother, model of delivery indicators) but less so on the healthy baby outcome. Most outcomes varied considerably by clinical risk; the one exception was intact perineum (43% for both lower and higher risk) but intact perineum varied more by ethnicity than any other indicator (26% to 63%).

For a number of indicators, deprived mothers had better outcomes (healthy mother, healthy mother/healthy baby dyad, delivery with bodily integrity, normal birth, spontaneous vaginal delivery, intact perineum and elective caesarean) than less deprived mothers. The two exceptions were the healthy baby indicator and emergency caesareans (see *Table 19*).

It was difficult to discern any clear pattern of variation by the rural/urban classification or by SHA, although London seemed to perform less well than other regions on a number of indicators. The amount of variation attributable to trust-level factors was generally low (*Table 20*).

University hospitals consistently performed less well than non-university hospitals although the differences tended to be quite small. Configuration appeared not to have much bearing upon outcomes. Trusts of the size 5.03 to 5.79 seemed to have highest intervention rates and lowest positive measures of birth. However, it is possible that larger trusts may have been often been split between OU sites, so that largest OUs were in the fourth quintile). Levels of staffing did not markedly impact upon the variation in outcomes but there were a few exceptions (*Table 21*).

Higher levels of doctor staffing improved a woman's chance of delivery with bodily integrity (lowest quintile 31.5%; highest quintile 33.2%) and intact perineum (lowest quintile 42.0%; highest quintile 44.5%). Higher levels of midwifery staffing improved a woman's chance of delivery with bodily integrity (lowest quintile 29.6%; highest quintile 33.7%) and intact perineum (lowest quintile 39.9%; highest quintile 44.5%) and having a birth where the mother was healthy (lowest quintile 25.5%; highest quintile 28.5%). More support staff improved a woman's chance of delivery with bodily integrity (lowest quintile 30.0%; highest quintile 33.9%) and intact perineum (lowest quintile 40.4%; highest quintile 45.3%). A doctor-to-midwife ratio of 0.22–0.25 : 1 (second quintile) generally seemed to be best.

A summary of unadjusted analyses has been presented to provide complete information, but because these data are not risk adjusted, these results should be interpreted with caution.

TABLE 20 Variation in outcomes by trust-level variables: unadjusted analysis (%)

| Variable | Healthy mother and healthy baby | | | | Mode of birth | | | | | Caesarean | | |
	Mother (*n* = 431,391)	Baby (*n* = 431,391)	Dyad (*n* = 431,391)	DwBI (*n* = 584,435)	NB (*n* = 467,022)	SVD (*n* = 584,435)	IP (*n* = 439,730)		Elective (*n* = 584,435)	Emergency (*n* = 584,435)	All (*n* = 584,435)
ONS maternities (thousands): quintiles											
1.21–2.95	30.7	85.4	27.3	34.8	40.8	63.8	46.1		9.7	13.7	23.4
2.96–3.93	29.2	86.2	26.0	33.7	40.7	63.3	45.1		10.0	14.5	24.5
3.95–5.03	27.3	84.8	24.2	32.1	40.0	63.8	42.7		9.7	14.5	24.3
5.03–5.79	24.0	85.8	21.4	28.5	38.4	60.0	39.1		10.8	15.7	26.5
5.86–10.68	28.3	83.8	24.8	33.6	40.1	62.8	44.9		9.7	14.6	24.3
University hospital											
No	28.6	85.5	25.4	33.0	40.4	63.4	43.9		9.8	14.4	24.2
Yes	25.8	84.3	22.7	30.8	38.9	60.8	42.0		10.3	15.5	25.8
Configuration as of September 2010											
OU	28.4	85.4	25.2	32.9	39.5	63.2	44.0		10.0	14.5	24.5
OU/AMU	24.9	84.7	21.9	29.6	38.6	60.2	40.5		10.4	15.8	26.1
OU/AMU/FMU	27.9	84.2	24.4	33.7	40.9	63.6	44.6		9.7	14.2	23.8
OU/FMU	30.7	86.0	27.6	35.6	43.1	65.2	46.5		9.4	13.6	22.9

DwBI, delivery with bodily integrity; IP, intact perineum; NB, normal birth; SVD, spontaneous vaginal delivery.
Note: totals may not add to 100% because of rounding.

TABLE 21 Variation in outcomes by staffing variables: unadjusted analysis (%)

| Staffing variable | Healthy mother and healthy baby | | | Mode of birth | | | | | Caesarean | | All (n = 584,435) |
| | Mother (n = 431,391) | Baby (n = 431,391) | Dyad (n = 431,391) | DwBI (n = 584,435) | NB (n = 467,022) | SVD (n = 584,435) | IP (n = 439,730) | | Elective (n = 584,435) | Emergency (n = 584,435) | |

FTE doctor per 100 maternities: quintiles

Staffing variable	Mother	Baby	Dyad	DwBI	NB	SVD	IP	Elective	Emergency	All
0.21–0.66	27.1	84.7	23.8	31.5	39.3	63.4	42.0	10.0	14.5	24.4
0.67–0.74	26.6	85.3	23.4	32.2	40.4	62.8	43.2	10.4	14.9	25.2
0.76–0.85	27.3	85.5	24.4	32.6	40.3	62.5	43.4	9.9	14.4	24.4
0.86–0.95	27.7	84.9	24.7	32.0	38.9	61.6	43.2	9.8	15.3	25.1
0.96–1.65	28.7	84.8	25.5	33.2	40.6	62.0	44.5	10.0	14.7	24.7

FTE midwives per 100 maternities: quintiles

Staffing variable	Mother	Baby	Dyad	DwBI	NB	SVD	IP	Elective	Emergency	All
1.11–2.65	25.5	85.8	22.6	29.6	39.0	61.8	39.9	9.8	15.5	25.3
2.66–2.93	26.2	84.0	22.9	30.4	39.7	61.1	41.4	10.6	15.3	25.8
2.95–3.20	28.3	84.7	25.0	33.9	40.0	63.5	45.3	10.1	14.4	24.5
3.20–3.37	29.6	85.7	26.5	34.0	40.1	62.6	45.4	9.8	14.5	24.2
3.40–4.71	28.5	85.2	25.4	33.7	40.8	63.5	44.5	9.8	14.0	23.7

FTE support staff per 100 maternities: quintiles

Staffing variable	Mother	Baby	Dyad	DwBI	NB	SVD	IP	Elective	Emergency	All
0.05–0.66	26.7	85.5	23.7	30.0	38.9	62.4	40.4	10.0	15.1	25.1
0.66–0.84	26.4	86.4	23.5	30.6	39.3	60.9	41.9	10.8	15.4	26.2
0.84–0.93	28.5	84.5	25.2	33.7	39.3	63.3	44.6	9.5	14.1	23.6
0.93–1.04	27.0	84.5	23.4	33.1	41.5	63.2	44.0	9.8	14.5	24.3
1.04–2.88	28.5	84.4	25.2	33.9	40.6	62.8	45.3	9.9	14.7	24.6

TABLE 21 Variation in outcomes by staffing variables: unadjusted analysis (%) (*continued*)

Staffing variable	Healthy mother and healthy baby			Mode of birth				Caesarean		
	Mother (n = 431,391)	Baby (n = 431,391)	Dyad (n = 431,391)	DwBI (n = 584,435)	NB (n = 467,022)	SVD (n = 584,435)	IP (n = 439,730)	Elective (n = 584,435)	Emergency (n = 584,435)	All (n = 584,435)
FTE all staff per 100 maternities: quintiles										
2.43–4.22	25.6	85.9	22.6	29.4	38.4	61.9	39.8	10.0	15.5	25.5
4.22–4.60	26.1	84.9	23.0	30.7	40.0	61.4	41.6	10.5	15.2	25.6
4.61–4.92	28.3	84.3	25.0	33.9	40.1	63.4	45.0	9.7	14.2	23.9
4.93–5.23	29.2	85.0	26.0	33.7	40.9	63.1	44.9	9.9	14.4	24.2
5.27–8.66	28.1	85.3	25.1	33.8	39.9	62.5	45.2	10.0	14.6	24.7
Doctor-to-midwife ratio: quintiles										
0.07–0.22	27.0	85.0	23.8	31.9	39.4	63.6	42.3	10.2	13.8	24.0
0.22–0.25	29.9	84.5	26.3	34.9	41.0	63.6	46.5	9.8	14.8	24.7
0.25–0.28	27.6	85.5	24.6	32.5	40.6	62.8	43.6	10.2	14.6	24.8
0.28–0.33	26.4	85.9	23.5	31.3	39.7	61.4	42.3	9.7	15.5	25.2
0.33–0.50	26.8	84.2	23.7	30.7	38.6	61.0	41.5	10.2	14.9	25.1
Support worker-to-midwife ratio: quintiles										
0.02–0.22	27.2	85.5	24.1	30.2	38.8	62.3	40.7	10.0	14.9	24.9
0.23–0.27	29.1	85.5	25.8	33.2	40.1	63.3	44.3	9.9	14.4	24.3
0.27–0.30	28.0	84.9	24.8	33.2	40.8	62.9	44.3	10.1	14.2	24.4
0.30–0.37	26.9	85.4	23.7	32.2	39.0	62.1	43.5	10.1	15.2	25.3
0.37–0.85	26.4	84.1	23.4	32.3	40.4	61.9	43.2	9.9	14.9	24.8

DwBI, delivery with bodily integrity; IP, intact perineum; NB, normal birth; SVD, spontaneous vaginal delivery.
Note: totals may not add to 100% because of rounding.

Multilevel modelling

Given the detailed nature of the tables reporting the results of the multilevel models, we have placed them in *Appendix 4* and have reproduced the results in summary form to aid the reader (*Tables 22–26*). Those mothers whose age and IMD were not known were excluded from the analysis. Note there were two different staffing models. The first model adds three variables to the model that contains mothers' characteristics, sociodemographic and trust-level variables: FTE doctors per 100 maternities, FTE midwives per 100 maternities and FTE support workers per 100 maternities. The second model replaces the three previous staffing variables with FTE all staff per 100 maternities, doctor-to-midwife ratio and support worker-to-midwife ratio.

The residual variance for the 10 intercept-only models (i.e. the model without any independent variables) ranged from 0.203 to 0.283. Based on these estimates, approximately 1–2% of the total variation in the outcome indicators is attributable to differences between trusts whereas 98–99% of the variation is attributable to differences between mothers within trusts. We are not able to say how much of the variation could be due to unknown characteristics that are predictive of outcome, or to variations in the quality of care and/or different models of care received by different women in the same trust.

The area under the curve (AUC) statistics discussed below provide some indication of how well these models fit the data. An AUC of 0.5 is no better than tossing a coin whereas an AUC of 1 implies perfect prediction.

Most of the variation observed for each individual indicator was explained by mothers' characteristics. The relative chi-squared values for the effects of mother's age, parity and clinical risk were often of a different order of magnitude from most of the other independent variables entered into the models (see *Table 22*). Age, parity and clinical risk were the only variables where the relative chi-squared value exceeded 1000 on one or more occasions (four, nine and nine occasions respectively). Parity and clinical risk were the only two variables where the relative chi-squared value exceeded 10,000 (three and nine occasions respectively). Ethnicity and the IMD were the only other two variables where the relative chi-squared value managed to exceed 100 on one or more occasions (two and four occasions respectively).

The effect sizes of the staffing variables were very small by comparison. For example, the relative chi-squared value for clinical risk was 15,841 for the healthy mother indicator whereas the largest effect size for a staffing variable for this indicator was only 3. In *Tables 22–25* effect sizes have been summarised using symbols (see key beneath each table). A monotonic relationship is one that increases consistently upwards (or downwards) in either a linear or a curvilinear fashion.

There was marginal improvement in a model's capacity to predict outcomes following the addition of sociodemographic, trust-level and staffing variables. The largest improvement was for intact perineum (AUC 0.722 to 0.732). The improvement in the AUC brought about by adding the staffing variables to the model was negligible with hardly any change in the AUC. Based on the AUC, most models meet the criteria for fair (0.70 to 0.80) to good (0.80 to 0.90) prediction. The model for elective caesarean provided the best predictions (AUC 0.814) and the model for emergency caesareans was least able to predict this outcome accurately (AUC 0.698). Potentially there is capacity to improve the fit of these models by adding further variables to raise the AUC to 0.9 and above, although a number of the variables that might help in this respect were inadequate for use in the analysis (e.g. smoking, body mass index) because they were either poorly or not consistently recorded.

The addition of variables representing mothers' characteristics into the intercept-only models (i.e. without any independent variables) more often than not resulted in an increase in the residual sigma (between trusts). The only indicators where this did not occur were normal birth (falling from 0.270 to 0.225) and spontaneous vaginal delivery (falling from 0.226 to 0.200). Conversely, sociodemographic variables had the reverse effect with a marked reduction in the variation for four of the indicators when these variables were

TABLE 22 Magnitude of model effects measured using the relative chi-squared value

Effect (df)	Healthy mother	Healthy baby	Health mother/ healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
Mothers' age group (6 df)	1083	14	850	1568	576	1746	915	713	487	1379
Parity (4 df)	10,018	615	9149	14,185	7238	7605	13,310	1785	4290	1061
Clinical risk (1 df)	15,841	26,718	23,436	17,470	63,030	49,583	945	20,858	18,388	54,882
Ethnicity (16 df)	74	6	62	124	28	20	158	21	60	32
IMD (4 df)	146	26	103	300	32	23	337	43	18	2
Rural/urban classification (7 df)	4	2	4	2	3	5	2	1	2	1
SHA (9 df)	2	2	2	2	2	1	2	1	1	2
Trust size (ONS maternities) (1 df)	4	0	3	4	2	1	5	0	0	0
University trust (1 df)	1	5	2	2	1	5	1	2	1	2
Configuration (3 df)	0	2	1	0	2	0	1	1	0	1
FTE doctors per 100 maternities (1 df)	0	1	1	1	0	0	0	2	1	3
FTE midwives per 100 maternities (1 df)	2	0	2	4	2	0	5	0	0	0
FTE support workers per 100 maternities (1 df)	3	0	2	0	0	1	0	1	0	0
FTE staff per 100 maternities (1 df)	1	1	1	6	1	0	6	0	1	1
Doctor-to-midwife ratio (1 df)	0	1	1	0	0	0	0	2	1	2
Support worker-to-midwife ratio (1 df)	2	1	2	0	0	1	0	1	0	1

df, degrees of freedom.

TABLE 23 Healthy mother and healthy baby indicators: summary of findings from the multilevel models

Level	Variable	Healthy mother (trusts = 113, deliveries = 431,391)	Healthy baby (trusts = 113, deliveries = 431,391)	Healthy mother/healthy baby dyad (trusts = 113, deliveries = 431,391)
Base model	(AUC, residual σ)	(0.500, 0.283)	(0.500, 0.227)	(0.500, 0.275)
Mother level	Mother's age (years)	↓↓↓	40–44 * 45 and over	↓↓
	Parity	↑↑↑↑	1 ** 0	↑↑↑
	Clinical risk – higher	↓↓↓↓	↓↓↓↓	↓↓↓↓
	Ethnicity	White and black Caribbean ↔ Indian	Chinese ↔ any other black, African, Caribbean background	Caribbean ↔ Indian
	(AUC, residual σ)	(0.735, 0.308)	(0.720, 0.268)	(0.747, 0.306)
Sociodemographics	IMD	↑↑	↓	↑↑
	Rural/urban	Town and fringe – less sparse ↔ Urban ≥ 10,000 – sparse	Village – less sparse ↔ Urban ≥ 10,000 – sparse	Town and fringe – less sparse ↔ Urban ≥ 10,000 – sparse
	SHA	North East ↔ London		East Midlands ↔ London
	(AUC, residual σ)	(0.741, 0.256)	(0.724, 0.250)	(0.752, 0.257)
Trust level	Size	(↓ p = 0.060)		(↓ p = 0.097)
	University trust		No ↔ Yes	
	Configuration			
	(AUC, residual σ)	(0.742, 0.242)	(0.726, 0.235)	(0.753, 0.241)
Staffing variables	FTE doctors per 100 maternities			
	FTE midwives 100 maternities			
	FTE support workers per 100 maternities			
	(AUC, residual σ)	(0.742, 0.238)	(0.726, 0.233)	(0.753, 0.236)
	FTE staff per 100 maternities			
	Doctor-to-midwife ratio			
	Support worker-to-midwife ratio			
	(AUC, residual σ)	(0.742, 0.239)	(0.726, 0.233)	(0.753, 0.236)

Monotonic positive relationship ↑ (minor), ↑↑ (moderate), ↑↑↑ (strong), ↑↑↑↑ (dominant).
Monotonic negative relationship ↓ (minor), ↓↓ (moderate), ↓↓↓ (strong), ↓↓↓↓ (dominant).
Non-monotonic relationship * (minor), ** (moderate), *** (strong), **** (dominant).
Categorical ↔ (minor), ↔↔ (moderate), ↔↔↔ (strong), ↔↔↔↔ (dominant); category most likely to achieve the outcome to the left of the two-headed arrow symbol and category least likely to the right; e.g. for the healthy mother indicator North East ↔ London means that mothers in the North East achieve the outcome most often and mothers in London the least often.

TABLE 24 Mode of birth indicators: summary of findings from the multilevel models

Level	Variable	Delivery with bodily integrity (trusts = 143, deliveries = 584,435)	Normal births (trusts = 119, deliveries = 467,022)	Spontaneous vaginal delivery (trusts = 143, deliveries = 584,435)	Intact perineum (trusts = 143, deliveries = 439,730)
Base model	(AUC, residual σ)	(0.500, 0.283)	(0.500, 0.270)	(0.500, 0.226)	(0.500, 0.286)
Mother level	Mother's age (years)	↓↓↓	↓↓	↓↓↓	≤19 **35–39
	Parity	↑↑↑	↑↑↑	↑↑↑	↑↑↑↑
	Clinical risk – higher	↓↓↓↓	↓↓↓↓	↓↓↓↓	↓↓
	Ethnicity	Caribbean ↔ Indian	Caribbean ↔ Irish	White and black Caribbean ↔ Indian	Caribbean ↔ Chinese
	(AUC, residual σ)	(0.724, 0.297)	(0.754, 0.225)	(0.722, 0.200)	(0.722, 0.334)
Sociodemographics	IMD	↑↑	↑	↑	↑↑
	Rural/urban	Village – sparse ↔ Hamlet and Isolated dwelling – sparse	Village – sparse ↔ Urban ≥ 10,000 – sparse	Village – sparse ↔ hamlet and Isolated dwelling – sparse	
	SHA	East Midlands ↔ London	(East Midlands ↔ London p = 0.071)		(North East ↔ South East Coast p = 0.064)
	(AUC, residual σ)	(0.732, 0.244)	(0.756, 0.209)	(0.724, 0.186)	(0.731, 0.280)
Trust level	Size	↓			↓
	University trust			No ↔ Yes	
	Configuration				
	(AUC, residual σ)	(0.733, 0.233)	(0.756, 0.199)	(0.724, 0.180)	(0.732, 0.267)

Level	Variable	Delivery with bodily integrity (trusts = 143, deliveries = 584,435)	Normal births (trusts = 119, deliveries = 467,022)	Spontaneous vaginal delivery (trusts = 143, deliveries = 584,435)	Intact perineum (trusts = 143, deliveries = 439,730)
Staffing variables	FTE doctors per 100 maternities				
	FTE midwives 100 maternities	↑			↑
	FTE support workers per 100 maternities				
	(AUC, residual σ)	(0.733, 0.230)	(0.756, 0.198)	(0.724, 0.179)	(0.732, 0.263)
	FTE staff per 100 maternities	↑			↑
	Doctor-to-midwife ratio				
	Support worker-to-midwife ratio				
	(AUC, residual σ)	(0.733, 0.230)	(0.756, 0.198)	(0.724, 0.179)	(0.732, 0.263)

Monotonic positive relationship ↑ (minor), ↑↑ (moderate), ↑↑↑ (strong), ↑↑↑↑ (dominant).
Monotonic negative relationship ↓ (minor), ↓↓ (moderate), ↓↓↓ (strong), ↓↓↓↓ (dominant).
Non-monotonic relationship * (minor), ** (moderate), *** (strong), **** (dominant).
Categorical ↔ (minor), ↔↔ (moderate), ↔↔↔ (strong), ↔↔↔↔ (dominant); category most likely to achieve the outcome to the left of the two-headed arrow symbol and category least likely to the right; e.g. for normal birth East Midlands ↔ London means that mothers in the East Midlands achieve the outcome most often and mothers in London the least often.

TABLE 25 Caesarean indicators: summary of findings from the multilevel models

Level	Variable	Elective caesarean (trusts = 143, deliveries = 584,435)	Emergency caesarean (trusts = 143, deliveries = 584,435)	Caesarean (trusts = 143, deliveries = 584,435)
Base model	(AUC, residual σ)	(0.500, 0.206)	(0.500, 0.203)	(0.500, 0.208)
Mother level	Mother's age (years)	↓↓	↑↑	↑↑↑
	Parity	2 *** 0	↓↓↓	↓↓↓
	Clinical risk – higher	↑↑↑↑	↑↑↑↑	↑↑↑↑
	Ethnicity	Irish ↔ white and black Caribbean	African ↔ Any other white background	African ↔ Chinese
	(AUC, residual σ)	(0.811, 0.233)	(0.692, 0.226)	(0.737, 0.215)
Sociodemographics	IMD	↓	↑	↓ ($p = 0.077$)
	Rural/urban		(Hamlet and Isolated dwelling – sparse ↔ Village – sparse $p = 0.071$)	
	SHA			London ↔ East Midlands
	(AUC, residual σ)	(0.814, 0.225)	(0.698, 0.198)	(0.740, 0.196)
Trust level	Size			
	University trust			
	Configuration			
	(AUC, residual σ)	(0.814, 0.222)	(0.698, 0.197)	(0.740, 0.194)
Staffing variables	FTE doctors per 100 maternities			
	FTE midwives 100 maternities			
	FTE support workers per 100 maternities			
	(AUC, residual σ)	(0.814, 0.219)	(0.698, 0.195)	(0.740, 0.192)
	FTE staff per 100 maternities			
	Doctor-to-midwife ratio			
	Support worker-to-midwife ratio			
	(AUC, residual σ)	(0.814, 0.219)	(0.698, 0.195)	(0.740, 0.191)

Monotonic positive relationship ↑ (minor), ↑↑ (moderate), ↑↑↑ (strong), ↑↑↑↑ (dominant).
Monotonic negative relationship ↓ (minor), ↓↓ (moderate), ↓↓↓ (strong), ↓↓↓↓ (dominant).
Non-monotonic relationship * (minor), ** (moderate), *** (strong), **** (dominant).
Categorical ↔ (minor), ↔↔ (moderate), ↔↔↔ (strong), ↔↔↔↔ (dominant); category most likely to achieve the outcome to the left of the two-headed arrow symbol and category least likely to the right; e.g. for caesarean London ↔ East Midlands means that mothers in the London have the most caesareans and mothers in the East Midlands the least.

TABLE 26 Ten indicators: model adjusted for variation between trusts

Indicator	Number of trusts	Mean (%)	SD	Range
Healthy mother and healthy baby				
Healthy mother	113	27.8	4.1	18.8–41.4
Healthy baby	113	84.9	2.9	75.6–90.7
Healthy mother/healthy baby dyad	113	24.6	3.7	16.0–38.5
Mode of birth				
Delivery with bodily integrity	143	32.6	4.3	21.8–43.0
Normal birth	119	40.0	3.9	26.0–50.1
Spontaneous vaginal delivery	143	62.6	3.6	52.8–71.5
Intact perineum	143	43.6	5.5	30.6–57.1
Caesarean				
Elective caesarean	143	10.1	1.8	5.2–15.2
Emergency caesarean	143	14.9	2.4	9.8–22.5
All caesareans	143	24.9	3.2	15.8–33.3

SD, standard deviation.
Note: Adjusted for mother's characteristics, sociodemographics, trust level factors and staffing (FTE doctors, FTE midwives, FTE support workers per 100 maternities).

added to the mother level model: healthy mother (falling from 0.308 to 0.256), healthy mother/healthy baby dyad (falling from 0.306 to 0.257), delivery with bodily integrity (falling from 0.297 to 0.244) and intact perineum (falling from 0.334 to 0.280). This suggests that when studying variability between trusts it is important to include in the model both mothers' characteristics, which increase variation between trusts, and sociodemographic variables, which decrease variation, otherwise our assessment of the variation will be biased. This provides a better indication of unexplained variation between trusts that adjusts for trust differences in women's risk profile.

Healthy mother and healthy baby

The results for the three healthy mother and healthy baby indicators (healthy mother, healthy baby, healthy mother/healthy baby dyad) are summarised in *Table 23*. The detailed results from the multilevel modelling are shown in *Appendices 3 and 5*.

Mothers' characteristics

Clinical risk was the most dominant predictor, followed by mother's parity; there were moderate effects for mother's age and IMD. The relationship with parity was linear for the healthy mother indicator, with mothers more likely to achieve a healthy outcome with increasing parity, and curvilinear for the healthy baby indicator. Babies were least likely to be born healthy if their mothers were nulliparous, and more likely to if this was their mother's second live birth (parity 1). Parity had a much more dominant effect upon the healthy mother indicator (relative $\chi^2 = 10,018$) than on the healthy baby indicator (relative $\chi^2 = 615$).

For mother and baby indicators, and the healthy mother/healthy baby dyad, clinical risk was a strong and dominant predictor with relative chi-squared values of 15,841, 26,718 and 23,436 respectively. The effect of mother's age on these three indicators was noticeably smaller, with relative chi-squared values ranging from 14 for the healthy baby indicator to 1083 for the healthy mother indicator (noting that mother's age is a confounder for parity and clinical risk). The relationship with age was linear for the healthy mother indicator, with mothers aged ≤ 19 years [odds ratio (OR) 3.744] most likely and those aged ≥ 45 years

(OR 1.00) least likely to have a healthy mother outcome. By comparison, this relationship was far weaker and flatter for the healthy baby indicator with upwards step changes at 25–29 years (OR 1.213) and 40–44 years (OR 1.306) and then a step down at ≥ 45 years (OR 1.000). Babies born to this last age group had the lowest chance of being born healthy.

Ethnicity had a stronger effect upon the healthy mother indicator (relative $\chi^2 = 74$) than the healthy baby indicator (relative $\chi^2 = 6$). Mothers of white and black Caribbean origin were most likely to experience a healthy mother outcome (OR 1.776) and mothers of Indian origin (OR 0.605) the least likely. For the healthy mother/healthy baby dyad Caribbean mothers (OR 1.657) were most likely and Indian mothers (OR 0.597) were least likely to experience a healthy outcome.

Sociodemographic factors

The chance of a healthy mother outcome was negatively associated with deprivation. Mothers belonging to the most deprived IMD quintile were more likely to achieve a healthy mother outcome than mothers belonging to the least deprived quintile [OR 1.382, 95% confidence interval (CI) 1.344 to 1.422]. This relationship was reversed, and less strong, for the healthy baby indicator (OR 0.854, 95% CI 0.826 to 0.883). Mothers from the most deprived IMD quintile were most likely to have a birth that resulted in a healthy outcome for both mother and baby (OR 1.323, 95% CI 1.285 to 1.363) noting that the proportion of elective caesareans increases as deprivation decreases (see *Table 19*: Most deprived 8.5% vs. Least deprived 12.1%). The healthy mother/healthy baby dyad was clearly weighted towards the healthy mother indicator.

There was some variation by SHA and rural/urban classification, although the size of these effects was of a much lower order of magnitude than for mothers' characteristics. Mothers giving birth in the North East were most likely to be have healthy mother outcome, while those in London were least likely (OR 1.240 vs. 0.891). For the healthy mother/healthy baby dyad the outcome was achieved most often in the East Midlands and least often in London (OR 1.253 vs. 0.907).

Trust factors

Mothers attending a university hospital trust were less likely to give birth to a healthy baby (OR 1.134, 95% CI 1.016 to 1.265). Size of trust had no impact although there was a negative effect (i.e. the more births in a trust the poorer the outcome) that approached statistical significance for the healthy mother indicator (OR 0.972, 95% CI 0.944 to 1.001; $p = 0.060$). This may be because sick mothers and babies were referred to large units, skewing their proportions of healthy mothers and healthy babies, although clinical risk has been controlled for in these models.

Staffing factors

Staffing levels were not statistically related to any of the three healthy mother and healthy baby indicators.

Model fit

The overall fit of these models based on the AUC were all in the range 0.70 to 0.80 (fair). The AUC for the healthy mother/healthy baby dyad was highest for the three indicators reported in this section (0.753 vs. 0.742 and 0.726 for the separate healthy mother and healthy baby indicators respectively). There was some reduction in the between-trust variation, after adding the independent variables to the intercept-only model, for the healthy mother indicator model (residual sigma from 0.283 to 0.238), and the mother and baby dyad model (from 0.275 to 0.236), but there was a slight increase in variation for the healthy baby indicator model (from 0.227 to 0.233).

Mode of birth

The results for the four mode of birth indicators are summarised in *Table 23*. The detailed results from the multilevel modelling are shown in *Appendices 3 and 5*.

Mothers' characteristics

Parity and clinical risk were the two dominant predictors of outcome. The effect of increasing parity increased the chance of a delivery with bodily integrity, a normal birth, a spontaneous vaginal delivery and an intact perineum. Being at increased clinical risk of complications during the birth reduced the chances of these outcomes. The effect of clinical risk was stronger than parity for normal birth and spontaneous vaginal delivery. The difference was less for delivery with bodily integrity (relative χ^2 17,470 vs. 14,185). Parity was far more important than clinical risk in determining whether or not a woman gave birth with her perineum intact (relative χ^2 13,310 vs. 945).

The relationship with age was curvilinear for all four indicators. A positive outcome was most likely for mothers aged $\leq$ 19 years [($\leq$ 19 years vs. $\geq$ 45 years) delivery with bodily integrity OR 3.638, 95% CI 3.151 to 4.201; normal birth OR 4.116, 95% CI 3.448 to 4.915; spontaneous vaginal delivery OR 5.877, 95% CI 5.175 to 6.676; intact perineum OR 1.871, 95% CI 1.545 to 2.266]. A positive outcome was least likely for mothers aged $\geq$ 45 years, although an intact perineum was least likely for mothers aged 35–39 years (OR 0.762, 95% CI 0.630 to 0.922).

The effect of ethnicity was stronger for delivery with bodily integrity and intact perineum than the other two indicators (relative χ^2 124 and 158 vs. 28 and 20). Mothers of Caribbean (black or black British) origin were most likely to deliver with bodily integrity (OR 1.799), have a normal birth (OR 1.249) and have an intact perineum (OR 2.242). In contrast, mothers of Indian and Chinese ethnicity were least likely to experience a positive outcome for delivery with bodily integrity (OR 0.594 and 0.627 respectively). Indian mothers were least likely to experience a spontaneous vaginal delivery (OR 0.830) and Chinese mothers to have an intact perineum (OR 0.563). Mothers of Irish origin were least likely to experience a normal birth (OR 0.731) and white and black Caribbean mothers (mixed) were most likely to experience a spontaneous vaginal delivery (OR 1.304).

Sociodemographic factors

Deprivation had a stronger effect upon delivery with bodily integrity and intact perineum than the other two indicators (relative χ^2 300 and 337 vs. 32 and 23). Women living in more deprived areas were more likely to deliver with bodily integrity, have a normal birth, experience a spontaneous vaginal delivery and have an intact perineum. The effects of geographical location, defined by SHA, and type and density of the area in which mothers lived were of a much lower order of magnitude than the effects of mothers' characteristics. The following outcomes were more likely to occur for mothers living in the East Midlands (delivery with bodily integrity, normal birth), North West (spontaneous vaginal delivery) and North East (intact perineum) and less likely to occur for those living in London (delivery with bodily integrity, normal birth, spontaneous vaginal delivery) and the South East Coast (intact perineum).

Trust factors

Giving birth in larger trusts with more deliveries lowered the chances of delivery with bodily integrity (OR 0.975, 95% CI 0.952 to 0.999) and an intact perineum (OR 0.971, 95% CI 0.945 to 0.998) but the effects were small (relative χ^2 4 and 5 respectively). For spontaneous vaginal delivery (OR 1.090, 95% CI 1.012 to 1.175) the outcome was better in trusts not attached to a university but again the effect was small (relative χ^2 5). Trust configuration, i.e. whether or not it had midwife-led units, appeared to have no effect upon mode of birth outcomes.

Staffing factors

A higher number of midwives (FTE per 100 maternities) was associated with improved chance of delivery with bodily integrity (OR 1.110, 95% CI 1.005 to 1.227) and an intact perineum (OR 1.132, 95% CI 1.010 to 1.268). The second staffing model suggests for both these indicators that higher levels of overall staffing increased the chances of a positive outcome (delivery with bodily integrity OR 1.079, 95% CI 1.016 to 1.147; intact perineum OR 1.092, 95% CI 1.019 to 1.170).

Model fit

The overall fits of these models based on the AUC were all in the range 0.70 to 0.80 (fair), similar to the healthy mother and healthy baby indicators. The AUC for the normal birth was highest (0.756 vs. 0.733, 0.724 and 0.732 for delivery with bodily integrity, spontaneous vaginal delivery and intact perineum respectively). There was a reduction in the between-trust variation (residual sigma) across all four indicators, particularly for normal birth, from 0.270 to 0.198 when mothers' characteristics, sociodemographics, trust-level and staffing variables were added to the intercept-only model (i.e. the model that contains no independent variables). This reduction was somewhat smaller for delivery with bodily integrity (0.283 to 0.230), spontaneous vaginal delivery (0.226 to 0.179) and intact perineum (0.283 to 0.263).

Caesareans

The results for the three caesarean indicators are summarised in *Table 25*. The detailed results from the multilevel modelling are shown in *Appendices 3 and 5*.

Mothers' characteristics

The chances of a caesarean were lowest for mothers aged ≤ 19 years [(≤ 19 years vs. ≥ 45 years) elective caesarean OR 0.130, 95% CI 0.112 to 0.152; emergency caesarean OR 0.348, 95% CI 0.302 to 0.402], rising thereafter with increasing age.

The relationship between the chance of caesarean and parity depended on whether it was elective or emergency. For elective caesareans, the probability was lowest for nulliparous mothers and highest for mothers with two previous live births. For emergency caesareans, nulliparous women were the most susceptible, with a sharp decline in risk thereafter. Women at increased clinical risk were far more likely to undergo a caesarean whether this was elective (OR 22.666, 95% CI 21.726 to 23.647) or emergency (OR 3.263, 95% CI 3.208 to 3.319). The relative chi-squared value for increased clinical risk (20,858) was of a much higher order of magnitude than for either mother's age (713) or parity (1785) for elective caesarean. The difference was smaller for emergency caesarean (parity 4290 vs. clinical risk 18,388).

Some variation between ethnic groups remains in the model having adjusted for all other independent variables. Irish women were most likely, and white and black Caribbean (mixed) women least likely, to undergo an elective caesarean (OR 1.233 vs. 0.692). African women were most likely, and women of any other white background were least likely, to have an emergency caesarean (OR 1.453 vs. 0.841).

Sociodemographic factors

Women living in the most deprived area, based on the IMD, were less likely to undergo an elective caesarean than those living in the least deprived areas [(most deprived vs. least deprived) OR 0.816, 95% CI 0.789 to 0.844] having corrected for mothers' characteristics (age, parity, ethnicity, clinical risk), other sociodemographic, trust-level and staffing variables. The relationship is reversed for emergency caesarean and the effect of IMD is less strong [(most deprived vs. least deprived) OR 1.113, 95% CI 1.081 to 1.145]. There was some variation across SHAs for all caesareans, with women living in London most likely to have a caesarean (OR 1.119) and women in East Midlands least likely (OR 0.875).

Trust-level factors

None of the trust-level variables (trust size, university trust status or configuration) was statistically associated with the chances of undergoing a caesarean.

Staffing factors

Staffing variables had a non-significant effect upon the chances of a caesarean. The correlation between ONS maternities (size) and FTE doctors per 100 maternities is –0.20, $p = 0.017$, $n = 143$.

Model fit

The overall model for elective caesareans achieved a noticeably higher AUC statistic than for emergency caesareans (0.814 vs. 0.698) and therefore was far better at predicting the outcome.

Trust variation

Variation between trusts represents a comparatively small component of the overall variation (approximately 1–2%). This variation was not substantially reduced by the addition of the independent variables, although reduction was more evident amongst the mode of delivery indicators (see *Table 26*).

The funnel plots for the 10 indicators (see *Appendix 3*) provide further confirmation of this finding. The funnel plot limits suggest that there was more variability than expected by chance, with more data points outside the control limits. One possible reason could be that important explanatory variables have been omitted from the model. The funnel plots generally confirmed what was found in the multilevel models, based on the change in the residual variance estimates, that the models did not reduce the variability between trusts to any great degree.

Sensitivity analysis

The sensitivity analysis consisted of refitting the model using a three-category clinical risk variable (lower, individual assessment, higher), fitting the model to the 50 trusts that operated their maternity service solely through a single OU, testing the interactions between parity and individual staffing variables (FTE doctors per 100 maternities, FTE midwives per 100 maternities, FTE support workers per 100 maternities) and between clinical risk and the same staffing variables. It was not possible to fit all the interactions simultaneously to the pre-existing (main effects) model because of data dependency issues, so each interaction variable (e.g. parity × FTE doctors per 100 maternities) was fitted separately to the model. The final sensitivity analysis compared the healthy baby indicator multilevel model for all mothers with the model confined to deliveries that were not preterm (< 37 weeks) and were not antepartum stillbirths. Note that the model for the healthy mother indicator, confined to the same subset, was similar to the model for all mothers and for that reason is not reported here.

A consistent picture emerged when the parameter estimates from the pre-existing model (clinical risk dichotomised into lower and higher risk) were compared with the model using the three-category clinical risk variable. Those women whose risk was based on an individual assessment nearly always had a better outcome than those at higher risk (see *Appendix 5*). The one exception was for intact perineum; women in this group had poorer outcomes than women in the higher-risk group (OR 0.782 vs. 0.812).

The single OU analysis (see *Appendix 5*) seemed to throw up differences in the relationships between trust-level variables and the indicators from the all trusts model. There was a tendency for London SHA to improve its position. For example, London rose from last place to joint fifth for the healthy mother outcome and from ninth to fourth for intact perineum. For the healthy baby indicator there was statistically significant variation between SHAs that was less evident previously ($p = 0.079$ vs. $p < 0.001$). The effect of trust size upon the healthy mother outcome strengthened, and the healthy mother/healthy baby dyad outcome was now less likely to occur in bigger trusts. This finding was replicated for delivery with bodily integrity, normal birth, spontaneous vaginal delivery, intact perineum, emergency caesarean and all caesareans. Attending a university trust was now advantageous, compared with previously, in terms of normal birth outcome ($p = 0.050$). Non-university trusts were no longer better for the healthy baby

indicator or for spontaneous vaginal delivery. More support workers reduced the chance of a healthy baby outcome ($p = 0.048$). The effect of FTE midwives upon delivery with bodily integrity was no longer significant despite the β-coefficient increasing in size from 0.105 to 0.113. This was also the case for intact perineum (β 0.124 to 0.147).

A decision was taken to ascertain whether or not the effect of staffing levels upon outcomes could vary according to either a woman's parity or her clinical risk by fitting interaction terms to the main model. These tests are shown in full in *Appendix 6* and are summarised in *Tables 27–29*. Five tests involving the midwifery staffing variable were not possible because starting values for the model covariance matrix could not be obtained. *Tables 27–29* show the OR at each level of the categorical variable (parity, clinical risk), the combined OR in brackets and the probability values associated with the chi-squared interaction test.

For example, in *Table 27* the odds of a healthy mother outcome associated with staffing levels of doctors varies by parity. The effect of increasing the number of doctors is stronger for nulliparous women than it is for women of higher parity. The overall effect of FTE doctors for nulliparous women is calculated by multiplying 0.95 (shared across all parities) and 1.20 to give an OR of 1.14. For women of parity 4 and above, the OR is 0.95.

Therefore, increasing FTE doctors improves outcomes for nulliparous women but has the reverse effect for women of parity 3 or more, although it should be noted that the ORs for woman falling into these two higher parity groups are close to 1 (parity 3 OR = 0.96, parity 4 or more OR = 0.95). The results in *Table 27* suggest that higher staffing levels for doctors will result in improved outcomes for nulliparous women. The chances of the healthy mother outcome being met are reduced when the number of support workers is increased, irrespective of parity (ORs range from 0.87 to 0.93). Support worker staffing levels are not associated with a healthy baby outcome (ORs range from 0.91 to 1.00). The chance of a healthy mother/healthy baby dyad outcome mirrors the finding for the healthy mother outcome (ORs range from 0.87 to 0.90).

Midwives have a more positive bearing upon the outcomes of women at lower risk for all three indicators in *Table 27*, whereas this observation applies to only the healthy mother indicator for doctors.

The results are presented in the same manner for mode of birth indicators (*Table 28*). It was not possible to fit the FTE midwives × parity and FTE midwives × clinical risk interactions to the normal birth model because obtaining initial estimates for the model covariance matrix was not possible.

Higher staffing levels of doctors decreased the chances of a normal birth in the lower parities and increased the chances of a spontaneous vaginal delivery in the higher parities. Women giving birth with an intact perineum were associated with higher staffing levels of midwives and support workers, especially in the higher parities. The effect of higher support worker levels on delivery with bodily integrity was advantageous for lower-risk women (OR 1.04) and not advantageous for higher-risk women (OR = 0.96). This broad trend that favoured lower-risk women was also apparent for normal birth (OR 1.06 vs. 0.96) and intact perineum (OR 1.04 vs. 0.99). Conversely, more doctors decreased the chances of a spontaneous vaginal delivery in lower-risk women (OR = 0.94) and increased the chances in higher-risk women (OR = 1.06).

Higher staffing levels of doctors were associated with bigger reductions in the proportion of elective caesareans in the higher parities ($p = 0.001$) (see *Table 29*). The combined ORs for FTE doctors ranged from 1.00 for parity 0 women to 0.75 for women of parity 2. The trend for midwives was similar but less strong statistically ($p = 0.003$). The effect of more doctors upon the chances of an emergency caesarean was felt most by women of parity 2 (OR 1.03) and least by women of parity 4 or more (OR 0.84). A higher number of doctors therefore reduces the odds of an emergency caesarean amongst women in the lowest (OR 0.87) and highest parities (OR 0.84). For all caesareans, a higher number of doctors is associated with fewer caesareans in the highest parities.

TABLE 27 Healthy mother and healthy baby indicators: effect of staffing levels by parity and clinical risk (from the multilevel model)

Risk factor	Parameter	Healthy mother			Healthy baby			Healthy mother/healthy baby dyad		
		Doctors	Midwives[a]	Support workers	Doctors	Midwives[a]	Support workers	Doctors	Midwives	Support workers
Parity	FTE per 100 maternities	0.95		0.87	1.07		0.87	0.99		0.87
	Parity × FTE per 100 maternities									
	0	1.20 (1.14)		1.02 (0.88)	1.17 (1.25)		0.98 (0.97)	1.22 (1.20)		1.04 (0.90)
	1	1.10 (1.04)		1.03 (0.89)	1.02 (1.09)		1.01 (1.00)	1.10 (1.08)		1.04 (0.90)
	2	1.03 (0.98)		1.07 (0.93)	1.09 (1.17)		0.92 (0.91)	1.05 (1.04)		1.04 (0.90)
	3	1.01 (0.96)		1.04 (0.90)	0.84 (0.90)		0.91 (0.90)	0.98 (0.96)		1.04 (0.90)
	4 or more	1.00 (0.95)		1.00 (0.87)	1.00 (1.07)		1.00 (0.99)	1.00 (0.99)		1.00 (0.87)
	p-value (test of interaction)	0.041		0.500	0.005		0.124	0.014		0.938
Clinical risk	FTE per 100 maternities	1.14	1.12	0.92	1.16	1.09	1.04	1.13	1.12	0.93
	Clinical risk × FTE per 100 maternities									
	Lower	1.00 (1.14)	1.00 (1.12)	1.00 (0.92)	1.00 (1.16)	1.00 (1.09)	1.00 (1.04)	1.00 (1.13)	1.00 (1.12)	1.00 (0.93)
	Higher	0.88 (1.00)	0.95 (1.06)	0.93 (0.85)	1.00 (1.15)	0.93 (1.02)	0.92 (0.95)	0.95 (1.08)	0.95 (1.06)	0.92 (0.86)
	p-value (test of interaction)	<0.001	<0.001	<0.001	0.977	0.009	0.018	0.248	0.007	<0.001

ORs, combined ORs (in brackets) and p-values for each interaction test shown in the table.
a Not possible to fit parity interaction model for midwives.

TABLE 28 Mode of birth indicators: effects of staffing levels by parity and clinical risk (from the multilevel model)

Risk factor	Parameter	Delivery with bodily integrity			Normal birth			Spontaneous vaginal delivery			Intact perineum		
		Doctors	Midwives	Support workers	Doctors	Midwives[a]	Support workers	Doctors	Midwives	Support workers	Doctors	Midwives	Support workers
Parity	FTE per 100 maternities	1.07	1.08	1.00	1.30		1.05	1.27	1.02	0.97	0.99	1.25	1.15
	Parity × FTE per 100 maternities												
	0	1.06 (1.13)	1.04 (1.02)	1.01 (1.01)	0.68 (0.88)		0.96 (1.01)	0.77 (0.98)	1.01 (1.02)	0.98 (0.94)	1.10 (1.09)	0.90 (1.12)	0.87 (1.00)
	1	0.98 (1.04)	1.01 (1.09)	0.98 (0.98)	0.69 (0.90)		0.97 (1.02)	0.80 (1.02)	1.01 (1.02)	1.00 (0.96)	1.03 (1.02)	0.89 (1.11)	0.87 (0.99)
	2	0.99 (1.06)	1.04 (1.12)	1.02 (1.02)	0.73 (0.95)		0.93 (0.98)	0.86 (1.09)	1.02 (1.03)	1.02 (0.98)	1.05 (1.04)	0.95 (1.18)	0.93 (1.07)
	3	1.02 (1.09)	1.00 (1.08)	1.03 (1.04)	0.80 (1.04)		0.95 (1.00)	0.90 (1.14)	1.01 (1.03)	0.97 (0.94)	1.09 (1.08)	0.92 (1.15)	1.01 (1.16)
	4 or more	1.00 (1.07)	1.00 (1.08)	1.00 (1.00)	1.00 (1.30)		1.00 (1.05)	1.00 (1.27)	1.00 (1.02)	1.00 (0.97)	1.00 (0.99)	1.00 (1.25)	1.00 (1.15)
	p-value (test of interaction)	0.014	0.331	0.362	< 0.001		0.626	0.002	0.977	0.611	0.492	0.007	< 0.001
Clinical risk	FTE per 100 maternities	1.10	1.12	1.04	0.88		1.06	0.94	1.03	0.98	1.08	1.14	1.04
	Clinical risk × FTE per 100 maternities												
	Lower	1.00 (1.10)	1.00 (1.12)	1.00 (1.04)	1.00 (0.88)		1.00 (1.06)	1.00 (0.94)	1.00 (1.03)	1.00 (0.98)	1.00 (1.08)	1.00 (1.14)	1.00 (1.04)
	Higher	1.01 (1.10)	0.98 (1.10)	0.92 (0.96)	1.13 (0.99)		0.91 (0.96)	1.13 (1.06)	1.00 (1.03)	0.96 (0.95)	0.99 (1.06)	1.00 (1.13)	0.95 (0.99)
	p-value (test of interaction)	0.877	0.146	< 0.001	0.001		< 0.001	< 0.001	0.982	0.044	0.675	0.771	0.017

ORs, combined ORs (in brackets) and p-values for each interaction test shown in the table.
a Not possible to fit parity or clinical risk interaction model for midwives.

TABLE 29 Mode of birth indicators: effects of staffing levels by parity and clinical risk (from the multilevel model)

Risk factor	Parameter	Elective caesarean			Emergency caesarean			Caesarean		
		Doctors	Midwives	Support workers	Doctors	Midwives	Support workers	Doctors	Midwives	Support workers
Parity	FTE per 100 maternities	0.78	0.99	0.95	0.84	1.06	1.13	0.73	1.04	1.05
	Parity × FTE per 100 maternities									
	0	1.29 (1.00)	1.11 (1.10)	1.14 (1.08)	1.03 (0.87)	0.92 (0.97)	0.87 (0.99)	1.18 (0.86)	0.96 (1.00)	0.97 (1.02)
	1	1.16 (0.90)	1.04 (1.03)	1.17 (1.11)	1.12 (0.94)	0.92 (0.97)	0.85 (0.96)	1.23 (0.90)	0.97 (1.01)	1.00 (1.05)
	2	0.96 (0.75)	0.99 (0.98)	1.10 (1.05)	1.22 (1.03)	0.95 (1.01)	0.89 (1.01)	1.12 (0.82)	0.96 (1.00)	0.99 (1.04)
	3	0.98 (0.76)	0.99 (0.98)	1.11 (1.05)	1.12 (0.95)	0.92 (0.97)	0.92 (1.04)	1.08 (0.79)	0.95 (0.99)	1.01 (1.06)
	4 or more	1.00 (0.78)	1.00 (0.99)	1.00 (0.95)	1.00 (0.84)	1.00 (1.06)	1.00 (1.13)	1.00 (0.73)	1.00 (1.04)	1.00 (1.05)
	p-value (test of interaction)	0.001	0.003	0.088	0.104	0.280	0.105	0.060	0.636	0.632
Clinical risk	FTE per 100 maternities	1.43	1.01	0.98	0.93	0.95	0.95	0.96	0.96	0.98
	Clinical risk × FTE per 100 maternities									
	Lower	1.00 (1.43)	1.00 (1.01)	1.00 (0.98)	1.00 (0.93)	1.00 (0.95)	1.00 (0.95)	1.00 (0.96)	1.00 (0.96)	1.00 (0.98)
	Higher	0.58 (0.84)	1.03 (1.04)	1.10 (1.08)	0.96 (0.89)	1.03 (0.98)	1.06 (1.01)	0.86 (0.83)	1.04 (1.00)	1.08 (1.06)
	p-value (test of interaction)	<0.001	0.546	0.132	0.375	0.072	0.033	<0.001	0.027	0.001

ORs, combined ORs (in brackets) and p-values for each interaction test shown in the table.

A higher number of doctors is associated with a reduction in the number of elective caesareans in the clinically higher-risk group (OR 0.84), but is associated with an increase in the number for lower-risk women (OR 1.43). This statistically significant interaction ($p < 0.001$) is not replicated for emergency caesareans ($p = 0.38$). The effect of doctor staffing levels upon the chances of an emergency caesarean is similar in both the higher- and lower-risk groups (OR 0.89 vs. 0.93). All caesareans incorporates the characteristics of both elective and emergency caesareans so a higher number of doctors reduces the chances of any caesarean but more so in the clinically higher-risk group [(higher vs. lower) OR 0.83 vs. 0.96].

Clinical risk is a less dominant predictor of the healthy baby outcome when preterm births (< 37 weeks) and antepartum stillbirths were excluded from the analysis sample (*Table 30*). The number of deliveries contributing to the analysis dropped from 431,391 to 403,052, a fall of 6.6%. The results from the sensitivity analysis show the model parameters and global statistical tests for model effects for all deliveries and deliveries that were not preterm births or antepartum stillbirths. While remaining the dominant factor, the effect of clinical risk, as measured by the chi-squared value, fell from 26,718 to 9606, and the trust-level random variance increased from 0.233 to 0.324 for the model that included mother-level, sociodemographic, trust-level and staffing variables. The model excluding preterm births and antepartum stillbirths is less able to predict the healthy baby outcome, with the area under the curve falling from 0.726 to 0.675. There is a stronger effect of midwife staffing when preterm births and antepartum stillbirths are excluded (OR 1.172, 95% CI 0.991 to 1.387; $p = 0.063$) than found for all deliveries (OR 1.029, 95% CI 0.912 to 1.161; $p = 0.65$).

TABLE 30 Healthy baby multilevel model with and without preterm births and antepartum stillbirths

	All mothers		Excluding preterm and antepartum stillbirths	
	β	SE(β)	β	SE(β)
Intercept	2.817	0.272	3.033	0.379
Mother's age group (years)				
≤19	0.092	0.086	−0.117	0.127
20–24	0.116	0.085	−0.149	0.126
25–29	0.193	0.084	−0.103	0.125
30–34	0.208	0.084	−0.098	0.125
35–39	0.212	0.085	−0.079	0.126
40–44	0.267	0.087	0.115	0.129
≥45	0.000		0.000	
Parity				
0	−0.317	0.020	−0.503	0.028
1	0.185	0.020	0.013	0.028
2	0.154	0.022	0.014	0.030
3	0.080	0.025	0.012	0.035
4 or more	0.000		0.000	
Clinical risk				
Lower	0.000		0.000	
Higher	−1.986	0.012	−1.276	0.013

TABLE 30 Healthy baby multilevel model with and without preterm births and antepartum stillbirths (*continued*)

	All mothers		Excluding preterm and antepartum stillbirths	
	β	SE(β)	β	SE(β)
Ethnicity				
Not given/not known/not stated	−0.058	0.038	−0.122	0.047
English/Welsh/Scottish/Northern Irish/British	−0.110	0.030	−0.124	0.039
Irish	−0.166	0.072	−0.255	0.089
Any other white background	−0.016	0.033	−0.028	0.042
White and black Caribbean	−0.165	0.071	−0.063	0.093
White and black African	0.080	0.089	0.001	0.112
White and Asian	−0.151	0.095	0.060	0.128
Any other mixed/multiple ethnic background	−0.046	0.069	−0.038	0.089
Indian	−0.136	0.039	−0.101	0.050
Pakistani	−0.118	0.038	−0.133	0.048
Bangladeshi	−0.161	0.048	−0.178	0.063
Chinese	0.173	0.069	0.286	0.090
Any other Asian background	−0.028	0.044	0.024	0.057
African	−0.017	0.038	−0.081	0.049
Caribbean	−0.138	0.052	0.036	0.072
Any other black/African/Caribbean background	−0.180	0.053	−0.213	0.069
Any other ethnic group, please describe	0.000		0.000	
IMD				
1 = most deprived	−0.158	0.017	−0.129	0.022
2	−0.096	0.017	−0.080	0.021
3	−0.059	0.017	−0.058	0.021
4	−0.032	0.017	−0.022	0.021
5 = least deprived	0.000		0.000	
Rural/urban classification				
No information/other postcode				
Urban ≥10,000 – sparse	−0.227	0.115	−0.206	0.140
Town and fringe – sparse	0.049	0.085	0.046	0.103
Village – sparse	0.078	0.090	−0.005	0.106
Hamlet and isolated dwelling – sparse	−0.121	0.112	−0.192	0.131
Urban ≥10,000 – less sparse	0.060	0.035	0.059	0.043
Town and fringe – less sparse	0.048	0.038	0.065	0.047
Village – less sparse	0.099	0.040	0.093	0.050
Hamlet and isolated dwelling – less sparse	0.000		0.000	

TABLE 30 Healthy baby multilevel model with and without preterm births and antepartum stillbirths (*continued*)

	All mothers		Excluding preterm and antepartum stillbirths	
	β	SE(β)	β	SE(β)
SHA				
North East	−0.089	0.118	−0.114	0.163
North West	0.234	0.092	0.297	0.127
Yorkshire and Humber	0.103	0.097	0.081	0.134
East Midlands	0.209	0.114	0.246	0.157
West Midlands	0.144	0.108	0.160	0.149
East of England	0.169	0.108	0.162	0.149
London	0.263	0.100	0.376	0.139
South East Coast	0.173	0.120	0.216	0.166
South Central	0.098	0.114	0.189	0.157
South West	0.000		0.000	
Trust size				
ONS maternities (thousands)	−0.002	0.015	−0.004	0.020
University trust				
Yes	0.000		0.000	
No	0.125	0.056	0.124	0.078
Configuration				
OU	−0.040	0.071	−0.006	0.098
OU/AMU	−0.126	0.081	−0.118	0.112
OU/AMU/FMU	−0.167	0.097	−0.235	0.134
OU/FMU	0.000		0.000	
Staffing variables				
FTE doctors per 100 maternities	0.149	0.133	0.030	0.184
FTE midwives per 100 maternities	0.028	0.062	0.159	0.086
FTE support workers per 100 maternities	−0.034	0.071	−0.121	0.098
Random variation (trust level)	σ	SE(σ)	σ	SE(σ)
Intercept only	0.227	0.011	0.350	0.016
Mother level	0.268	0.019	0.370	0.026
Mother level, sociodemographics	0.250	0.017	0.347	0.024
Mother level, sociodemographics, trust level	0.235	0.016	0.331	0.023
Mother level, sociodemographics, trust level, staff 1	0.233	0.016	0.324	0.023

TABLE 30 Healthy baby multilevel model with and without preterm births and antepartum stillbirths (*continued*)

	All mothers		Excluding preterm and antepartum stillbirths	
	β	SE(β)	β	SE(β)
Global tests (df)	χ^2	**p-value**	χ^2	*Pr > χ^2*
Mother's age group (6 df)	85.881	< 0.0001	58.348	< 0.0001
Parity (4 df)	2461.647	< 0.0001	1819.462	< 0.0001
Clinical risk (1 df)	26,717.623	< 0.0001	9606.228	< 0.0001
Ethnicity (16 df)	89.968	< 0.0001	76.682	< 0.0001
IMD (4 df)	103.297	< 0.0001	44.019	< 0.0001
Rural/urban classification (7 df)	16.186	0.0235	11.712	0.1105
SHA (9 df)	15.450	0.0793	15.532	0.0773
ONS maternities (1 df)	0.019	0.8890	0.042	0.8383
University trust (1 df)	5.005	0.0253	2.566	0.1092
Configuration (3 df)	4.507	0.2116	4.793	0.1876
FTE doctors per 100 maternities (1 df)	1.246	0.2642	0.027	0.8702
FTE midwives per 100 maternities (1 df)	0.211	0.6456	3.451	0.0632
FTE support workers per 100 maternities (1 df)	0.231	0.6308	1.533	0.2156
Area under the curve	**AUC**	**SE(AUC)**	**AUC**	**SE(AUC)**
Intercept only	0.500	0.000	0.500	0.000
Mother level	0.720	0.001	0.666	0.001
Mother level, sociodemographics	0.724	0.001	0.672	0.001
Mother level, sociodemographics, trust level	0.726	0.001	0.675	0.001
Mother level, sociodemographics, trust level, staff 1	0.726	0.001	0.675	0.001

β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.
Pr > χ^2 refers to the probability of exceeding the χ^2-statistic.

Staffing, outcome and costs

The final objective was to examine the relationship between maternity workforce staffing levels, quality and safety outcomes and cost. Data on medical staffing were not detailed enough to ascertain the split between obstetric and gynaecological responsibilities in a trust. Therefore, this analysis investigated the relationships between midwifery staffing levels, where data quality was very good, midwifery-related outcome measures and the cost of providing maternity services for NHS trusts in England. This section draws on analyses conducted for the RCM.[95]

For this analysis we used trust-level data to investigate relationships between outcome measures, midwifery staffing levels and the cost of providing maternity services for NHS trusts in England.

Initial analyses

Number of deliveries

Maternities for each site derived from ONS data matched well with the number of deliveries calculated using reference cost data. Two NHS trusts had discrepancies, with >70% more deliveries according to ONS data than recorded in the activity from reference costs. These were excluded from further analyses. For further analyses, the number of deliveries derived from reference costs was used. In the following text the expression $y \sim x \oplus z$ is used for the schematic relationships of 'y depends on x and z' being tested at each stage, although the functional relationship may be more complicated if transformations or link functions have been used.

Trust profiles and average costs

The average and range of variables used for the analyses are shown in *Table 31*.

Hypothesis 1: investing in providing good antenatal services results in reduced operative delivery rates and lower delivery costs

Operative delivery rate ~ antenatal cost per delivery (⊕ risk ⊕ parity ⊕ age ⊕ IMD) When the case-mix adjustments were included in the model there was no relationship with antenatal spend. Including the reference costs category 'attendance' (either face to face or not face to face, which is largely thought to be antenatal contact time) in the definition of antenatal expenditure reduced the dependence further.

Delivery cost per delivery ~ antenatal cost per delivery ⊕ midwives per delivery ⊕ trust size ⊕ (age ⊕ % nulliparous ⊕ % risk ⊕ IMD) An initial analysis showed that delivery costs have a significant dependence on the antenatal spend, with increased antenatal costs associated with a decrease in delivery costs. However, taken at face value, the extra amount incurred on antenatal care exceeded the reduction in delivery costs.

Figure 1 shows that there are around a dozen trusts with very high antenatal costs and correspondingly low delivery costs which are responsible for the regression relationship. When the analysis was repeated excluding trusts with antenatal costs in excess of £1500 there was no relationship between antenatal cost and delivery cost per delivery.

TABLE 31 Variables used in cost analysis

Variables	Average of trusts	Lowest trust	Highest trust	SD
Average mother's age (years)	29	27	33	1
Nulliparous (%)	44	26	61	8
Increased risk (%)	47	28	69	7
IMD ranking	14,422	4657	25,453	4783
Operative delivery rate (%)	37	27	53	4
Normal birth rate (%)	41	26	59	5
Intact perineum rate (%)	43	26	66	7
Total cost per delivery (£)[a]	4128	2606	6429	648
Proportion antenatal cost (%)	27	5	54	10
Delivery cost per delivery (£)[a]	1974	1108	3479	462
Deliveries per midwife[a]	34	21	53	5

a Based on number of deliveries derived from reference cost data.
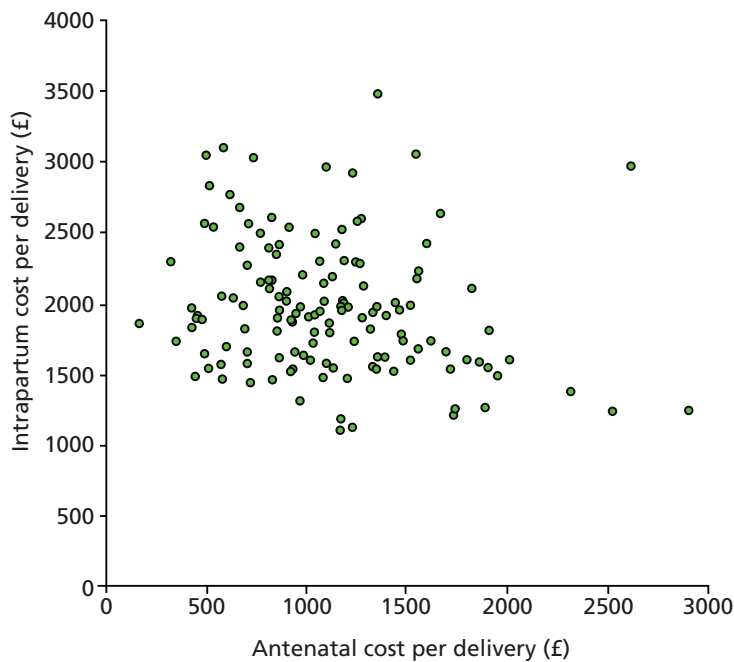
FIGURE 1 Relationship between antenatal cost/delivery and delivery cost/delivery (adjusted for Market Forces Factor).

Hypothesis 2: trusts with a high antenatal expenditure as a proportion of all maternity expenditure will have lower operative delivery rates

Operative delivery rates ∼ antenatal/total costs (⊕ risk ⊕ parity ⊕ age ⊕ IMD) ⊕ trust size
No relationship was found between the proportion of maternity expenditure spent on antenatal care and operative delivery rates after adjusting for maternal characteristics and trust size.

Hypothesis 3: there is a relationship between delivery cost per delivery and midwifery staffing levels

Delivery cost per delivery ∼ midwives per delivery ⊕ trust size (⊕ risk ⊕ parity ⊕ age ⊕ IMD)
Higher midwifery staffing levels were associated with higher costs of each delivery, although the relationship was not strong. This relationship strengthened when antenatal expenditure was included as an explanatory variable in the model.

Only around 17% of the variation between trusts' delivery costs could be accounted for by variables included in this model (midwives per delivery, trust size and case mix), rising to 23% when antenatal expenditure was taken into account (hypothesis 2b). The remaining variation in the average cost of a delivery was not accounted for by maternal characteristics, size of trust, number of FTE registered midwives employed or antenatal spend and must be due to other factors not included in the analysis.

Hypothesis 4: outcomes measures are correlated with the delivery cost per delivery

Delivery cost per delivery ∼ operative delivery rate (⊕ risk ⊕ parity ⊕ age ⊕ IMD) The analysis found that higher operative delivery rates were not significantly associated with higher delivery costs after adjusting for maternal characteristics. Variations in costs between trusts were not related to the numbers of women having operative deliveries.

Delivery cost per delivery ∼ normal birth rate (⊕ risk ⊕ parity ⊕ age ⊕ IMD) There was no association between delivery cost per delivery and the normal birth rate.

Delivery cost per delivery ~ intact perineum rate (⊕ risk ⊕ parity ⊕ age ⊕ IMD) There was no association between delivery cost per delivery and the intact perineum rate. As the denominator for intact perineum was vaginal deliveries, the analysis was repeated used delivery cost per vaginal delivery but this also showed no association.

Delivery cost per delivery ~ healthy mother and healthy baby outcomes (⊕ risk ⊕ parity ⊕ age ⊕ IMD) There was no association between delivery cost per delivery and any of the three healthy mother and healthy baby indicators (healthy mother, healthy baby and healthy mother/healthy baby dyad).

CQC scores ~ delivery cost per delivery No relationship could be established between delivery cost per delivery and women's experience of maternity care as measured by the average of the CQC scores.

CQC scores ~ antenatal cost per delivery No relationship could be established between antenatal cost per delivery and women's experiences of antenatal care as measured by the antenatal component of the CQC scores.

Hypothesis 5: trusts with high operative delivery rates have a high postnatal cost as a proportion of all maternity expenditure

Postnatal costs per delivery/total costs per delivery ~ operative delivery rates (⊕ risk ⊕ parity ⊕ age ⊕ IMD) A relationship could not be found that explained postnatal costs in terms of variations in operative delivery rates once adjustments were made.

Hypothesis 6: a trust's total expenditure on maternity services is related to the characteristics of the women who use the trust

Total costs per delivery ~ trust size ⊕ risk ⊕ parity ⊕ age Having a higher proportion of women at increased risk of complications was associated with more expensive maternity care. Level of social deprivation approached significance ($p = 0.052$) with deliveries in trusts with a higher proportion of women living in deprived areas having more costly deliveries. However, in this analysis trust size and case mix accounted for only 22% of the variation in total delivery costs.

Some of the previous analyses also showed relationships between costs and maternal characteristics. For trusts with a larger proportion of women at increased risk of complications, delivery costs were higher and they spent a higher proportion of their total maternity costs on postnatal care, and trusts with a higher proportion of nulliparous women had higher costs per delivery.

Hypothesis 7: the efficiency of a trust measured in terms of the expenditure per delivery is not correlated with the size of the trust measured by number of deliveries

A number of previously described analyses included trust size as an explanatory variable. Larger trusts, measured by the numbers of deliveries, appeared to offer maternity services which cost less than those in smaller trusts. Repeating the analysis for the 50 trusts whose maternity services are delivered by a single OU, the significant effect of size increased, suggesting that the economies of scale were greater in a single site.

The size of the trust had no relationship to the cost of a delivery only (i.e. when costs of antenatal or postnatal care were excluded). Deliveries were not cheaper in larger units, once adjustments have been made for maternal characteristics and numbers of midwives. This lack of relationship held when repeating the analysis for the 50 trusts that delivered maternity services through only a single OU.

Outcome measures and trust size

Larger trusts were associated with lower CQC scores of women's experience of maternity care. When broken down into components, women's experience of antenatal care and women's experience of labour and birth was poorer in larger trusts. No significant relationships were found between trust size and staff during labour and birth, care in hospital after the birth or feeding the baby during the first few days, although there was a trend close to significance for larger trusts to have lower scores. Overall, the analysis of CQC was very sensitive to individual data points, with significant results being lost once outliers were removed. There was no relationship between trust size and the other outcome measures: normal birth, intact perinea and number of complaints.

Economic modelling

Descriptive statistics

For the economic analysis, the unit of analysis was the hospital trust and therefore the patient-level data were aggregated to trust level. *Table 32* presents the descriptive statistics for the aggregated data. Two measures of maternity output are presented. The first measure of output is the total number of deliveries within the trust, which has a sample mean (SD) of 4600 (1991) deliveries. The second is the total number of deliveries weighted by relative cost, which combines vaginal and caesarean deliveries based upon the HRG tariff assigned to these modes of delivery. The purpose being to take account of the fact that caesarean deliveries require greater resources and have higher values (prices) attached to them. In comparison with total deliveries, the sample mean (SD) for the cost-weighted deliveries is 5740 (2491). The variation in both measures of output is relatively large.

As described in the methods and data section of the report, the staffing variables are grouped into four categories: midwives, support staff, consultants and all other doctors. Note that for completeness the econometric analysis was repeated with consultants and other doctors combined into one variable but the results were statistically indistinguishable. However, there is a strong theoretical reason to consider these two groups separately, especially in relation to their complementarity or substitutability with other staffing groups. By far the largest staff group is midwives (mean FTE 135), followed by support workers (mean FTE 42). There are roughly half as many consultants as other categories of doctors (mean FTE 11 compared with 24). There is very little variation in the average number of consultants (SD 0.60 FTE), which is likely to be the result of regulations regarding consultant cover. There is relatively large variation in the number of midwives (mean FTE 135, SD 6.5). It is worth repeating that the consultant numbers in particular, although it applies to the other workforce variables as well, are at best a proxy for the delivery suite staffing levels because staff will inevitably spend time delivering gynaecological services that could not be separated from time spent in obstetric services.

TABLE 32 Descriptive statistics for the aggregated data

Variable	Mean	SD	Minimum	Maximum
Deliveries	4599.82	1991.18	1207.00	10,703.00
Cost-weighted deliveries	5739.75	2490.58	1428.00	13,300.00
Midwives (FTE)	135.11	6.45	42.63	360.53
Support staff (FTE)	42.43	3.55	2.19	139.37
Consultants (FTE)	11.39	0.60	4.00	31.28
Other doctors (FTE)	24.03	1.46	2.12	73.27
Mean maternal age (years)	29.47	1.18	27.40	32.90
Mean parity	1.02	0.30	0.60	1.90
% high risk (NICE)	50.35	6.36	33.00	70.00

A number of control variables were included to adjust for any differences in case mix faced by trusts. The choice of variables was informed by the primary multilevel statistical analysis reported above, and these variables were aggregated to generate trust-level averages. The included variables were maternal age (mean 29 years), parity (mean 1), and the proportion of mothers classified as high risk using the NICE criteria[80] (mean 50%). In all cases, the variation across the sample was low, with SDs of 1, 0.30, and 6.36 respectively. This supports the findings from the multilevel analysis that reported a high degree of intrahospital variation at patient-level characteristics but relatively little interhospital variation.

Production function results

Table 33 reports the parameter estimates of an ordinary least squares regression of the Diewert production function[92] as specified in the methods and data section of the report. *Table 34* presents the calculated marginal productivities (estimated at the means of the variables) and the Hicks elasticities of complementarity,[64] again following the methods outlined earlier.

In *Table 33* the two alternative output measures are reported side by side in columns (1) and (2). The results are similar across the two output measures, with differences that are to be expected. The regression coefficients themselves are unhelpful in understanding the relationship between staffing levels, skill mix and the level of output produced by maternity units. Instead, the marginal productivities and the elasticities of complementarity are of greater interest and these are reported in *Table 34*. The marginal productivity – the number of additional deliveries that would be expected on average to occur if a staffing group was increased by one additional person per year – are calculated from the regression coefficients.

TABLE 33 Parameter estimates for generalised linear production function

Independent variables	(1) total deliveries		(2) total cost weighted	
	Coefficient	Standard errors	Coefficient	Standard errors
Midwives	−21.83	27.02	36.38	4.60
Support	−1.32	26.19	−2.50	33.53
Consultant	66.17	305.09	190.53	390.59
Other doctors	10.03	73.98	−27.98	94.71
Midwives$^{1/2}$ × Support$^{1/2}$	−10.07	48.01	16.43	61.47
Midwives$^{1/2}$ × Consultant$^{1/2}$	49.04	129.92	22.64	166.34
Midwives$^{1/2}$ × Other doctors$^{1/2}$	167.27[a]	80.04	233.26[a]	102.47
Support$^{1/2}$ × Consultant$^{1/2}$	153.22	187.23	121.93	239.70
Support$^{1/2}$ × Other doctors$^{1/2}$	−56.90	89.33	−88.50	114.36
Consultants$^{1/2}$ × Other doctors$^{1/2}$	−367.75	214.86	−397.86	275.08
Constant	305.06[b]		387.48[b]	203.73
Observations	141		141	
Adjusted R^2	0.885		0.878	

a $p < 0.10$.
b $p < 0.05$.
Controlling for average maternal age, average parity, the proportion of women classified as high risk, the proportion of women receiving a higher reimbursement and the proportion of multibirth deliveries.

TABLE 34 Estimates of Hicks elasticities of complementarity

Output measure – all deliveries		Midwives	Support staff	Consultants	Other doctors
Marginal productivity: F_i		17.93	10.47	32.31	42.81
Hicks elasticities: F_{ij}	Support staff	−0.03	0.00	0.00	
	Consultants	0.30	1.63	0.00	
	Other doctors	0.70	−0.42	−5.26	
Output measure – cost-weighted deliveries		**Midwives**	**Support staff**	**Consultants**	**Other doctors**
Marginal productivity: F_i		93.85	50.15	58.72	51.01
Hicks elasticities: F_{ij}	Support staff	0.05	0.00	0.00	
	Consultants	0.14	1.30	0.00	
	Other doctors	0.97	−0.65	−5.69	

F_i, marginal productivity.

Despite a relatively high adjusted R-squared (>88% in both models), very few parameters are statistically significant in either model at any of the commonly used levels of significance. This is probably because of the relatively low degrees of freedom in this aggregated data set (141 observations to fit 15 parameters) and the relatively high correlation among the explanatory variables which results from a production function such as the Diewert[92] that includes multiple cross-products (interaction effects). Therefore the model suffers from multicollinearity, which is further evidenced by the table of variance inflation factors (VIFs) (*Table 35*). In *Table 35*, all of the variables except maternal age and proportion of mothers classified as high risk using the NICE criteria[80] have a VIF in excess of 10, with many of the cross-products having VIFs in the thousands.

TABLE 35 Variance inflation factors (ordered by magnitude)

Independent variables	VIF	1/VIF
Midwives$^{1/2}$ × Consultant$^{1/2}$	1465.170	0.001
Support$^{1/2}$ × Consultant$^{1/2}$	1248.060	0.001
Midwives$^{1/2}$ × Other doctors$^{1/2}$	1212.360	0.001
Midwives$^{1/2}$ × Support$^{1/2}$	962.520	0.001
Consultant	860.310	0.001
Consultants$^{1/2}$ × Other doctors$^{1/2}$	849.530	0.001
Midwives	785.250	0.001
Support$^{1/2}$ × Other doctors$^{1/2}$	564.670	0.002
Other doctors	256.520	0.004
Support	125.520	0.008
Maternal age	1.380	0.723
Higher payment (%)	1.370	0.729
Multiple deliveries (%)	1.320	0.757
Nulliparous (%)	1.200	0.832
High risk (%)	1.170	0.856

Multicollinearity is a data problem rather than an econometric problem and the optimal solution is to gather more data to reduce the relative size of the SEs. We have an exhaustive sample and, therefore, obtaining further data is impossible in this case. Given the theoretical importance of the variables, omitting, combining or transforming the variables is not a satisfactory solution either. Ridge regression was attempted but the results differed little from the original ordinary least squares estimates, which we therefore continue to report here.

The marginal products are reported in *Table 34*, which estimates the change in output that would occur from a marginal change in each labour input holding the remaining labour variables constant. The top panel (A) in *Table 34* reports these for the production function that uses all deliveries as the output measure, and the bottom panel (B) reports the marginal products for the model that uses a cost-weighted output measure. The results are qualitatively similar, so we shall confine ourselves to a discussion of the total deliveries marginal products. The marginal productivities are all positive, indicating that increasing any staff group would increase the production of deliveries that a provider could handle. This means (in non-economic terms) the number of deliveries an institutional provider could handle or accommodate. For example, adding an additional FTE midwife would allow a hospital to produce an additional 18 deliveries on average per annum. The marginal productivities for all types of delivery are highest for consultants (32 additional deliveries) and other doctors (43 additional deliveries). When the deliveries are weighted by their HRG tariff or cost then midwives become the most productive, followed by consultants. It is important to remember that the cost-weighted results tell us nothing about the actual number of deliveries and should not be relied upon for workforce planning or policy decisions.

Table 34 also reports the estimates of the elasticities of complementarity between the different staffing groups in the production of 'deliveries'. Again, as the conclusions that are drawn from the two alternative output measures are the same, we shall limit ourselves to the discussion of the top panel (A) for total deliveries. To reiterate, a positive elasticity indicates that the two staffing groups are complements (need to be used together) and a negative elasticity indicates that the two staffing groups are substitutes (can be used in place of one another). Of the six possible combinations of staffing groups (the cross-products in the regression model), three of them are complements and three are substitutes. Therefore it is clear that a flexible production function such as ours which allows inputs to be complements is superior to more rigid, but simpler, specifications which force all inputs to be substitutes.

Midwives are quantity-complements with consultants and other doctors in the total number of deliveries produced by an organisation, as indicated by the positive Hick's elasticities of complementarity in the top panel of *Table 34*. However, they are quantity-substitutes with support workers. Besides being complements with midwives, consultants are also complements with support workers but are substitutes with other doctors. This makes intuitive sense given their scope of work overlap. It is important to remember that these are quantity complements or substitutes. However, substituting support workers for midwives may also have an impact on some aspects of the quality of care depending on the groups of women involved or the care setting.

Chapter 4　Discussion

The aim of this study was to understand the relationships between maternity workforce staffing, skill mix, organisational factors and a range of outcomes including patient safety and quality indicators and efficiency. Therefore this research aims to answer the following questions:

- How do organisational factors affect variability in maternal interventions and maternal and perinatal outcomes?
- What is the relationship between maternity staffing, skill mix and maternal and perinatal outcomes?
- What is the relationship between maternity staffing, cost and outcomes?

We conducted secondary analysis of data from HES from 143 NHS trusts in England for the NHS year 2010–11 and NHS Workforce Statistics, England: 2010–11. We included mother's characteristics, measured at the individual level, that are known to affect the outcomes of interest. These included age, parity, clinical risk at the end of pregnancy as measured by the NICE guideline for intrapartum care, ethnicity, area socioeconomic deprivation as measured by the IMD, geographical location (urban/rural) and region.

Characteristics measured at trust level included size measured by number of deliveries, teaching status, maternity configuration (whether AMU and FMU are part of trust provision or not), and staffing variables. These included staffing levels (FTE obstetric medical staff, midwives and maternity support staff/100 maternities, FTE all staff/100 maternities) and skill mix (doctor/midwife and midwife/support worker ratio).

We looked at 10 process and outcome indicators derived from HES data informed by needing to have a balance of positive and negative indicators, the importance to women, costs, and the availability and quality of coding within the HES data set. Three indicators were derived to indicate a healthy mother and healthy baby, thus reflecting a concept of harm-free care, avoidance of longer-term morbidity and a positive outcome. The mode of birth indicators were chosen to compare important processes and outcomes across trusts and with other studies. These were:

- mode of birth: comprising (1) delivery with bodily integrity (without caesarean, uterine damage, second-/third-/fourth-degree tear, sutures and episiotomy), (2) normal birth (without induction, instrumental and caesarean birth, episiotomy, general and/or regional anaesthetic), (3) spontaneous vaginal delivery and (4) intact perineum
- caesarean indicators: (5) elective caesarean, (6) emergency caesarean, (7) all caesareans
- healthy mother indicator: (8) delivery with bodily integrity, mother returned home within 2 days, not readmitted within 28 days and without instrumental delivery, maternal sepsis or anaesthetic complication
- healthy baby indicator: (9) baby's weight 2.5–4.5 kg, gestational age 37–42 weeks, live baby
- healthy mother/healthy baby dyad: (10) healthy mother and healthy baby.

Two-thirds of women were aged between 20 and 34 years (67%); most either were nulliparous (43%) or had one previous live birth (32%). Women were more likely to be classified as higher risk than lower risk (55% vs. 45%) according to the definition of women at increased risk of complications based on the NICE intrapartum care guidelines.[80] Included in the 55%, were 4% of women who required individual assessment to determine if they were at increased risk of complications. About two-thirds of women were categorised as white British, a further 9% were from another white background, 4% were Pakistani, 3% were African, 3% were Indian and 2% were from other Asian backgrounds. A higher proportion of women were living in a deprived area based on the IMD: 28% compared with 15% from the least deprived quintile. Most women lived in denser urban areas (86%).

Multilevel logistic regression models, in which mothers were nested within trust, were fitted to the 10 maternity indicators. When considering the effect of maternal characteristics, women's outcomes were largely determined by their clinical risk (based on NICE guidance[80]) and parity, with higher-risk women and nulliparous women generally being more likely to have interventions and less favourable outcomes. Age was the next most important factor, increasing interventions, with ethnicity and deprivation also being significant but having variable impact and direction. The effect of mothers' age varied by indicator. For nine indicators the best outcome was achieved in the age group ≤ 19 years and the poorest outcome in mothers aged ≥ 45 years for all indicators except intact perineum, where the relationship was U-shaped. The proportion of elective caesareans increased with age and this trend was similar but less strong for emergency caesareans.

Higher clinical risk was associated with fewer mothers achieving various healthy outcomes. The effect size (relative chi-squared value) for clinical risk was highest for normal birth and all caesareans and weakest for intact perineum. Analysis using the three-category clinical risk variable found that women in the individual assessment group nearly always had better outcomes than women in the higher-risk group based on NICE guidelines; the one exception was intact perineum.

There were improved outcomes with increasing parity for six indicators (healthy mother, healthy mother/healthy baby dyad, delivery with bodily integrity, normal birth, spontaneous vaginal delivery and intact perineum). For a number of these indicators there were marked improvements in outcomes between nulliparous mothers and mothers who already had one child. Healthy baby outcome and all caesareans were less affected by parity than the other indicators. There was a sharp increase in the proportion of elective caesareans between nulliparous mothers and mothers of parity 1 and the reverse for emergency caesareans.

Ethnicity and deprivation had their strongest associations with delivery with bodily integrity and intact perineum. Caribbean (black and black British), and white and black Caribbean mothers (mixed race) were the most likely to deliver with bodily integrity and have an intact perineum, and Indian and Chinese mothers were the least likely. For eight indicators, the more deprived mothers had better outcomes. In comparative terms, the healthy baby outcome did not vary greatly by ethnicity, while deprivation had less of an effect upon all caesareans.

Bragg et al.[39] in a cross-sectional analysis explored whether or not variation in unadjusted rates of caesarean births could be explained by maternal characteristics and clinical risk factors. However, adjusting for maternal characteristics and clinical risk factors did not greatly reduce the variation between individual trusts, with the observed variation in caesarean rates, suggesting that organisational factors would need to be included in future analysis. This work was furthered by the RCOG in its Clinical Indicators Project,[25] which identified 11 potential performance indicators derived from HES data. This report similarly found unexplained variation across trusts. Organisational factors such as trust configuration, size, models of care, staffing levels and skill mix remained unknown factors and we explore these in the next section.

How do organisational factors affect variability in maternal interventions and maternal and perinatal outcomes?

Approximately 1–2% of the total variation in the outcome indicators was attributable to differences between trusts, whereas 98–99% of the variation was attributable to differences between mothers within trusts. There was marginal improvement in a model's capacity to predict outcomes following the addition of sociodemographic, trust-level and staffing variables. Based on the AUC, most models meet the criteria for fair to good prediction. The key factors are summarised in the next section.

Healthy mother and healthy baby

The results for the three healthy mother and healthy baby indicators (healthy mother, healthy baby, healthy mother/healthy baby dyad) are summarised. The overall fits of these models based on the AUC were all in the range 0.70 to 0.80 (fair).

For mother and baby indicators, and the healthy mother/healthy baby dyad, clinical risk was a strong and dominant predictor, then mother's parity with moderate effects for mother's age and IMD. The relationship with parity was linear for the healthy mother indicator, with mothers more likely to achieve a healthy outcome with increasing parity.

The chance of a healthy mother outcome was negatively associated with deprivation. Mothers belonging to the most deprived IMD quintile were more likely to achieve a healthy mother outcome than mothers belonging to the least deprived quintile. This relationship was reversed, and less strong, for the healthy baby indicator. There was some variation by SHA and rural/urban classification, although the size of these effects was of a much lower order of magnitude than for mothers' characteristics. For the healthy mother/healthy baby dyad the outcome was achieved least often in London.

Mothers attending a university hospital trust were less likely to give birth to a healthy baby. Size of trust had no impact, although there was a negative effect (i.e. the more births in a trust the poorer the outcome) that approached statistical significance for the healthy mother indicator. This may be because of sick mothers and babies being referred to large units, skewing their proportions of healthy mothers and healthy babies, although clinical risk has been controlled for in these models.

Mode of birth

The results for the four mode of birth indicators are summarised. The overall fits of these models based on the AUC were all in the range 0.70 to 0.80 (fair). Parity and clinical risk were the two dominant predictors of outcome. The effect of increasing parity increased the chance of a delivery with bodily integrity, a normal birth, a spontaneous vaginal delivery and an intact perineum. Being at increased clinical risk of complications during the birth reduced the chances of these outcomes.

Women living in more deprived areas were more likely to deliver with bodily integrity, have a normal birth, experience a spontaneous vaginal delivery and have an intact perineum. The effects of geographical location, defined by SHA, and type and density of the area in which mothers lived were of a much lower order of magnitude than the effects of mothers' characteristics.

Giving birth in larger trusts with more deliveries lowered the chances of delivery with bodily integrity and an intact perineum. For spontaneous vaginal delivery, the outcome was better in trusts not attached to a university but again the effect was small. Trust configuration, i.e. whether or not it had midwife-led units, appeared to have no effect upon mode of birth outcomes.

Caesareans

The results for the three caesarean indicators are summarised. The overall model for elective caesareans achieved a noticeably higher AUC statistic than for emergency caesareans and therefore was far better at predicting the outcome. The chances of a caesarean were lowest for mothers aged ≤19 years, rising thereafter with increasing age. Women living in the most deprived areas were less likely to undergo an elective caesarean than those living in the least deprived areas. The relationship is reversed for emergency caesarean and where the effect of deprivation was less strong. There was some variation across SHAs for all caesareans, with women living in London most likely to have a caesarean.

Summary

Variation between trusts represents a comparatively small component of the overall variation (approximately 1–2%). This variation was not substantially reduced by the addition of the independent variables, although reduction was more evident amongst the mode of delivery indicators.

The funnel plots suggest that there was more variability than expected by chance, with more data points outside the control limits. The funnel plots generally confirmed what was found in the multilevel models, based on the change in the residual sigma estimates, that the models did not reduce the variability between trusts to any great degree.

Potentially there is capacity to improve the fit of these models by adding further variables, although a number of the variables that might help in this respect were inadequate for use in the analysis (e.g. smoking, body mass index) because they were either poorly or not consistently recorded.

Overall, the effects of trust size and university status were small. It was more often the case that larger trusts performed less well than smaller trusts. Giving birth in a larger trust reduced a woman's chances of having an intact perineum, giving birth with bodily integrity and having a healthy mother or healthy mother/healthy baby outcome. When the analysis was restricted to trusts with a single OU, the analysis also showed that women were more likely to have an emergency caesarean in a larger unit and less likely to have a normal birth. University trusts seemed to perform as well as, and often better than, their non-university counterparts in the restricted analysis.

What is the relationship between maternity staffing, skill mix and maternal and perinatal outcomes?

Overall, the linear effects of the staffing variables were not statistically significant for eight indicators. However, all women benefited from an increase in midwifery staffing, in terms of retaining an intact perineum and bodily integrity.

Looking at mode of birth indicators, a higher number of midwives (FTE per 100 maternities) and higher levels of overall staffing were associated with improved chance of delivery with bodily integrity and an intact perineum. There was a reduction in the between trust variation across all four mode of birth indicators, particularly for normal birth, when mothers' characteristics, sociodemographics, trust-level and staffing variables were added to the intercept-only model (i.e. the model that contains no independent variables). Staffing variables had a non-significant effect upon the chances of a caesarean.

We then investigated whether the effect of staffing levels upon outcomes could vary according to either a woman's parity or her clinical risk. An analysis of the multiplicative effects of parity and clinical risk with the staffing variables was more revealing.

Risk status: doctors

Higher-risk women were more likely to have an increase in spontaneous vaginal delivery rates and reduced elective caesarean rates with an increase in doctors. Lower-risk women were more likely to remain healthy throughout birth with higher numbers of doctors.

In the presence of higher numbers of doctors, however, lower-risk women were more likely to have an elective caesarean and less likely to have spontaneous vaginal delivery.

Risk status: midwives

Lower-risk women were more likely to have an increase in healthy mother, healthy baby and healthy mother/healthy baby dyad outcomes with an increase in midwives, and less likely to have an emergency caesarean.

Risk status: support workers

Lower-risk women were more likely to have an increase in normal birth, intact perineum and bodily integrity and lower emergency caesarean rates with more support workers. However, women were less likely to have a healthy mother outcome whatever their risk. Higher-risk women were less likely to have a healthy mother, healthy baby or healthy mother/healthy baby dyad outcome.

Parity: doctors

With more doctors, nulliparous women and women with parity of ≥4 were less likely to have an emergency caesarean, and multiparous women were less likely to have an elective caesarean. Women with higher parities were more likely to have higher spontaneous vaginal birth rates as the number of doctors increased and nulliparous women were more likely to have healthy mother, healthy baby and healthy mother/healthy baby dyad outcomes, although rates decrease as parity increases.

Parity: midwives

With more midwives, women in the higher parities were more likely to deliver with an intact perineum and nulliparous women were more likely to have an increase in elective caesareans.

Parity: support workers

The level of support workers appeared to have little significant effect based on parity, although intact perineum rates were more likely to be increased for women with parity > 2.

Differing staff levels and configurations may have an impact on outcomes of quality and safety. Trusts that have higher levels of midwife staffing are more likely to have higher numbers of nulliparous women having elective caesareans. Increasing the number of doctors appears to have most benefit for women at higher risk of complications, but does not benefit lower-risk women. More doctors improve the chance of healthy outcomes in nulliparous women who labour, and midwives have the most beneficial effect when looking after low-risk women. Support workers are also best deployed with low-risk women, reducing intervention rates. However, caution needs to be exercised with any increase in number and deployment of support workers, as they can also have a negative impact on the healthy mother and healthy baby outcomes in all groups.

Murray *et al.*[77] found no significant differences in results between hospitals with high- and low-quality coded HES data, suggesting that hospitals with high birth record completeness may be generalisable and representative of all hospitals. However, caution should be exercised regarding relevance to other countries and health-care systems. The small relative effect of staffing may be due to limited variation in staffing levels and skill mix. This may be due to the influence of guidelines regarding medical and midwifery staffing. In countries with wider variations, staffing may be shown to have a greater effect.

What is the relationship between maternity staffing, cost and outcomes?

We examined the relationship between maternity workforce staffing levels, quality and safety outcomes and cost. Data on medical staffing were not detailed enough to ascertain the split between obstetric and gynaecological responsibilities in a trust. Therefore, this analysis investigated the relationships between midwifery staffing levels, where data quality was very good. For this analysis we used trust-level data to investigate relationships between outcome measures, midwifery staffing levels and the cost of providing maternity services for NHS trusts in England.

When the case-mix adjustments were included in the model there was no relationship between antenatal spend and reduced operative delivery rates and lower delivery costs. No relationship was found between the proportion of maternity expenditure spent on antenatal care and operative delivery rates after adjusting for maternal characteristics and trust size.

Higher midwifery staffing levels were associated with higher costs of each delivery, although the relationship was not strong. Only around 17% of the variation between trusts' delivery costs could be accounted for by variables included in this model. The remaining variation in the average cost of a delivery was not accounted for by maternal characteristics, size of trust, number of FTE registered midwives employed or antenatal spend and must be due to other factors not included in the analysis.

The analysis found that higher operative delivery rates were not significantly associated with higher delivery costs after adjusting for maternal characteristics. Variations in costs between trusts were not related to the numbers of women having operative deliveries.

There was no association between delivery cost per delivery and the normal birth rate. There was no association between delivery cost per delivery and the intact perineum rate, or any of the three healthy mother and healthy baby indicators, and women's experience of maternity care as measured by the average of the CQC scores. A relationship could not be found that explained postnatal costs in terms of variations in operative delivery rates once adjustments were made.

Having a higher proportion of women at increased risk of complications was associated with more expensive maternity care. Level of social deprivation approached significance with deliveries in trusts, with a higher proportion of women living in deprived areas having more costly deliveries. However, in this analysis trust size and case mix accounted for only 22% of the variation in total delivery costs. Some of the previous analyses also showed relationships between costs and maternal characteristics. For trusts with a larger proportion of women at increased risk of complications, delivery costs were higher and they spent a higher proportion of their total maternity costs on postnatal care, and trusts with a higher proportion of nulliparous women had higher costs per delivery.

Larger trusts, measured by the numbers of deliveries, appeared to offer maternity services that cost less than those in smaller trusts. Repeating the analysis for the 50 trusts that only consist of a single OU, the significant effect of size increased, suggesting that the economies of scale were greater in a single site. The size of the trust had no relationship with the cost of a delivery only (i.e. when costs of antenatal or postnatal care were excluded). Deliveries were not cheaper in larger units once adjustments had been made for maternal characteristics and numbers of midwives. This lack of relationship held when repeating the analysis for the 50 trusts that delivered maternity services through only a single OU each.

Larger trusts were associated with lower CQC scores for women's experience of maternity care. There was no relationship between trust size and the other outcome measures: normal birth, intact perinea and number of complaints.

From this study, the increased investment in staff did not necessarily have an effect on the outcome and experience measures chosen, where there was in general no relationship with midwifery staffing levels. However, there was a higher intact perineum rate in trusts with higher levels of midwifery staffing. Although this validates the result to some extent, any trusts submitting erroneous data will be correlated between years. It would be interesting to see if the relationship holds good if the data are analysed at individual record level. Maternity units with higher levels of midwifery staffing may find it easier to provide continuity of carer and one-to-one midwifery care. This finding would then be consistent with research evidence that one-to-one midwifery care can result in a significant reduction in perineal trauma.[70] This is an important outcome for women which impacts on the quality of their life with a new baby and could impact on future decisions about mode of birth.[96]

Although there was no significant improvement in women's experiences of care, as measured by CQC scores, as a result of higher staffing, there was a trend in that direction. These scores were not all directly related to midwifery staffing and covered a wide range of women's experience of care.

While overall levels of midwifery staffing are important, other factors will also affect outcomes and experience of care for women. The deployment of those staff within the maternity service, their attitude towards the women they care for, their skills and the culture within which they work will all play a part in women's care.

Despite the relationship of higher levels of midwifery staffing with improved intact perineum rates, and the higher costs associated with providing that staffing, when tested directly there was no relationship detected between delivery costs and improved perineal outcomes. In fact, few patterns connected the cost of providing maternity services with differences in the populations they serve or the complexity of births. Higher operative delivery rates were not associated with higher costs of intrapartum care (after adjusting for case mix) and this emphasises the inherent variability found in the reference costs, as found by Laudicella *et al.*[67] Excess bed-days were not included with delivery costs in these analyses (instead being included in postnatal costs). It may be that including the extra duration of stay following an operative birth would have an impact on this relationship,[68] although this is unlikely, as the cost of excess bed-days was very small in relation to total cost. Trusts with higher numbers of women at increased risk of complications do appear to be associated with higher delivery costs, as found by Gaughan *et al.*,[68] and also with higher postnatal costs. This pattern is reflected in the new maternity pathway system, where higher payments are made for higher-risk women, but interventions themselves are not rewarded.

Providing more costly antenatal care alone did not appear to have a direct impact on operative delivery rates and was not recouped by a commensurate reduction in delivery costs, nor did it result in a better experience of antenatal care as measured by the CQC antenatal summary score. Some trusts did appear to be able to increase antenatal expenditure and reduce delivery costs. However, these trusts may be attributing costs differently between antenatal spend and delivery cost, which would create the illusion of a trade-off. From the data it is not possible to decide if these trusts have a winning formula.

The analyses generally adjusted for differences in size of trust as measured by the annual number of deliveries. The biggest effect of trust size was seen with total costs, particularly when restricted to trusts which had a single site only. Costs per childbearing woman were lower for larger units after taking into account case mix, which may have been a result of economies of scale. As no relationship could be seen between trust size and delivery cost alone, it could be deduced that the differences are related to antenatal or postnatal costs. However, the wide variations in costs suggest that trusts may be allocating costs differently along the maternity pathway.

Women seemed to be slightly less satisfied with their experience of care, as measured by CQC scores, in larger trusts than in smaller ones. Again, because of the complexities of trust configurations, it may be that size of unit is more important than size of trust, but CQC scores are available at trust level only. Although larger trusts appear to spend less overall, and may offer a lower quality of service from the point of view of women's experience, this type of analysis does not show causality and therefore it is not possible to assume that spending less on maternity care results in lower-quality care.

Relationships between staffing, costs and outcomes are complex and it is perhaps unsurprising that there was often no clear correlation between these different variables. However, these results do not contradict the idea that quality of care is not directly related to the costs of providing the services and that quality and cost do not need to require a trade-off. We explore this in the economic modelling.

Economic modelling

The economic analysis was frustrated by the data limitations described in the next section, mainly with respect to the availability of detailed staffing data. Despite this, a number of interesting findings emerged from the economic modelling, which were consistent across the two definitions of maternity output: the total number of deliveries per trust and the cost-weighted deliveries per trust. The latter definition was an attempt to control for the greater complexity and resource use involved in performing caesarean births, which is accounted for in the higher reimbursement received by trusts.

To accommodate this, the cost-weighted deliveries combined these two broad types of deliveries based on their relative cost. However, the results were very similar across both models and we can therefore discuss them together. The Hick's elasticities of complementarity were estimated. This measures the extent to which staff groups are complements or substitutes, i.e. whether they should be used in combination or can be used in place of each other in terms of the number of deliveries a provider can handle each year. This measure indicated that midwives were complements with consultants and other doctors in the production of deliveries, that is, midwives and doctors and midwives and consultants should be used in combination to maximise the number of deliveries a provider handles. Consultants and other doctors were found to be substitutes for each other, as were midwives and support workers. In that sense, there are some tasks that a consultant and more junior doctor can both do, as there are tasks that both midwives and support workers can perform. This echoes findings from research by Goryakin *et al.*[62] that looks at the relationship between registered nurses and health-care assistants. The elasticities in *Table 34* indicate the degree to which staff groups are either complements or substitutes and none of them imply that you simply trade, for example, one FTE midwife for one support worker.

As described in the literature review, this is the first study that we are aware of to look at this issue and, therefore, there are no appropriate studies with which to compare our findings. Thurston and Libby[65] employed a similar methodology but applied to primary care physicians in America. They found that all staff groups were complements except for technicians, who were substitutes for both nurses and administrators. Our findings are similar in that midwives and support staff are also found to be substitutes, while most other groups are found to be complements. We also find that consultants and other doctors are substitutes for each other, which was not an issue considered by Thurston and Libby[65] because they considered primary care practices and the grade or experience of the physician was not modelled.

The marginal products of all staff groups were positive, indicating that adding an additional worker of any type would increase the total number of deliveries a provider would be able to cope with. Purely in terms of the number of additional deliveries and not considering the cost of adding an additional worker, doctors had the highest marginal productivity followed by consultants, then midwives and finally support workers. Adding an additional FTE worker in each of these categories would allow a hospital to handle an additional 43, 32, 18 and 10 deliveries per annum, respectively. Conversely, reducing staffing by one member of staff in each group would reduce a hospital's output by an equivalent amount. Given that the national average number of deliveries per FTE in these staff groups respectively is 180, 384, 32.5 and 100, the marginal productivities support the results that most staff groups are complementary, i.e. they should be used in combination. Adding additional workers of one category (e.g. midwives) will not be as productive as adding a combination of all the staff groups: the right skill mix is therefore critical to the efficient operating of a maternity service. However, substituting support workers for midwives may also have an impact on some aspects of the quality of care depending on the groups of women involved or the care setting.

Limitations

Secondary analysis is dependent on the quality of data. We used the full census of women's deliveries in HES (656,969 delivery records) so there was no bias caused by non-response. Any biases would therefore be caused by missing data, poorly recorded data or omitted variables from the risk adjustment model. A scoring system was used to select records with the largest amount of most useful and relevant data of greatest relevance to the project. Extensive data cleaning was conducted to remove duplicates and records which did not relate to a delivery episode, identifying units with inconsistent or missing data. A decision was taken not to include any trust where fewer than 80% of women could be coded for a particular indicator. This limited the use of some potential indicators.

A common problem when analysing routinely collected data is that it is only possible to work with the available data that are of a sufficient quality to use. For example, we could not include body mass index or an indicator of smoking status in our models because of data quality issues, although they are known to be important risk factors. There may well be other measures of clinical well-being and lifestyle that go beyond the NICE risk classification that would help to reduce variability.

Only a limited set of trust-level organisational variables were used. We were not able to include measures that tell us something about the organisation (e.g. organisational climate, local climate) and models of care that could be predictive of outcome. Our models may also have omitted other variables, either known or unknown, that are predictive of outcome.

Staffing data were available only at trust level so we could not explore the effects of staffing at the unit level. The data for trusts that have multiple units could not be disaggregated. Aggregated trust-level data makes the assumption that unit-level effects within a trust are similar, which may not be true. The staffing data are taken from a census undertaken every September. This single-point estimate will hide any fluctuations that may occur over time. We analysed data that were aggregated over a period of a year. These data will therefore miss those occasions when the service is placed under stress, or reaching a critical point, because of excess deliveries, low staffing levels or other factors.

This study, like many others in the literature, relies on a single cross-section of data that makes causal inference problematic if not impossible. At best we can claim an association between our independent variables and the outcomes. An obvious concern is omitted variable bias. Some potentially omitted variables have been listed above, including inter alia smoking status (for mothers) and a number of trust level variables. It is possible that trusts with higher staffing levels also have higher levels of other inputs that affect organisational performance and the quality of care, such as advanced medical equipment, high performance management teams or a culture of patient safety. A major problem involves the potential endogeneity between staffing levels and the outcomes we have used. In effect, trusts make decisions about factors that have an impact on the quality of care, for example staffing, subject to a set of constraints such as regulation, limited budget and case mix.

Adding additional years of data would allow for some control over unobserved variables that vary across providers but do not vary over time. However, only in an experiment that deliberately (or fortuitously as the result of a policy design) allows for the manipulation and randomisation of staffing levels could researchers make casual claims about the relationship between staffing and outcomes. Therefore, this concern is not unique to this study.

Limitations of cost analysis

This investigation brought together data from a wide variety of sources, including trust profiles and outcome data from an extensively cleaned HES data set. The analysis included measures of women's experience of care.

There are a number of limitations to this analysis. Firstly, the analyses were done using data aggregated at trust level, rather than at the level of individual patient records. This analysis, because of data limitations, has considered only the number of registered midwives and not taken into account the use of maternity support workers, nor has it considered the medical workforce or nursing staff working alongside midwives.

Reference costs and staffing data were available only at trust level, and not for individual maternity sites within trusts. There were a wide variety of configurations of maternity services with trusts operating one or more obstetric sites and varying numbers of AMUs and FMUs, and differences in costs, staffing and outcomes between these units may have masked some associations. Because of this, some analyses were undertaken using trusts that operate only a single obstetric site. Extreme caution needs to be exercised regarding the use of reference costs versus actual costs.

Ethnicity was not included as a variable in this analysis. As a categorical variable with a large number of potential coefficients, its inclusion would have been detrimental to the determination of the other coefficients. Grouping ethnicity would have reduced the number of coefficients, but ethnicities that seem similar have different outcomes (e.g. Indian and Pakistani women). As ethnicity is only a minor driver of outcome, it has been excluded from the analyses.

Limitations of the economic modelling

The economic analysis was limited by a number of, primarily, data-related factors. While the quality of the workforce data has steadily improved since the introduction of the Electronic Staff Record system, the data remain limited for this type of research. The data were reported at trust rather than unit level, and there was no account for time spent in different roles or departments (e.g. obstetric vs. gynaecology). There were also limitations in the availability of data on bank and agency staff used.

It was not possible to obtain credible data on the other inputs such as capital and variable inputs such as drugs. Therefore, it was not possible to investigate any input substitution between, say, capital and labour. Given the focus of this study on skill mix, this was less of a concern for the research team.

Finally, the functional form adopted for the production function analysis induces multicollinearity in the variables due to the inclusion of cross-products to test for substitution and complementarity of inputs. Given the relatively low degrees of freedom in the models resulting from a small number of trusts (144) in the data set, few of the variables were statistically significant. It is not possible to ascertain whether this was due to the high degree of multicollinearity or the lack of a statistically significant relationship in the data. Yet as there was a high R-squared value, it is likely that multicollinearity is the cause and that the variables are indeed significant predictors of output. There are few options to resolve this issue because it is not possible to collect additional data on trusts, nor is it appropriate to reduce the functional form to a simpler specification such as a Cobb–Douglas production function.

Public and patient involvement

This project has had PPI at every stage of the research project, from the development of the proposal, through undertaking some of the research, providing an advisory role, drafting and commenting on the report, to dissemination of the findings.

As a result the findings will be relevant not only to clinicians, policy-makers and NHS managers, but to pregnant women and their families. A particular impact has been on the choice of indicators used as measures of the quality of maternity care. These include a number of positive measures, rather than just a series of interventions or harms. These have included existing measures such as normal birth. This used a consensus definition agreed by the Maternity Care Working Party,[17] but based on a definition originally proposed by the user organisation BirthChoiceUK. Further innovative positive measures which indicate an absence of physical harm are those of intact perineum and delivery with bodily integrity. Being without damage or sutures in the early postnatal days enables women to recover more quickly from the birth and to feel more physically comfortable as they begin the tiring work of looking after a new baby. A trio of further measures reflect the health of the mother, the baby and the mother and baby dyad. These combine the concepts of safety, clinical effectiveness and women's experience to view how the well-being of both the mother and baby as they emerge from the birth process, whatever that may have been.

Our lay collaborator was also able to undertake part of the research herself, taking on responsibility with the obstetrician (SB) for identifying codes in women's HES records which might indicate that a women would be considered to be at increased risk of complications for birth. This proved to be one of the most important characteristics in determining outcomes for women, according to the multilevel model, along with parity. This work has implications for future analyses, where outcomes can be stratified by risk and parity, and potentially provide women with more personalised information about their likelihood of particular outcomes.

One of the challenges of involving recent service users in this study has been that it has largely been a paper exercise, analysing routinely collected data and understanding the results of multilevel logistic regression, together with the selection of quality indicators. For someone familiar with quality metrics and HES maternity data and with a basic statistical understanding, this has not been as challenging as it might have been to a less expert user.

Although joining the project team and undertaking some of the research herself has brought benefits to the project, it has required a time commitment that might have been difficult for other service user representatives. In this case, the project was not able to provide full funding for this (as the time commitment and contribution were recognised only once the project had started). However, it is pertinent to consider for the future how such time should be costed in for 'expert patients'. They may have little academic record but a wealth of experience and generally go unfunded. Universities must consider how to support payment to such self-employed individuals within an increasingly tightly regulated environment.

There also appears to be a trend for those reviewing grant proposals to discount the PPI contribution of experienced service users, and user-researchers who work independently. While having recent service users with a variety of experiences and sociodemographic backgrounds is clearly important to many projects, experienced service users can bring a wider, less personal perspective, bringing together views of many consumers, and often with a knowledge of the research literature and policy from a patient perspective also. It is vital that this not be undervalued by reviewers of grant proposals.

Chapter 5 Conclusions

Implications for health care

Much of the variation in outcomes that was measured at trust level was explained by clinical risk and parity. The powerful impact of changing age, parity and clinical risk factors (even without the information about obesity and smoking in this study) have important implications for the 'fitness' of the reproducing population in terms of costs as well as outcomes. This study draws attention to the need to consider all these factors in terms of public health policy.

This study is the most comprehensive analysis of all women giving birth in England, which has adjusted for all the clinical risk factors present at the end of pregnancy and prior to labour as outlined in the NICE intrapartum care guidelines. It provides useful data for trusts, clinicians, women and their families in relation to variations in outcome. Knowing the extent to which rates of normal birth, for example, are affected by parity and risk provides a useful basis for planning maternity care for a given population, and for discussing likelihood of different outcomes with women. It also highlights the importance of adjusting for age, parity and clinical risk when comparing outcomes between trusts.

There is some indication that staffing levels have positive and negative effects on some outcomes, and deployment of doctors and midwives where they have most beneficial impact is important. Approximately 1–2% of the total variation in the outcome indicators was attributable to differences between trusts, whereas 98–99% of the variation was attributable to differences between mothers within trusts. There was less variation in staffing level, possibly because of existing staffing standards for the midwifery and obstetric workforce. Levels of midwifery staffing were associated with only 2 of the 10 indicators, delivery with bodily integrity and intact perineum. The models estimated population effects, which may not be a true reflection for all trusts particularly those with more diverse characteristics. Certain multiplicative effects revealed themselves and showed that the effect of staffing upon outcomes sometimes varied according to mother's parity and clinical risk. However, there is potential here for reverse causation where units with higher proportions of high-risk women increase their staff to meet that demand.

Our findings are based upon current staffing regimes. Any extrapolations based on staffing levels not often encountered in this sample would have to be interpreted with a high degree of caution. The measure for medical staffing was problematic because it was not possible to disaggregate this into contributions made to obstetric care (over the course of a whole pregnancy) and maternal delivery (on one day) as opposed to gynaecology care. So, while the findings are accurate, they were not strictly 'fit for purpose' and could have led to bias in estimated effects. The preferred measure to full-time equivalence for all staff groups would have been hours committed to maternal delivery care, or FTE for all groups over a whole pregnancy, birth and postnatal pathway.

The staffing measures related to only one time point. Therefore, our results do not reflect the temporal fluctuations that would be experienced on a daily or weekly basis, where both number of maternities and staff availability would vary. Those critical points when the service is most under strain would remain hidden. A measure of bank and agency staff usage would also have added to our overall understanding, but good-quality data on bank and agency use in NHS maternity care were not available to this study. The results, therefore, assume that what occurs at the aggregated level (year) would have been replicated had staffing data been available over shorter time periods (e.g. a month or a week).

In the era of 'Big Data' it is essential that more detailed information on staffing be made available and disaggregated over shorter periods so that questions similar to those pursued in this research can be better answered. As each maternity unit will have its own unique set of circumstances, the acquisition of additional contextual data is an imperative, for example measures of the local and organisational environment that reflect the way the unit is run.

In sum, it is very hard to say what a 'staffing' ratio or model is, when the most cost-effective care is continuity of midwifery care antenatally/intrapartum/postnatally where midwives provide care for women either in the community or in an acute trust facility, which could be an obstetric or midwife-led setting. Thus, staffing measures that focus on ratios on a ward will miss this, especially when the care takes place over 6 months both in and out of hospital.

What these data add to what is already known is some indication that midwifery staffing levels do make a difference for a small number of indicators, although we cannot quantify accurately how this could be reliably deployed into workload models.

Managers may wish to exercise caution in increasing the number of support workers who care for higher-risk women. Careful attention needs to be paid to the relationships between staff groups and the potential and limits of skill mix and role substitution. Most staff groups were complements rather than substitutes, suggesting that substituting staff groups may actually harm productivity. Collecting more detailed workforce data at the level of the individual unit and greater detail over the time spent on different activities (e.g. obstetrics vs. gynaecology) would enable a number of the limitations of this research to be overcome.

Data on medical staffing were not detailed enough to ascertain the split between obstetric and gynaecological responsibilities in a trust. Therefore, this analysis investigated the relationships between midwifery staffing levels, where data quality was very good. For this analysis we used trust-level data to investigate relationships between outcome measures, midwifery staffing levels and the cost of providing maternity services for NHS trusts in England.

Higher midwifery staffing levels were associated with higher costs of each delivery, although the relationship was not strong. Only around 17% of the variation between trusts' delivery costs could be accounted for by variables included in this model. The remaining variation in the average cost of a delivery was not accounted for by maternal characteristics, size of trust, number of FTE registered midwives employed or antenatal spend and must be due to other factors not included in the analysis.

After adjusting for maternal characteristics and trust size, no relationship was found between the proportion of expenditure spent on antenatal care and operative delivery rates, or between higher operative delivery rates and higher delivery costs. Variations in costs between trusts were not related to the numbers of women having operative deliveries.

There was no association between cost per delivery and the normal birth rate, intact perineum rate, or any of the three healthy mother and healthy baby indicators, and women's experience of maternity care as measured by the average of the CQC scores. A relationship could not be found that explained postnatal costs in terms of variations in operative delivery rates once adjustments were made.

Having a higher proportion of women at increased clinical risk was associated with more expensive maternity care, and area deprivation approached statistical significance. These factors are currently reflected in the maternity pathway payment system. There appeared to be economies of scale across the total maternity episode (antenatal, intrapartum and postnatal care) by trust size, which were increased in trusts on a single site. However, larger trusts were associated with lower scores for women's experience, although there was no relationship found between trust size and any clinical outcome.

86

From this study, the increased investment in staff did not necessarily have an effect on the outcome and experience measures chosen, where there was in general no relationship with midwifery staffing levels, apart from a higher intact perineum rate, and delivery with bodily integrity in trusts with higher levels of midwifery staffing.

Data quality in HES data is improving but a significant number of trusts have poor-quality data, and some important variables such as smoking and body mass index are poorly recorded. Measures that aim to influence improvement in data quality, such as the RCOG benchmarking project, are to be encouraged.

Current benchmarking indicators that are derived from HES data do not cover a range of outcomes of importance to women. However, this study has developed new indicators that can be derived from HES data and provide an indication of a healthy outcome such as healthy mother, healthy baby and the healthy mother/healthy baby dyad. These should be considered for inclusion in future work. These indicators provide a wider range with which to assess quality of care in maternity care where the overall aim is to promote health and well-being and prevent harm.

The relative impact of maternity workforce staffing is a key policy topic. There is a paucity of evidence regarding the effects of the number and grade of medical and midwifery staffing, and the impact of skill mix with unqualified staff. The evidence underpinning current workforce models such as BR+ is slight. This study has highlighted some effects, but also shown that the workforce needs to be assessed as a whole rather than separately.

The development of appropriate quality indicators for maternity care is a global initiative and has often focused on measurement of interventions as a proxy. The development of harm-free indicators and indicators of a healthy mother and healthy baby are an important development. In addition, there need to be valid indicators of women's experience. Our analysis of the CQC women's experience survey indicates that this may be a fruitful avenue to pursue, in providing another source of data to measure quality.

Recommendations for research

To examine the impact of policy, studies comparing different approaches in similar industrialised nations (e.g. across Europe), should be performed to understand, and potentially modify, current adverse trends for reproduction.

There are wide variations in a range of outcomes that remain after adjustment for sociodemographic and background risk. Further research is required on what may be influencing unexplained variation such as organisational climate and culture, use of NICE guidelines in practice and variation of models of care within trusts.

Further research is required on the validation and use of the new indicators developed to assess quality of care in future years as HES data improve.

The longer-term impact on the baby needs to be assessed by linking the mother and baby records within HES.

Further research is also needed on how to measure more accurately which factors in levels of consultant obstetrician, midwife and support worker staffing impact on outcomes and quality, in addition to productivity.

Further research is required on developing indicators of quality derived from women's experience and these should be informed by what is important to women and their families.

Acknowledgements

We acknowledge the contribution of the advisory group, which involved key stakeholders and representatives from user groups.

Contributions of authors

Jane Sandall (Professor of Women's Health) supervised analysis and drafted report.

Trevor Murrells (statistician/research data manager) developed and tested the risk model, conducted multilevel analysis and contributed to writing the report.

Miranda Dodwell (user representative, BirthChoiceUK and visiting research associate, King's College London) provided a user perspective, developed and tested the risk classification, conducted costing analysis and contributed to writing the report.

Rod Gibson (data analyst) cleaned and validated the data, developed algorithms, developed and tested the risk classification, conducted costing analysis and contributed to writing the report.

Susan Bewley (Professor of Complex Obstetrics) provided a clinical perspective to development of the risk model and contributed to writing the report.

Kirstie Coxon (research associate) provided project and report management and contributed to writing the report.

Debra Bick (Professor of Evidence Based Midwifery Practice) provided a clinical perspective and contributed to writing the report.

Graham Cookson (Professor of Economic and Public Policy) provided an economic perspective and contributed to writing the report.

Cathy Warwick (chief executive officer, Royal College of Midwives) provided a midwife and support worker workforce perspective and contributed to writing the report.

Diana Hamilton-Fairley (Director of Education and Quality, Health Education South London) provided an obstetric workforce perspective.

References

1. Department of Health. *The Operating Framework for the NHS in England 2010/11*. London: Department of Health; 2009.

2. Department of Health and Department for Education and Skills. *National Service Framework for Children, Young People and the Maternity Services*. London: Department of Health; 2004.

3. Department of Health. *Maternity Matters: Choice, Access and Continuity of Care in a Safe Service*. London: The Stationery Office; 2007.

4. Department of Health. *The Mandate: A Mandate from the Government to the NHS Commissioning Board: April 2013 to March 2015*. London: Department of Health; 2012.

5. Department of Health. *2012/13 Choice Framework*. London: Department of Health; 2012.

6. Department of Health. *The NHS Outcomes Framework 2013–14*. London: Department of Health; 2012.

7. Royal College of Anaesthetists, Royal College of Obstetricians and Gynaecologists, Royal College of Paediatrics and Child Health, Royal College of Midwives. *Safer Childbirth: Minimum Standards for the Organisation and Delivery of Care in Labour*. London; RCOG Press; 2007.

8. Royal College of Midwives. *Staffing Standard in Midwifery Services (Position Statement No. 15)*. London: RCM; 2009.

9. Royal College of Obstetricians and Gynaecologists. *High Quality Women's Health Care: A Proposal for Change*. London: RCOG Press; 2011.

10. Department of Health. *The Nursing Roadmap for Quality: A Signposting Map for Nursing*. London: Department of Health; 2010.

11. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. 1st edn. Washington, DC: National Academies Press; 2001.

12. Darzi A. *High Quality Care for All: NHS Next Stage Review Final Report*. London: The Stationery Office; 2008.

13. Chief Nursing Officers of England, Northern Ireland, Scotland and Wales. *Midwifery 2020: Delivering Expectations*. London: Department of Health; 2010.

14. NHS Scotland. *The Healthcare Quality Strategy for NHS Scotland*. Edinburgh: Scottish Government; 2010.

15. Department of Health. *Equity and Excellence: Liberating the NHS*. London: Department of Health; 2010.

16. Department of Health. *The NHS Outcomes Framework 2011/12*. London: Department of Health; 2010.

17. Maternity Care Working Party. *Making Normal Birth a Reality: Consensus Statement from the Maternity Care Working Party*. 2007. URL: www.rcog.org.uk/en/guidelines-research-services/guidelines/making-normal-birth-a-reality (accessed 3 October 2014).

18. NHS Commissioning Board. *Commissioning for Quality and Innovation (CQUIN): 2013/14 Guidance*. 2013. URL: www.england.nhs.uk/wp-content/uploads/2012/12/cquin-guidance.pdf (accessed 5 September 2014).

19. QIPP Safe Care Team. *Delivering the NHS Safety Thermometer CQUIN 2012/13: A Preliminary Guide to Delivering 'Harm Free' Care*. Salford: Department of Health; 2012.

20. NHS Information Centre. *Indicators for Quality Improvement*. 2013. URL: https://mqi.ic.nhs.uk/ (accessed 5 September 2014).

21. Advancing Quality Alliance (AQuA). *Advancing Quality Alliance*. 2013. URL: www.advancingqualityalliance.nhs.uk/ (accessed 5 September 2014).

22. Child and Maternal Health Intelligence Network. *NHS Atlas of Variation in Healthcare for Children and Young People*. Public Health England; 2010. URL: www.rightcare.nhs.uk/index.php/atlas/ children-and-young-adults/ (accessed 5 September 2014).

23. Donabedian A. The quality of care: how can it be assessed? *J Am Med Assoc* 1988;**260**:1743–8. http://dx.doi.org/10.1001/jama.1988.03410120089033

24. Raleigh VS, Foot C. *Getting the Measure of Quality: Opportunities and Challenges*. London: The King's Fund; 2010.

25. Royal College of Obstetricians and Gynaecologists. *Patterns of Maternity Care in English NHS Hospitals 2011/12*. London: RCOG Press; 2013.

26. The Stationery Office. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry* (chaired by Robert Francis, QC). London: The Stationery Office; 2013.

27. Health and Social Care Information Centre. *Hospital Episode Statistics, Admitted Patient Care – England 2011–2012*. 2012. URL: www.hscic.gov.uk/article/2021/Website-Search? productid=9161&q=admitted+patient+care+England+2011-12&sort=Relevance&size=10& page=1&area=both#top (accessed 10 September 2014).

28. Gregory KD, Korst LM, Lu MC, Fridman M. AHRQ patient safety indicators: time to include hemorrhage and infection during childbirth. *Jt Comm J Qual Patient Saf* 2013;**39**:114–22.

29. Bailit JL. Measuring the quality of inpatient obstetrical care. *Obs Gynecol Surv* 2007;**62**:207–13. http://dx.doi.org/10.1097/01.ogx.0000256800.21193.ce

30. Agency for Healthcare Research and Quality. *Quality Improvement: Inpatient Quality Indicators #33: Technical Specifications. Primary Cesarean Delivery Rate*. Rockville, MD: Agency for Healthcare Research and Quality; 2012.

31. Drösler S. *Facilitating Cross-National Comparisons of Indicators for Patient Safety at the Health-System Level in the OECD Countries*. *DELSA/ELSA/WD/HTP(2008)1*. Organisation for Economic Co-operation and Development (OECD): Paris; 2008.

32. Sawyer A, Ayers S, Abbott J, Gyte G, Rabe H, Duley L. Measures of satisfaction with care during labour and birth: a comparative review. *BMC Pregnancy Childbirth* 2013;**13**:108. http://dx.doi.org/ 10.1186/1471-2393-13-108

33. Paranjothy S, Frost C, Thomas J. How much variation in CS rates can be explained by case mix differences? *BJOG* 2005;**112**:658–66. http://dx.doi.org/10.1111/j.1471-0528.2005.00501.x

34. World Health Organization. *International Classification of Diseases, Version 10*. 2010. URL: http://apps.who.int/classifications/icd10/browse/2010/en (accessed 5 September 2014).

35. NHS Connecting for Health. *Office of Population, Censuses and Surveys Classification 4.4*. 2012. URL: http://webarchive.nationalarchives.gov.uk/20130502102046/http://connectingforhealth. nhs.uk/systemsandservices/data/clinicalcoding/codingstandards/opcs4/index_html (accessed 5 September 2014).

36. Raleigh VS, Cooper J, Bremner SA, Scobie S. Patient safety indicators for England from hospital administrative data: case–control analysis and comparison with US data. *BMJ* 2008;**337**:a1702.

37. Bottle A, Aylin P. Intelligent information: a national system for monitoring clinical performance. *Health Serv Res* 2008;**43**:10–31. http://dx.doi.org/10.1111/j.1475-6773.2007.00742.x

38. Essex H, Green J, Baston H, Pickett K. Which women are at an increased risk of a caesarean section or an instrumental vaginal birth in the UK: an exploration within the Millennium Cohort Study. *Br J Obstet Gynaecol* 2013;**120**:732–43. http://dx.doi.org/10.1111/1471-0528.12177

39. Bragg F, Cromwell DA, Edozien LC, Gurol-Urganci I, Mahmood TA, Templeton A, *et al.* Variation in rates of caesarean section among English NHS trusts after accounting for maternal and clinical risk: cross sectional study. *BMJ* 2010;**341**:c5065. http://dx.doi.org/10.1136/bmj.c5065

40. Department of Health. *The Year: NHS Chief Executive's Annual Report 2008/09*. London: Department of Health; 2009.

41. Health Care Commission. *Towards Better Births: A Review of Maternity Services in England*. London: Commission for Healthcare Audit and Inspection; 2008.

42. The King's Fund. *Safe Births, Everybody's Business*. London: The King's Fund; 2008.

43. National Patient Safety Agency. *Quarterly Data Summary Issue 13: Learning from Reporting – Staffing*. 2009. URL: www.nrls.npsa.nhs.uk/resources/collections/quarterly-data-summaries/?entryid45=62052 (accessed 5 September 2014).

44. Jarman B, Gault S, Alves B, Hider A, Dolan S, Cook A, *et al.* Explaining differences in English hospital death rates using routinely collected data. *BMJ* 1999;**318**:1515–20. http://dx.doi.org/10.1136/bmj.318.7197.1515

45. Rafferty AM, Clarke SP, Coles J, Ball J, James P, McKee M, *et al.* Outcomes of variation in hospital nurse staffing in English hospitals: cross-sectional analysis of survey data and discharge records. *Int J Nurs Stud* 2007;**44**:175–82. http://dx.doi.org/10.1016/j.ijnurstu.2006.08.003

46. Schubert M, Glass TR, Clarke SP, Aiken LH, Schaffert-Witvliet B, Sloane DM, *et al.* Rationing of nursing care and its relationship to patient outcomes: the Swiss extension of the International Hospital Outcomes Study. *Int J Qual Heal Care* 2008;**20**:227–37. http://dx.doi.org/10.1093/intqhc/mzn017

47. Shuldham C, Parkin C, Firouzi A, Roughton M, Lau-Walker M. The relationship between nurse staffing and patient outcomes: a case study. *Int J Nurs Stud* 2009;**46**:986–92. http://dx.doi.org/10.1016/j.ijnurstu.2008.06.004

48. Van den Heede K, Sermeus W, Diya L, Clarke SP, Lesaffre E, Vleugels A, *et al.* Nurse staffing and patient outcomes in Belgian acute hospitals: cross-sectional analysis of administrative data. *Int J Nurs Stud* 2009;**46**:928–39. http://dx.doi.org/10.1016/j.ijnurstu.2008.05.007

49. Van den Heede K, Lesaffre E, Diya L, Vleugels A, Clarke SP, Aiken LH, *et al.* The relationship between inpatient cardiac surgery mortality and nurse numbers and educational level: analysis of administrative data. *Int J Nurs Stud* 2009;**46**:796–803. http://dx.doi.org/10.1016/j.ijnurstu.2008.12.018

50. Kane R, Shamliyan T, Mueller C, Duval S, Wilt T. The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis. *Med Care* 2007;**45**:1195–204. http://dx.doi.org/10.1097/MLR.0b013e3181468ca3

51. Harris DA, Vernon JP, Sidman RD, Sucov A, Aarbelaez C, Monks A. Is there a doctor in the house? Comparing attendings, residents, and midlevel providers to increase clinician staffing in an academic emergency department: a cost-efficiency analysis. *Ann Emerg Med* 2004;**44**(Suppl. 1):S18–19. http://dx.doi.org/10.1016/j.annemergmed.2004.07.061

52. Sucov A, Sidman R, Valente J. A cost-efficiency analysis to increase clinician staffing in an academic emergency department. *Acad Med* 2009;**84**:1211–16. http://dx.doi.org/10.1097/ACM.0b013e3181b187fc

53. Pronovost PJ, Angus DC, Dorman T, Robinson KA, Dremsizov TT, Young TL. Physician staffing patterns and clinical outcomes in critically ill patients: a systematic review. *JAMA* 2002;**288**:2151–62. http://dx.doi.org/10.1001/jama.288.17.2151

54. Judge A, Chard J, Learmonth I, Dieppe P. The effects of surgical volumes and training centre status on outcomes following total joint replacement: analysis of the Hospital Episode Statistics for England. *J Public Health* 2006;**28**:116–24. http://dx.doi.org/10.1093/pubmed/fdl003

55. Kc D, Terwiesch C. Impact of workload on service time and patient safety: an econometric analysis of hospital operations. *Manage Sci* 2009;**55**:1486–98. http://dx.doi.org/10.1287/mnsc.1090.1037

56. Masnick K, McDonnell G. A model linking clinical workforce skill mix planning to health and health care dynamics. *Hum Resour Health* 2010;**8**:11. http://dx.doi.org/10.1186/1478-4491-8-11

57. Jones S, Bottle A and Griffiths P. *An Assessment of Failure to Rescue Derived from Routine NHS Data as a Nursing Sensitive Patient Safety Indicator*. National Nursing Research Unit, King's College London 2011.

58. Smith A, Dixon A. *Healthcare Professionals' Views about Safety in Maternity Services*. London: The King's Fund; 2008.

59. Joyce R, Webb R, Peacock JL. Associations between perinatal interventions and hospital stillbirth rates and neonatal mortality. *Arch Dis Child Fetal Neonatal Ed* 2004;**89**:F51–6. http://dx.doi.org/10.1136/fn.89.1.F51

60. Gerova V, Griffiths P, Jones S, Bick D. *The Association between Midwifery Staffing and Outcomes in Maternity Units in England: Observational Study Using Routinely Collected Data. Preliminary Report and Feasibility Assessment*. London: National Nursing Research Unit; 2010.

61. Ball J, Bennett B, Washbrook M, Webster F. Birthrate Plus programme: a basis for staffing standards? *Br J Midwifery* 2003;**11**:264–6. http://dx.doi.org/10.12968/bjom.2003.11.5.11218

62. Goryakin Y, Griffiths P, Maben J. Economic evaluation of nurse staffing and nurse substitution in health care: a scoping review. *Int J Nurs Stud* 2011;**48**:510–12. http://dx.doi.org/10.1016/j.ijnurstu.2010.07.018

63. Hicks JR. *The Theory of Wages*. London: Macmillan; 1932.

64. Hicks JR. Elasticities of substitution again: substitutes and complements. *Oxford Econ Papers* 1970;**22**:289–96.

65. Thurston NK, Libby AM. A production for physician services revisited. *Rev Econ Stat* 2002;**84**:184–91. http://dx.doi.org/10.1162/003465302317332017

66. Øvretveit J. *Does Improving Quality Save Money: A Review of Evidence of Which Improvements to Quality Reduce Costs to Health Service Providers*. London: The Health Foundation; 2009.

67. Laudicella M, Olsen KR, Street A. Examining cost variation across hospital departments: a two-stage multi-level approach using patient-level data. *Soc Sci Med* 2010;**71**:1872–81. http://dx.doi.org/10.1016/j.socscimed.2010.06.049

68. Gaughan J, Mason A, Street A, Ward P, Centre for Health Economics. *English Hospitals Can Improve Their Use of Resources: An Analysis of Costs and Length of Stay for Ten Treatments*. York: Centre for Health Economics; 2012.

69. Bellanger MM, Or Z. What can we learn from a cross-country comparison of the costs of child delivery? *Health Econ* 2008;**17**:S47–57. http://dx.doi.org/10.1002/hec.1325

70. Sandall J, Soltani H, Gates S, Shennan A, Devane D. Midwife-led continuity models versus other models of care for childbearing women. *Cochrane Database Syst Rev* 2013;**8**:CD004667. http://dx.doi.org/10.1002/14651858.CD004667.pub3

71. Devane D, Brennan M, Begley C, Clarke M, Walsh D, Sandall J, *et al. Socio-economic Value of the Midwife: A systematic review, meta-analysis, meta-synthesis and economic analysis of midwife-led models of care*. London: Royal College of Midwives; 2010.

72. Ryan P, Revill P, Devane D, Normand C. An assessment of the cost-effectiveness of midwife-led care in the United Kingdom. *Midwifery* 2013;**29**:368–76. http://dx.doi.org/10.1016/j.midw.2012.02.005

73. Schroeder E, Petrou S, Patel N, Hollowell J, Puddicombe D, Redshaw M, *et al.* Cost effectiveness of alternative planned places of birth in woman at low risk of complications: evidence from the Birthplace in England national prospective cohort study. *BMJ* 2012;**344**:e2292. http://dx.doi.org/10.1136/bmj.e2292

74. Dodwell M, Gibson R. *BirthChoiceUK*. London: BirthChoiceUK; 2001. URL: www.BirthChoiceUK.com (accessed 17 September 2014).

75. Which? *Which? Birth Choice*. London: Consumers' Association; 2014. URL: www.which.co.uk/birth-choice (accessed 17 September 2014).

76. HM Government. *NHS Reference Costs 2010/11 CSV files*. 2013. URL: http://data.gov.uk/dataset/nhs-reference-costs-2010-11/resource/3f62aa69-4d8b-421a-bb6e-f26a4dc199f2 (accessed 5 September 2014).

77. Murray J, Saxena S, Modi N, Majeed A, Aylin P, Bottle A. Quality of routine hospital birth records and the feasibility of their use for creating birth cohorts. *J Public Health* 2012;**35**:298–307. http://dx.doi.org/10.1093/pubmed/fds077

78. Knight HE, Gurol-Urganci I, Mahmood TA, Templeton A, Richmond D, van der Meulen JH, *et al.* Evaluating maternity care using national administrative health datasets: how are statistics affected by the quality of data on method of delivery? *BMC Health Serv Res* 2013;**13**:200. http://dx.doi.org/10.1186/1472-6963-13-200

79. Sinha S, Peach G, Poloniecki JD, Thompson MM, Holt PJ. Studies using English administrative data (Hospital Episode Statistics) to assess health-care outcomes: systematic review and recommendations for reporting. *Eur J Public Health* 2013;**23**:86–92. http://dx.doi.org/10.1093/eurpub/cks046

80. NICE. *Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth*. NICE Clinical Guideline 55. 2007. URL: www.nice.org.uk/guidance/cg55/resources/guidance-intrapartum-care-pdf (accessed 10 September 2014).

81. Care Quality Commission. *Survey of Women's Experiences of Maternity Care 2010*. London: Care Quality Commission; 2011. URL: www.cqc.org.uk/surveys/maternity (accessed 8 October 2014).

82. Office for National Statistics. *Framework for National Statistics*. Cardiff: Office for National Statistics; 2000.

83. Department of Health. *The Operating Framework for the NHS in England 2012/13*. London: Department of Health; 2011.

84. Australian Council on Healthcare Standards (ACHS). *Australasian Clinical Indicator Report 2005–2012*. 14th edn. Ultimo, New South Wales: ACHS; 2013.

85. Department of Communities and Local Government. *The English Indices of Deprivation 2010, Neighbourhood Statistics Release*. 2011. URL: http://webarchive.nationalarchives.gov.uk/20120919132719/http://www.communities.gov.uk/documents/statistics/pdf/1871208.pdf (accessed 5 September 2014).

86. Birthplace in England Collaborative Group. Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study. *BMJ* 2011;**343**:d7400. http://dx.doi.org/10.1136/bmj.d7400

87. HES/NHS Information Centre. *Data Dictionary: Inpatients*. The NHS Information Centre for Health and Social Care. 2010. URL: www.hscic.gov.uk/media/1358/HES-Hospital-Episode-Statistics-HES-Admitted-Patient-Data-Dictionary/pdf/HES_Inpatients_DD_Sept10.pdf (accessed 10 September 2014).

88. Association of University Hospital Trusts. *List of AUKUH Members*. URL: www.aukuh.org.uk/index.php/members/aukuh-members (accessed 10 September 2014).

89. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;**24**:1185–202. http://dx.doi.org/10.1002/sim.1970

90. Reinhardt U. A production function for physician services. *Rev Econ Stat* 1972;**54**:55–66. http://dx.doi.org/10.2307/1927495

91. Berndt E, Christensen L. The translog function and the substitution of equipment, structures and labour in U.S. manufacturing. *J Econ* 1973;**1**:81–114. http://dx.doi.org/10.1016/0304-4076(73)90007-9

92. Diewert WE. An application of the Shephard duality theorem: a generalized Leontif production function. *J Polit Econ* 1971;**79**:507. http://dx.doi.org/10.1086/259764

93. Sato R, Koizumi T. On the elasticities of substitution and complementarity. *Oxford Econ Papers* 1973;**25**:44–56.

94. Syrquin M, Hollender G. Elasticities of substitution and complementarity: the general case. *Oxford Econ Papers* 1982;**34**:515–19.

95. Gibson R, Dodwell M. *A Preliminary Investigation into Midwifery Staffing, Maternity Service Costs and Maternity Outcomes in English NHS Trusts 2010/11: A Report by BirthChoiceUK for the Royal College of Midwives*. London: Royal College of Midwives; 2013.

96. Kitzinger S, Walters R. *Some Women's Experiences of Episiotomy*. London: National Childbirth Trust; 1993.

97. HM Government. *NHS Trust Reference Cost Schedules 2010–11*. 2013. URL: http://data.gov.uk/dataset/nhs-reference-costs-2010-11/resource/2e929335-bfe2-4257-b5d8-69bafad74821 (accessed 3 October 2014).

98. CQC. *Maternity Survey: Core Questionnaire (Core Questionnaire for the NHS Maternity Survey 2010)*. CQC; 2010. URL: www.nhssurveys.org/survey/866 (accessed 5 September 2014).

99. National Centre for Health Outcomes Development. *Hospital Episode Statistics (HES) – Construction of Continuous Inpatient (CIP) Spells and Assessment of Data Quality Annex 4 of National Centre for Health Outcomes Development (NCHOD) (2009) Compendium of Health Outcomes User Guide. 2009*. National Centre for Health Outcomes Development. URL: https://indicators.ic.nhs.uk/download/Additional%20Reading/Methods%20annexes/Compendium%20User%20Guide%202009%20(October)%20Annex%204%20V1.doc (accessed 12 September 2014).

Appendix 1 Advisory group membership

Member	Organisation
Carmel Bagness	Royal College of Nursing
David Richmond, Alana Cameron, Hannah Knight	RCOG
Susanne Tyler	NHS South Central (commissioner for maternity care)
Alison Macfarlane	City University
Rachel Plachcinski	National Childbirth Trust
Helen Duncan, David Knox	ChiMat
Jennifer Hollowell, Maggie Redshaw	National Perinatal Epidemiology Unit
Ann Farenden	Care Quality Commission
Chris Owen, Sarah Faulkner	Centre for Workforce Intelligence
Mike Wee	Poole Hospital/Bournemouth University
Sue Eardley	RCPCH
Chris Graham	Picker Institute
Marie Washbrook	BR+
Andrew Canter	National Maternity Support Foundation
Janet Scott, Charlotte Bevan	SANDS
Dharmintra Pasupathy (Chair)	King's College London
Debby Gould	NHS QUEST, Salford Royal NHS Foundation Trust

Appendix 2 Data sources, cleaning and derivation of indicators

Hospital Episode Statistics data

The data set available to the study consisted of:

- HES inpatient records 2000/01 to 2010/11
- HES outpatient records 2003/04 to 2010/11
- HES A&E records 2007/08 to 2010/11
- ONS-HES linked mortality statistics (including neonatal deaths) 2000/01 to 2010/11.

Further details of these data sets are provided below.

Hospital Episode Statistics inpatient records 2000/01 to 2010/11

This data set included 184 million inpatient episodes, of which:

- 7.2 million were delivery episodes
- 7.3 million were birth episodes
- 15 million were mothers' non-delivery episodes
- 6 million were babies' non-birth episodes.

Every inpatient record consisted of a general section which was applicable to all inpatient episodes. Inpatient records with the HES field EPITYPE = 2 or 5 were considered to be a record of a mother delivering a baby during that hospital episode. These also had a 'maternity tail' where specific maternity-related information was recorded, and nine 'baby tails' that recorded details on the status of each baby born (up to a total of nine babies) during the delivery episode. In addition each baby had his or her own inpatient birth record, being of EPITYPE 3 or 6, but this was not linked to the mother's record.

Outpatient 2003/04 to 2010/11

The data set included 610 million outpatient episodes, of which:

- 70 million related to mothers, but were not necessarily maternity related
- 24 million related to children who had been born since 2000.

Clinical information was very sparse in this data set and thus it was not deemed an adequate source of information on which to base maternity indicators. It was used to help determine a woman's ethnicity when this was missing in the inpatient records.

Accident and emergency records 2007/08 to 2010/11

Although the quality of the information in this data set is rapidly improving, for the available years it was not sufficiently reliable or complete to use.

Office for National Statistics–Hospital Episode Statistics linked mortality statistics (including neonatal deaths) 2000/01 to 2010/11

These data derived from death certification give the cause of death of an individual. The record is linked to the HES data set, enabling the individual's previous medical history to be examined.

Office for National Statistics birth registrations by communal establishment code 2000/01 to 2010/11

The number of deliveries by communal establishment code for 2010/11 was obtained from the ONS under Open Government Licence v1.0. These data were matched with communal establishment place using a list obtained from NHS Connecting for Health. Communal establishment place was matched to NHS trust (or PCT) using the BirthChoiceUK database. This data set was used to establish the extent of duplicate records in HES.

NHS reference costs

These are available under Open Government Licence v2.0.[76]

Currency description

This gives the HRG activity being undertaken, for example NZ11A – Normal Delivery with Complication or Comorbidity (CC).

Department description

A delivery episode that lasts longer than usual generates two cost line items, the second one indicating an extended stay. These have to be accounted for so that the number of deliveries can be correctly calculated. The line items corresponding to excess bed-days can be identified by the entry in the department code (= EI_XS OR NEI_L_XS OR NEI_S_XS)

Deliveries and delivery cost

The number of deliveries is given by the sum of the activity column in the reference costs where the currency description indicates a delivery and excess bed-days entries are excluded by filtering on the department code. The delivery cost is computed by summing the actual cost column.

Currency code	Currency description
NZ11A	Normal Delivery with CC
NZ11B	Normal Delivery without CC
NZ11C	Normal Delivery with Epidural with CC
NZ11D	Normal Delivery with Epidural without CC
NZ11E	Normal Delivery with Induction with CC
NZ11F	Normal Delivery with Induction without CC
NZ11G	Normal Delivery with Post-partum Surgical Intervention
NZ12A	Assisted Delivery with CC
NZ12B	Assisted Delivery without CC
NZ12C	Assisted Delivery with Epidural with CC
NZ12D	Assisted Delivery with Epidural without CC
NZ12E	Assisted Delivery with Induction with CC
NZ12F	Assisted Delivery with Induction without CC
NZ12G	Assisted Delivery with Post-partum Surgical Intervention
NZ13Z	Planned Lower Uterine Caesarean Section
NZ14Z	Emergency or Upper Uterine Caesarean Section
NZ15Z	Caesarean Section with Eclampsia, Pre-eclampsia or Placenta Praevia

Reproduced from *NHS Trust Reference Cost Schedules 2010–11*.[97] Contains public sector information licensed under the Open Government Licence v2.0. URL: www.nationalarchives.gov.uk/doc/open-government-licence/version/2/.

Antenatal costs

Currency code	Currency description
CMANV	Ante-Natal Visits
NZ04C	Ante-natal or Post-natal Observation age between 16 and 40 years with length of stay 0 days
NZ04D	Ante-natal or Post-natal Observation age < 16 or > 40 years with length of stay 0 days
NZ05C	Ante-natal or Post-natal Investigation age between 16 and 40 years with length of stay 0 days
NZ05D	Ante-natal or Post-natal Investigation age < 16 or > 40 years with length of stay 0 days
NZ06Z	Ante-natal or Post-natal Full Investigation with length of stay 0 days
NZ07C	Ante-natal or Post-natal Observation age between 16 and 40 years with length of stay 1 day or more
NZ07D	Ante-natal or Post-natal Observation age < 16 or > 40 years with length of stay 1 day or more
NZ08C	Ante-natal or Post-natal Investigation age between 16 and 40 years with length of stay 1 day or more
NZ08D	Ante-natal or Post-natal Investigation age < 16 or > 40 years with length of stay 1 day or more
NZ09Z	Ante-natal or Post-natal Full Investigation with length of stay 1 day or more

Reproduced from *NHS Trust Reference Cost Schedules 2010–11*.[97] Contains public sector information licensed under the Open Government Licence v2.0. URL: www.nationalarchives.gov.uk/doc/open-government-licence/version/2/.

In 2010/11 Currency Codes NZ0[4, 5, 6, 7, 8, C, D, Z] cover antenatal or postnatal stays of 0 or 1+ days. Although this does not separate ante- and postnatal costs, a separate analysis in 2011/12 shows only 2.5% of the activity is postnatal so it is assumed that effectively all the recorded costs are antenatal.

In 2011/12 Currency Codes NZ[16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27]Z are taken to represent antenatal costs, even though this does contain a small element of postnatal costs.

There are a number of face-to-face attendances and non-face-to-face attendances (Currency Code WF0 [1, 2, A, B, C, D]). The service codes 501 and 560* identify these as either obstetric of midwife episodes. Although a cost, they cannot be identified with either ante- or postnatal periods.

Postnatal cost
Postnatal costs are made up of excess bed-days and postnatal visits. Cost of excess bed-days is defined as

Currency Code = NZ1[1,2,3,4,5][A,B,C,D,E,F,G,Z] as above AND (Department Code = EI_XS OR NEI_L_XS OR NEI_S_XS)

and cost of postnatal visits as

Currency Code = CMPNV.

Total maternity expenditure
In calculating the total maternity expenditure all the above line items are added. The following items are not included:

- Paediatric Critical Care (XB*)
- Mother and Baby Units 11/12(MHCOM08, MHIPMB); 10/11(MHCS*, MHCT9) – not in units with maternity units
- Neonatal critical care (XA*).

Care Quality Commission survey questions[81,98]

Questions reproduced with permission from CQC.[81,98]

Antenatal care

B5. Were you given a choice of having your baby at home? (Those women who had a choice about where to have their baby, i.e. Selected option 1 to B4)

B15. Was the reason for this scan clearly explained to you? (ALL)

B17. Were the reasons for having a screening test for Down's syndrome clearly explained to you? (ALL)

B19. Was the reason for this scan clearly explained to you? (ALL)

Labour and birth

C2. During your labour, were you able to move around and choose the position that made you most comfortable? (ALL)

C4. During your labour and birth, did you feel you got the pain relief you wanted? (ALL)

C9. If you had an episiotomy (cut) or tear requiring stitches, how long after your baby was born were the stitches done? (Those women who had a normal or assisted vaginal delivery, i.e. Selected option 1 or 2 to C6)

C10. Did you have skin to skin contact (baby naked, directly on your chest or tummy) with your baby shortly after the birth? (ALL)

Staff during labour and birth

C12. Did you have confidence and trust in the staff caring for you during your birth and labour? (ALL)

C13. If you had a partner or a companion with you during your labour and delivery, were they made welcome by the staff? (ALL)

C14. Were you (and/or your partner or a companion) left alone by midwives or doctors at a time when it worried you? (ALL)

C15. Thinking about your care during labour and birth, were you spoken to in a way you could understand? (ALL)

C16. Thinking about your care during labour and birth, were you involved enough in decisions about your care? (ALL)

C17. Overall, how would you rate the care received during your labour and birth? (ALL)

Care in hospital after the birth

D2. Looking back, do you feel that the length of your stay in hospital after the birth was . . . (Those women who went to hospital and did not have a home birth)

D3. Thinking about the care you received in hospital after the birth of your baby, were you given the information or explanations you needed? (Those women who went to hospital and did not have a home birth)

D4. Thinking about the care you received in hospital after the birth of your baby, were you treated with kindness and understanding? (Those women who went to hospital and did not have a home birth)

Feeding the baby during the first few days

E4. (Thinking about feeding your baby in the first few days after the birth) Did you feel that midwives and other carers gave you consistent advice? (ALL)

E5. (Thinking about feeding your baby in the first few days after the birth) Did you feel that midwives and other carers gave you active support and encouragement? (ALL)

Hospital Episode Statistics data cleaning

The information which follows provides an account of full data cleaning conducted for the study, prior to data analysis. The project restricted the records to NHS hospital deliveries resulting in a registrable birth. Duplicate delivery records were removed from the mother's records.

The babies' birth records also contained duplicate records. These duplicates were not removed, as the majority of the project's work concentrated on the mother's delivery record and the resources were not available to clean both.

Gestational age
HES field: GESTAT.

Because gestational age was used to help identify delivery episodes that do not result in a registrable birth, it was the first field to be cleaned.

Some trusts submit the gestational age to HES in days, rather than in weeks as required. The HES system truncates the last digit, leading to full-term births having an apparent gestational age of around 27 weeks. These cases were identified as follows.

For a trust with >200 deliveries a year where the gestational age is known *and* the percentage of these records with gestational age ≤30 weeks exceeds 10% in any given year, all records in that trust have their gestational age set to 'unknown' if the birthweight is not known to be ≤1kg.

For all trusts the following additional cleaning rule was also imposed: where the gestational age is between 10 and 20 weeks and the birthweight exceeds 1 kg, the gestational age is set to 'unknown', as this combination of gestational age and birthweight is unlikely.

Removal of records associated with abortions and miscarriages
A record was marked as not a registrable delivery if any of the following applies:

- delivery was at <21 weeks (cleaned) gestational age
- gestational age <24 weeks and a stillbirth is recorded (BIRSTAT_1 equal 2, 3 or 4)
- the OPCS-4 code for induced abortion (Q14) was found in the procedure codes
- an ICD-10 code for miscarriage or abortion (O00 to O08) was found in the diagnostic codes.

Multiple records
Some delivery events have many delivery records associated with them that are not necessarily identical. Out of the many alternatives, one record has to be chosen to represent the delivery. This was achieved using the following procedure:

All of a mother's registrable delivery records were put into chronological order. If two successive records were separated by >30 weeks then they were considered to relate to different delivery episodes. Early miscarriages would have already been removed by the procedure to clean the registrable birth data (discussed earlier) so the genuine occurrence of two deliveries within 30 weeks should be highly improbable.

If two successive records were separated by <30 weeks and baby sex, baby weight or gestational age was recorded in both records, and they were different, then the two records were considered to relate to different delivery episodes and thus both were kept.

For the group of records that all relate to one delivery, a single record was chosen by scoring the information in the record (see below) and choosing the record with the highest score. When two or more records had the same high score the most recent one was chosen.

All the multiple records relating to a single delivery which were not chosen were flagged as invalid records and excluded from any future analysis. A delivery resulting in more than one baby generates only a single delivery record so this cleaning procedure does not adversely affect the recording of multiple births.
A scoring system was used to select record with the greatest amount of relevant data to the project.

Data item	Score
DELMETH_1 ('delivery method') corresponds to known delivery type	10
HRG code corresponds to a delivery	7
Method of onset of labour known	5
Method of anaesthetic known	4
Start of episode known	1
Gestational age known (after cleaning)	1
Birthweight known	1
Non-trivial procedural codes	2
Non-trivial diagnostic codes	2

Place of delivery

HES Field: DELPLAC_1.

In HES data, there is a separate field for place of birth for each baby in the maternity tail of the mother's delivery record. Although multiples are occasionally born in different locations, this rare event was ignored and the place of delivery of the first recorded baby in HES was used.

From a comparison with ONS birth registration data, it is known that home births are under-recorded in HES, with many trusts recording no home births. However there are a few trusts that record very many home births. ONS birth registration data are available only by local authority and are generally believed to be accurate. Nationally the ONS recorded home birth rate is <2.5%, with local peaks never exceeding 10%, so a trust rate in excess of 40% is clearly not believable.

Data cleaning was undertaken using the following steps: first, the proportion of deliveries in each trust that did not take place in a NHS hospital was calculated for each year. If a trust had >200 delivery records and the percentage of records with a non-NHS hospital delivery exceeded 15%, then the place of delivery was set to unknown for all records in the trust for the year. The maximum home birth rate for a local authority is about 10%, so it is unlikely that >15% would not take place in a NHS hospital setting.

A NHS hospital birth was taken to be DELPLAC_1 being any one of the following {0, 2, 3, 4, 8, 9}.

0 = in NHS hospital: delivery facilities associated with midwife ward.

1 = at a domestic address.

2 = in NHS hospital: delivery facilities associated with consultant ward.

3 = in NHS hospital: delivery facilities associated with general medical practitioners ward.

4 = in NHS hospital: delivery facilities associated with consultant, general medical practitioners or midwife ward, or any combination of two of these.

5 = in private hospital.

6 = in other hospital or institution.

7 = in NHS hospital: ward or unit without delivery facilities.

8 = other than those above.

9 = not known.

Parity

HES Field: NUMPREG.

Parity was frequently not reported. Even when it was, some trusts recorded it incorrectly. For instance, some trusts reported that nearly all their mothers were first-time mothers, or nearly all second-time mothers. Any trust that had a percentage of nulliparous women outside the range 20–70% had all their parity data set to 'unknown' for that year.

For all women in the data set the following algorithm was applied to identify and correct inconsistent parity, or to provide a value when it was unknown.

The women's ICD-10 codes were then searched for evidence of a previous delivery (e.g. a previous caesarean). The following codes were used as indicators of a previous delivery episode: O34.2, O75.7, Z35.4, Z87.6 and Z87.5. When found, the parity of the woman was assumed to be one if there was no other evidence to suggest otherwise. The actual parity may have been higher, but as the largest difference in outcomes is between nulliparous and primiparous women this is the most important distinction to make.

Each woman's delivery history was then examined to estimate her parity. If the parity for any of her deliveries was given either by the cleaned HES field NUMPREG or from the examination of ICD-10 codes, then this value was counted forward or back as appropriate. If this resulted in a number which was less than the number of deliveries observed to date in HES, it was overwritten with the higher number.

The final value of parity derived was consistent in that it increased in accordance with the chronological order of a woman's delivery history. Jumps still occurred, but this is natural, as some woman will have given birth to subsequent babies out of area. It is also consistent in the sense that there were no mothers recorded in the current year as nulliparous but who had had a previous caesarean.

Ethnicity

HES changed the way ethnicity is recorded in 2000/01. The older codes were mapped onto the newer ethnicity categories.

If a woman's ethnicity was recorded as either 'Unknown' or 'Not Stated' then both her inpatient and outpatient history were searched to see if a valid ethnicity could be found. This was successful 60% of the time.

No attempt was made to make ethnicity consistent when it varied across a woman's history, so a woman's recorded ethnicity can change between deliveries. Her ethnicity recorded in 2010/11 was used for the subsequent analyses.

Mother's age

Mother's age was computed using the mother's date of birth and the episode start date. Some records had maternal age miscoded. Women who fell into the following age ranges (see below), and were reported as having given birth, had their age set to 'unknown', as it was assumed that the recorded age was a miscoding of their date of birth.

Age 60–120 years: there is a small number of women in this age range. It is likely that some of the women in the 50–59 range are also miscodings but these were left unchanged.

Age > 400 years: some trusts use a default date in the sixteenth century which was used to calculate a woman's age. In earlier years some trusts had a high proportion of their women in this category.

Age 0 years: this is an obvious miscoding of entering the current year rather than actual year of date of birth. There were several hundred of these every year.

Age 1–11 years: there were a few women in this range each year, which were assumed to be miscodings. From age 12 years onwards the number of births per year increased rapidly so these were assumed to be real.

Health Reference Group codes in inpatient non-delivery records

Some 10,000 non-delivery inpatient HES records had an HRG code associated with a delivery. Examination of these records indicated that, although the episode was often associated with a maternity issue, an actual delivery was unlikely to have taken place. These were ignored.

Baby's weight, mother's record

In HES records, the baby's weight is recorded in grams. A value of 9999 is used to represent unknown. A value of 7000 is used to represent the weight of babies that exceed 7 kg.

In 2010/11 only 24 babies were in the range 6–7 kg while 189 were recorded as 7 + kg. This suggested miscoding, so all those babies with weight in excess of 7 kg had their weight set to 'unknown'.

For live births any baby with a weight of < 450 g had his or her weight set to 'unknown'. For stillbirths the lower limit is taken as 250 g.

Plurality

HES field: NUMBABY.

In 2001/02 the HES field NUMBABY indicated > 40,000 sextuplets, with excess numbers also recorded in neighbouring years. Values of 6 for NUMBABY were thus ignored throughout the data set. Multiplicity is also recorded in the ICD-10 codes, so the loss of information is minimal. As the project focused on recent years, this particular miscoding of NUMBABY is of no consequence; however, the practice of recording six maternity tails persists.

The mother's ICD-10 codes were queried, looking for the codes Z37 and Z38 as an indication of multiplicity.

A delivery was declared a singleton if there was no evidence of a multiple delivery and the plurality was known from either NUMBABY or the ICD-10 codes.

Anaesthetics

There are three sources of anaesthetic data:

1. DELPREAN: a maternity tail field that records anaesthetics used before delivery
2. DELPOSAN: a maternity tail field that records anaesthetics used after delivery
3. OPCS-4: a procedure code that records anaesthetics used during a delivery episode.

DELPREAN and DELPOSAN codes

These HES fields take on the following values.

1 = general anaesthetic: the administration by a doctor of an agent to produce unconsciousness.

2 = epidural or caudal anaesthetic: the injection of a local anaesthetic into the epidural space.

3 = spinal anaesthetic: the injection of a local anaesthetic agent into the subarachnoid space.

4 = general anaesthetic and epidural or caudal anaesthetic.

5 = general anaesthetic and spinal anaesthetic.

6 = epidural or caudal, and spinal anaesthetic.

7 = other than 1 to 6.

8 = not applicable.

9 = not known.

C = caesarean performed but no mention of anaesthetic.

The category 'caesarean without anaesthetic' is not an official HES designation but was introduced here to capture the records where a woman would have had an anaesthetic but the HES record did not record it.

There is no separate code for a combined spinal–epidural increasingly used in practice. For simplicity, epidurals and spinals were combined into a category called regional anaesthetic. A woman who had a caesarean without any mention of anaesthetic use was assumed to have had a regional although a few will have had a general anaesthetic.

DELPREAN/DELPOSAN	Type	Anaesthetic used
1	General	Yes
2	Regional	Yes
3	Regional	Yes
4	General and regional	Yes
5	General and regional	Yes
6	Regional	Yes
7	Other/not used	
8	Other/not used	
9	Unknown	
C	Regional allocated	Yes

Office of Population Censuses and Surveys anaesthetic

The following OPCS-4 codes indicate anaesthetic use. A woman may have several such codes, including both a general and a regional.

Code	Description	Type	Anaesthetic used
Y80.1	Inhalation anaesthetic using muscle relaxant	General	Yes
Y80.2	Inhalation anaesthetic using endotracheal intubation NEC	General	Yes
Y80.3	Inhalation anaesthetic NEC	General	Yes
Y80.4	Intravenous anaesthetic NEC	General	Yes
Y80.5	Rapid sequence induction of anaesthetic	General	Yes
Y80.8	Other specified general anaesthetic	General	Yes
Y80.9	Unspecified general anaesthetic	General	Yes
Y81.1	Epidural anaesthetic using lumbar approach	Regional	Yes
Y81.2	Epidural anaesthetic using sacral approach	Regional	Yes
Y81.8	Other specified spinal anaesthetic	Regional	Yes
Y81.9	Unspecified spinal anaesthetic	Regional	Yes
Y82.1	Local anaesthetic nerve block	Other/not used	
Y82.2	Injection of local anaesthetic NEC	Other/not used	
Y82.3	Application of local anaesthetic NEC	Other/not used	
Y82.8	Other specified local anaesthetic	Other/not used	
Y82.9	Unspecified local anaesthetic	Other/not used	
Y84.1	Gas and air analgesia in labour	Other/not used	
Y84.2	Sedation NEC	Other/not used	
Y84.8	Other specified other anaesthetic	Other/not used	
Y84.9	Unspecified other anaesthetic	Other/not used	
Y90.1	Application of transcutaneous electrical nerve stimulator	Other/not used	
	No mention	Other/not used	

NEC, not elsewhere classifiable.

Some trusts recorded anaesthetic uses that were implausible, for example giving all women a general anaesthetic. The following cleaning rules were thus used on each source of anaesthetic separately:

All the trust's anaesthetic data for a given source of information were set to 'unknown' if:

- DELPREAN: anaesthetic use exceeds 70% or general anaesthetic use exceeds 10%
- OPCS-4 coded: anaesthetic use exceeds 70% or general anaesthetic use exceeds 10%
- DELPOSAN: anaesthetic use exceeds 70% or general anaesthetic use exceeds 5%.

In practice, trusts with implausible data were captured by these rules and trusts with plausible data were well clear of the trigger points.

Because OPCS-4 codes do not distinguish between pre- and post-delivery anaesthetics, and some trusts appear to confuse the two when coding, data from these three sources were combined. If any of the

three sources mentioned a general anaesthetic, it was assumed for data-cleaning purposes that the woman had one; likewise regional anaesthetics.

After the three sources had been combined, the 'anaesthetic use exceeds 70% or general anaesthetic use exceeds 10%' rules was reapplied. This rule was triggered only once in 2010/11.

Records were flagged if the anaesthetic use had been assumed because the woman had a caesarean and DELPREAN did not mention the use of anaesthetics. If any of the other sources recorded anaesthetic use then this flag was cleared. Trusts that still had a high incidence of 'assumed anaesthetic use', but had a low anaesthetic use in vaginal deliveries, suggested a general under-reporting. As a consequence any trust with an 'assumed anaesthetic use' in excess of 5% of deliveries had all its anaesthetic data set to unknown.

If the assumption is made that those women who have an ICD-10 code indicating a complication arising from an anaesthetic represent a random sample of women who have actually received an anaesthetic, then capture–recapture analysis can be used to get an estimate of the coverage of these various methods of recording anaesthetic use.

Code	Estimated coverage (%)
DELPREAN + 'Assumed from caesarean'	86.8
DELPREAN and DELPOSAN + 'Assumed from caesarean'	88.0
DELPREAN and DELPOSAN and OPCS-4 + 'Assumed from caesarean'	96.8

Derivation of maternal and infant indicators used in the analysis

The indicators discussed here contributed to study outcomes. Further cleaning was undertaken on the following fields to ensure that each was as accurate as possible before the analysis took place.

Maternal indicators

Anaesthetic complication
This relates to any mention of ICD-10 code O74 (Complications of anaesthesia during labour and delivery) or O89 (Pulmonary complications of anaesthesia during the puerperium) where it was known that the woman was administered an anaesthetic (see above re cleaning of anaesthetic data). Absence of anaesthetic complications is part of the composite 'healthy mother' outcome.

Postnatal duration
The time between delivery of the baby and discharge as given by the HES field POSTDUR.

Antepartum stillbirth (derived from mother's delivery record)
HES field BIRSTAT=2 for any one of the mother's nine maternity tails. In calculating the stillbirth rate, women with an unknown birth status for the first baby tail (BIRSTAT_1=9) were excluded from the calculation.

Intrapartum stillbirth (derived from mother's delivery record)
HES field BIRSTAT=3 or 4 for any one of the mother's nine maternity tails. In calculating the rate, women with unknown birth status for the first maternity tail (BIRSTAT_1=9) were excluded from the calculation.

Stillbirth
Either an antepartum or intrapartum stillbirth as previously defined above.

Method of delivery
The OPCS-4 codes for each woman were searched for one of the codes R17–R25, which was then mapped onto a synthetic version of DELMETH which mirrored the definitions used in the HES field DELMETH. If the method of delivery could not be determined by this procedure the HES field DELMETH_1 was used.

Emergency caesarean
Derived from the method of delivery.

Elective caesarean
It was often the case that an elective caesarean was recorded in the method of onset of labour (DELONSET), yet the method of delivery recorded an emergency caesarean. So that elective caesareans were measured on the same basis as other methods of delivery, the method of delivery (rather than method of onset of labour) was used in the definition. Typically the elective rate as measured by method of delivery is one percentage point lower than the rate indicated by the method of onset.

Caesarean
Derived from the method of delivery. On rare occasions the HES field DELONST recorded that a woman had an elective caesarean, yet the method of delivery recorded a vaginal birth. In such cases it was assumed that the women did not have a caesarean.

Delivery with bodily integrity
Delivery without caesarean, second-, third- or fourth-degree tear, episiotomy, uterine damage or sutures, where the following definitions are used:

Uterine damage   Mention of any of the ICD-10 codes O71.1 to O71.7 inclusive (other obstetric trauma) or any one of the following OPCS-4 codes:

R30.1: repositioning of inverted delivered uterus.

R30.9: unspecified other operations on delivered uterus.

R32.1: repair of obstetric laceration of uterus or cervix uteri.

Second-/third-/fourth-degree tear   ICD-10 codes O70.1–O70.4.

Sutures   Repair of laceration mentioned in the procedure codes (R32.*).

Episiotomy   Any mention of the OPCS-4 codes R27.*.

Tears are recorded in ICD-10 diagnostic codes and repairs in OPCS-4 procedure codes. By referring to both types of codes, a more robust indicator was developed which was more tolerant of poor coding. As a consequence of this a women was not classified as having bodily integrity if she had either a first-degree tear that was stitched or a second-degree tear that was not stitched.

Induction
The HES field DELONSET taking any of the values 3, 4 or 5.

Maternal sepsis
Mention of an ICD-10 code associated with sepsis, but excluding pyrexia. ICD-10 codes O75.3, O85.* and O86.* excluding O86.4.

Mother's readmission within 28 days of discharge

Readmission within 28 days of the end of the delivery episode (Inpatient.EPISTART – Delivery.DISDATE <29). The delivery record must indicate that the woman had left hospital and was not transferred to a different ward or hospital (Delivery.DISDEST = 51 or 52). The inpatient record had to be the first episode in a spell to distinguish it from a transfer (Inpatient.EPIORDER = 1).

Spontaneous delivery

Vaginal delivery without the use of instruments.

Unassisted delivery

Spontaneous onset, spontaneous delivery and no mention of an episiotomy in the procedure codes.

Normal birth

Spontaneous labour onset, spontaneous delivery, no episiotomy and no mention of either pre- or postdelivery use of general or regional anaesthetic in DELPREAN, OPCS-4 codes or DELPOSAN.

This represents a slight tightening of the Maternity Care Working Party (MCWP)[17] definition of normal birth which excludes women from the normal birth category if they have had a predelivery anaesthetic but not if they have had a postdelivery anaesthetic.

The advantage of our more stringent definition is that it makes use of anaesthetic information contained in the OPCS-4 codes as well as in DELPREAN. This increases the accuracy of the anaesthetic data used (as DELPREAN is often incomplete) and increases the number of women in the normal birth denominator. The disadvantage is that it is different from the established MCWP definition. However, it is estimated that only a small percentage of women having a normal birth according to the MCWP definition would then go on to have complications requiring a postdelivery anaesthetic.

Previous caesarean

There are two ICD-10 codes that record previous caesarean, O75.7 (vaginal delivery following previous caesarean) and O34.2 (maternal care due to uterine scar from previous surgery). In addition to these two codes a woman's previous delivery history was searched, looking for evidence of a caesarean.

A caesarean was identified by searching the ICD-10 and OPCS-4 codes, the method of onset field DELONSET and the method of delivery field DELMETH. Once it was determined that a woman had had a previous caesarean, the information was propagated forwards through her other delivery records.

Vaginal birth after caesarean

A woman was said to have had a vaginal birth after caesarean if she had had a caesarean in any previous pregnancy and her method of delivery in 2010/11 explicitly stated a vaginal delivery.

Baby indicators

Derivation of baby's indicators

As the baby's record is not linked to the mother's record, only the limited information available in the mother's record's maternity tail can be associated with the mother.

Healthy birthweight

Birthweight was taken from the first baby tail Birweit_1 after cleaning. A weight in the range 2.5–4.5 kg was defined as a healthy weight.

Term gestational age

Gestational age was taken from the first baby tail Gestat_1 after cleaning. Term was taken to be 37 to 42 completed weeks of pregnancy.

Live baby
All babies born in the episode alive as defined by BIRSTAT in the mother's delivery record.

Healthy mother/healthy baby dyad indicator
Combined outcomes:

- delivery whilst retaining bodily integrity and without the use of instruments
- no mention of sepsis
- no mention of anaesthetic complication
- mother returned home in 2 days or less
- mother not readmitted within 28 days and
- all babies in delivery episode alive and born at term with healthy birthweights.

Indicators not used in the analysis
The following quality indicators have been suggested in a range of policy initiatives. We outline the rationale for not using the following in the analysis.

Smoking
There are a number of ICD-10 codes that indicate smoking (F16, Z72.0, Z71.6 and P04.2). There was a wide spectrum of smoking rates for trusts, with many reporting zero rates, and all other possibilities up to 25% were observed. It is evident that many trusts underreport but it is not possible to distinguish underreporting from genuine low rates.

First antenatal appointment by 12 weeks
This was taken from the HES field ANGEST. Even after allowing for those trusts that misreported gestational age there was a wide spectrum of rates from 0% to 90%.

Antenatal duration
The time between the admission of the woman to a NHS trust and the delivery of the baby, as given by the HES field ANTEDUR.

Elective caesarean before 37 weeks' gestation
An elective caesarean where the cleaned gestational age was <37 weeks.

Mother's transfer
Method 4 taken from the National Centre for Health Outcomes Development *Annex 4*[99] is used. In this procedure delivery records were linked to non-delivery inpatient records using the combined fields:

(PROCODET + PROVSPNO + ADMIDATE + EXTRACT_HESID).

There are two types of transfer: from general inpatients into a delivery episode and from a delivery episode to a general inpatient episode. The latter was distinguished by requiring that the inpatient HES field EPIORDER exceed that of the delivery, i.e. the inpatient episode followed the delivery episode. When examined by trust, the results were hard to interpret, quite possibly because many transfers were due to administrative procedures rather than clinical necessity.

Mother's length of stay
The duration of a woman's spell in hospital is the desirable measure (SPELDUR). On the rare occasions this was not available, the length of the episode (EPIDUR) was used instead. A value equating to more than a year was discounted to filter out the occasional miscoding which could contribute some excessively long stays running to decades. The indicator was restricted to women where there was evidence that they returned home after discharge (DISDEST = 19 or 29).

Blood transfusion

Blood transfusions are indicated by the OPCS-4 code X33. Most trusts reported a near zero rate of blood transfusion; a few reported around 0.5%, which is perhaps more realistic.

Postpartum haemorrhage

These are identified by ICD-10 codes O72.0 and O72.1, neither of which gives any indication of the volume of blood loss. The range of rates for trusts was from 3% to 25%. It was considered that postpartum haemorrhage is under-reported.

Induction or acceleration

OPCS-4 codes R14, R15 and X35.1 indicate procedures associated with induction or acceleration. However, it was not possible to determine whether the procedure was performed before or after the onset of labour.

Caesarean hysterectomy

The numbers identified by the OPCS-4 code R25.1 were very small, making this an unsuitable indicator.

Ulcer

Identified by ICD-10 codes L89 and L97. The numbers recorded were too small to make an adequate indicator.

Foreign body left during procedure

These can be identified by both ICD-10 and OPCS-4 codes; however, only a handful of records had any mention of these codes. Possibly the subsequent removal took place in an outpatient episode; these were very poorly reported.

Delivery-induced faecal incontinence

This can be identified by the ICD-10 diagnostic code R15 and a collection of OPCS-4 codes indicating a repair when found in the general inpatients HES records following a delivery. The lag between the delivery and the subsequent repair can run to years and hence was not a suitable indicator for the study, which is concerned with recent data.

Maternal mortality within 42 days

This is determined using ONS death registration records linked to HES. The numbers recorded were very low, 69 for England in 2010/11, some of which would be unrelated to maternity. The small numbers make this an unsuitable indicator.

Maternal mortality between 43 days and 1 year

Again the small numbers make this an unsuitable candidate for an indicator.

Appendix 3 Funnel plots for the 10 indicators used in analysis
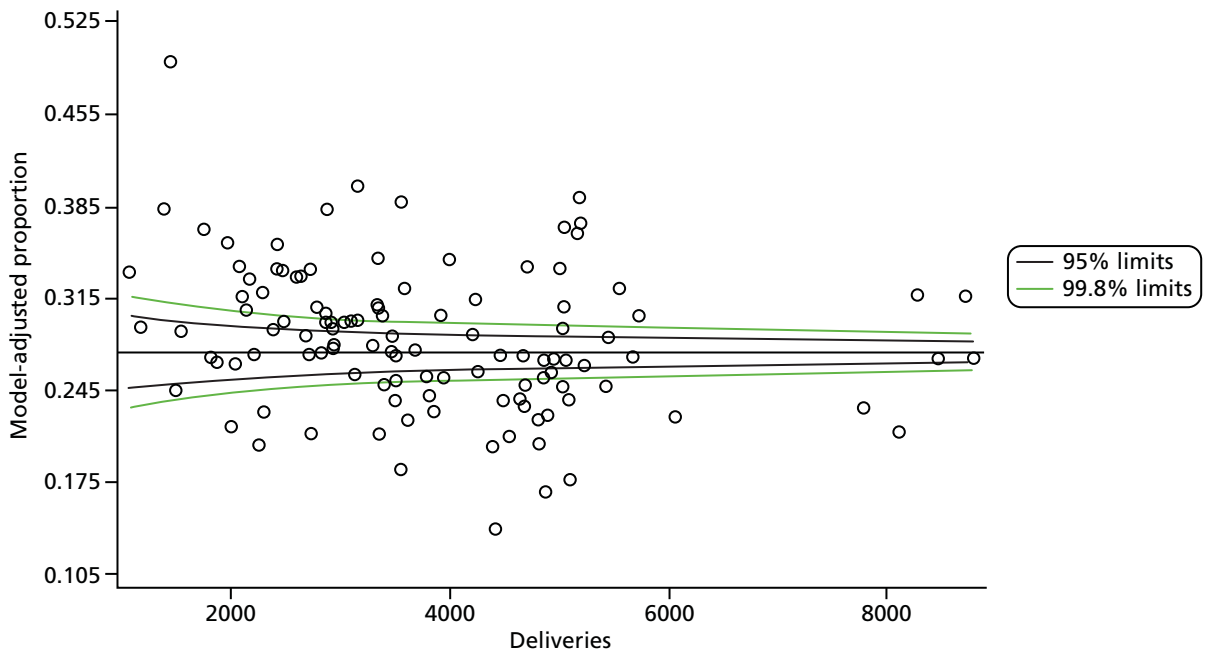
FIGURE 2 Funnel plot: healthy mother intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
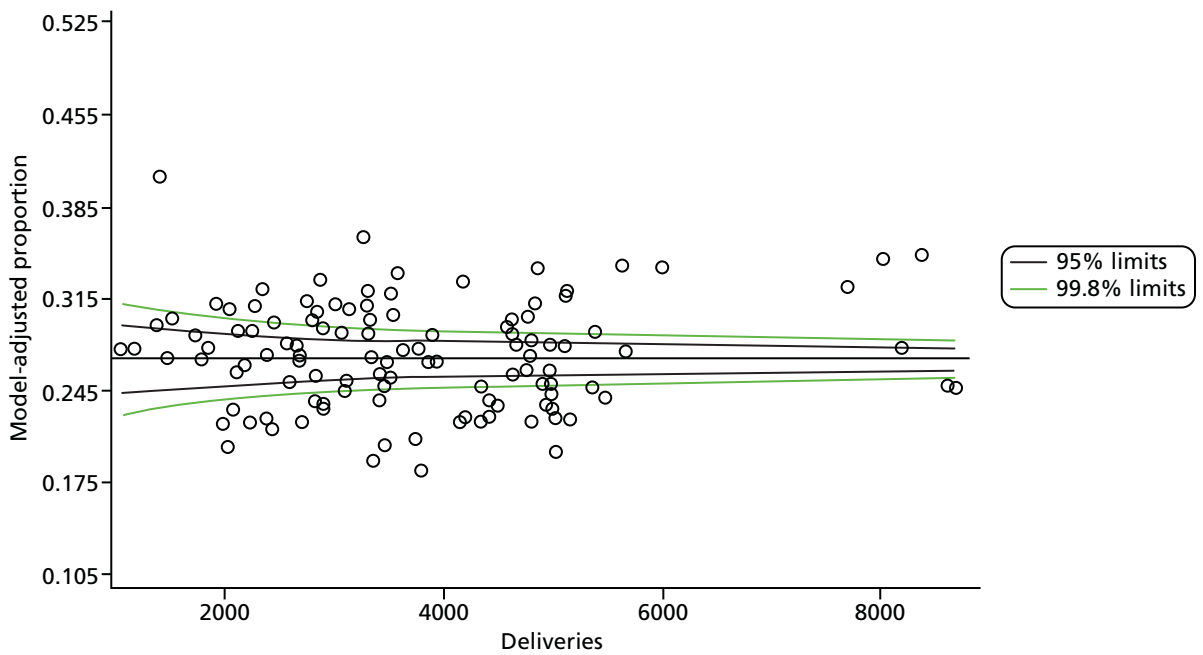
FIGURE 3 Funnel plot: healthy mother full model.
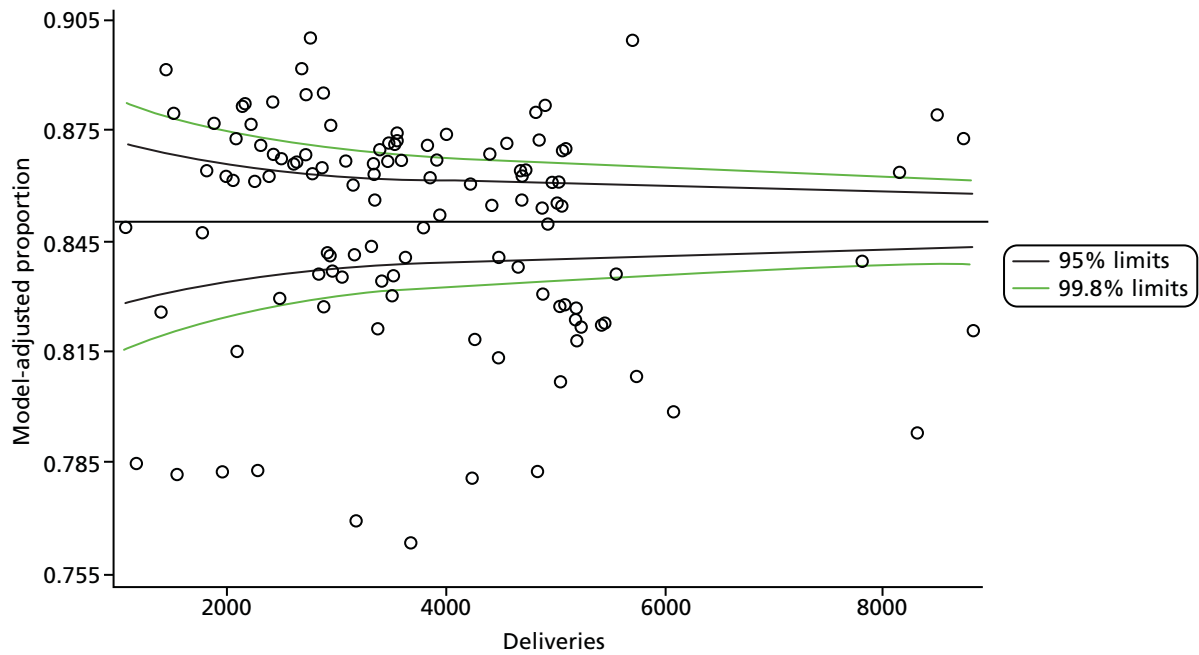
FIGURE 4 Funnel plot: healthy baby intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
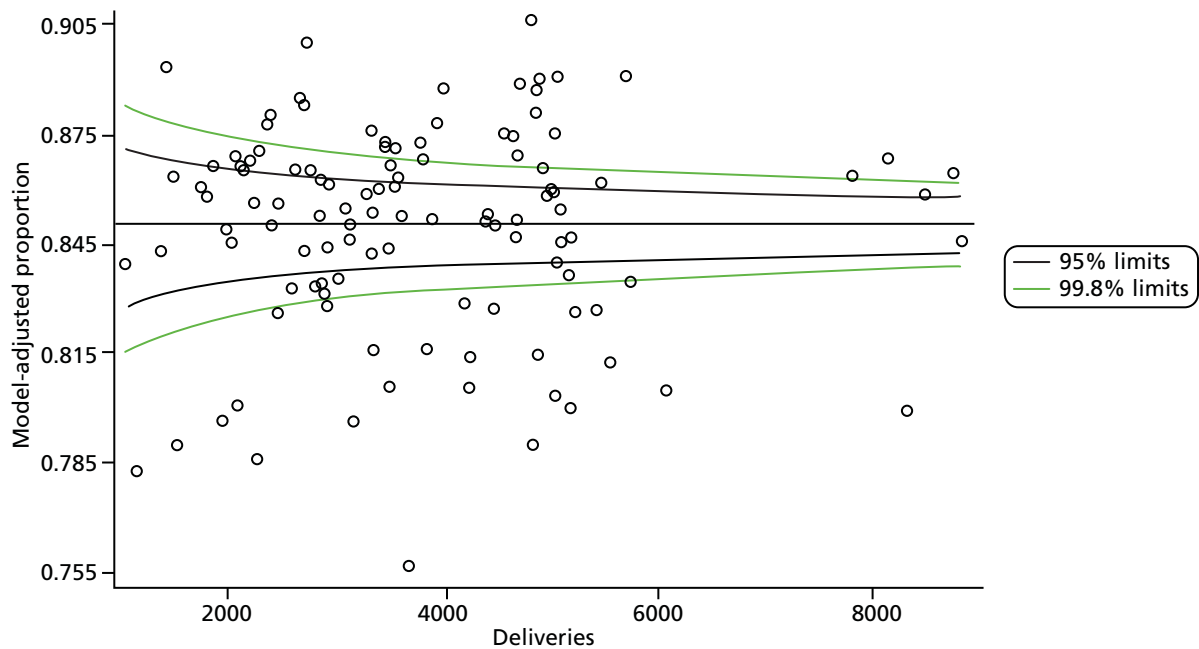
FIGURE 5 Funnel plot: healthy baby full model.
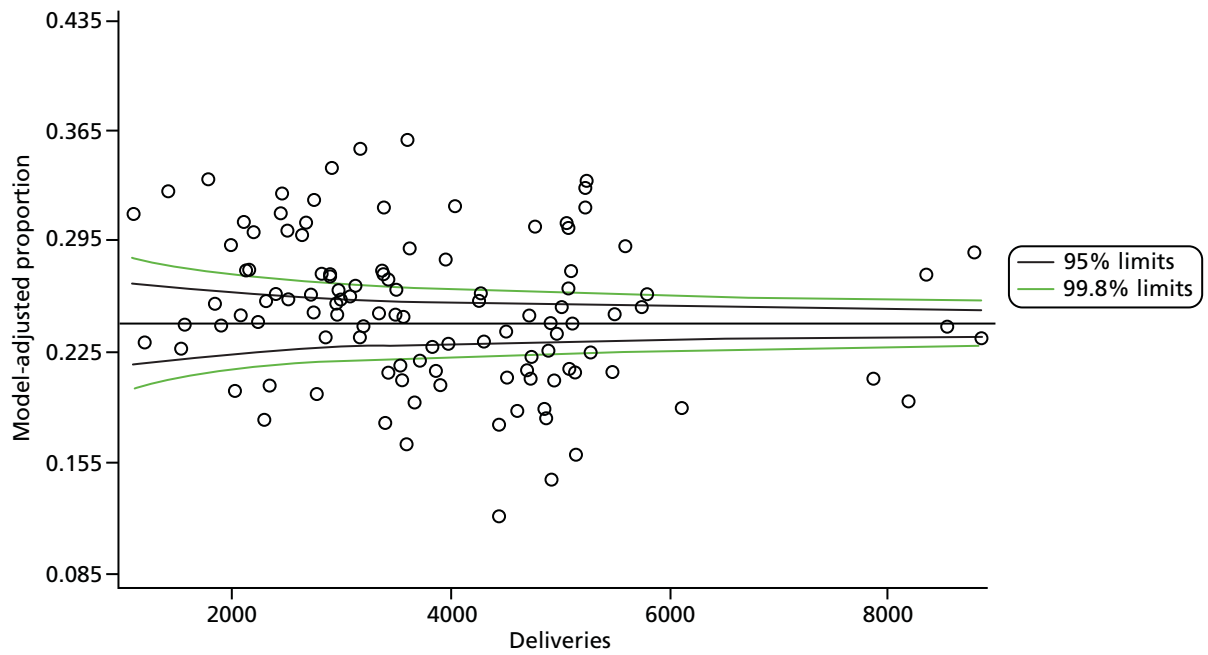
FIGURE 6 Funnel plot: healthy mother/healthy baby dyad intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
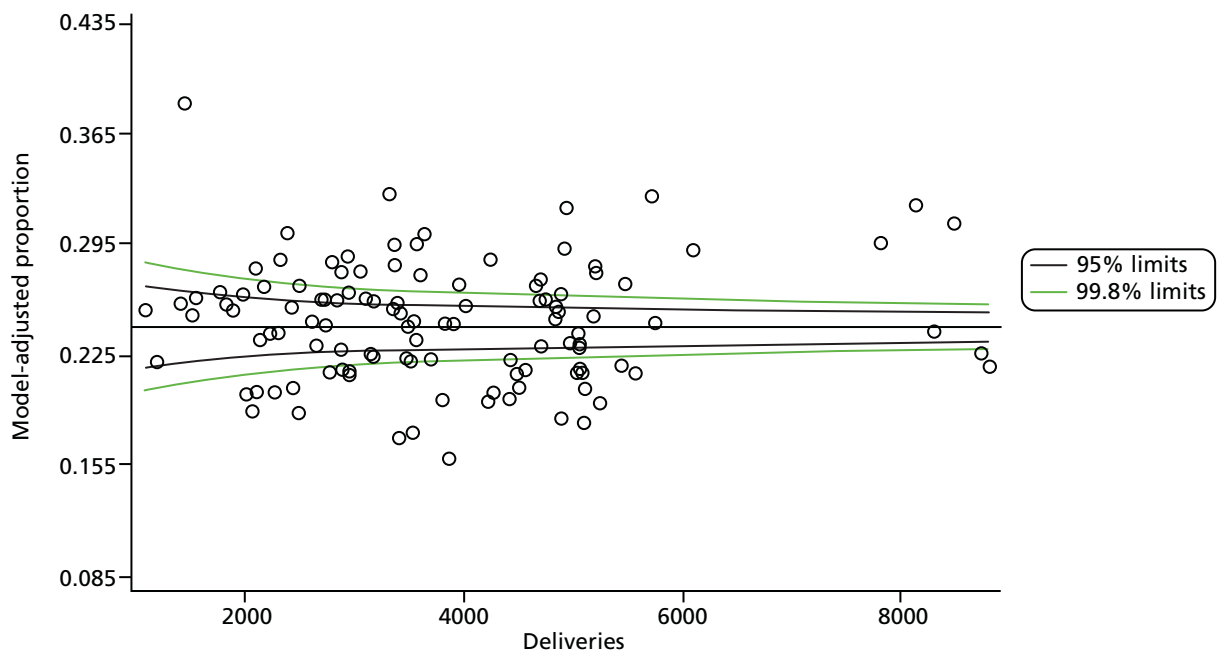
FIGURE 7 Funnel plot: healthy mother/healthy baby dyad full model.
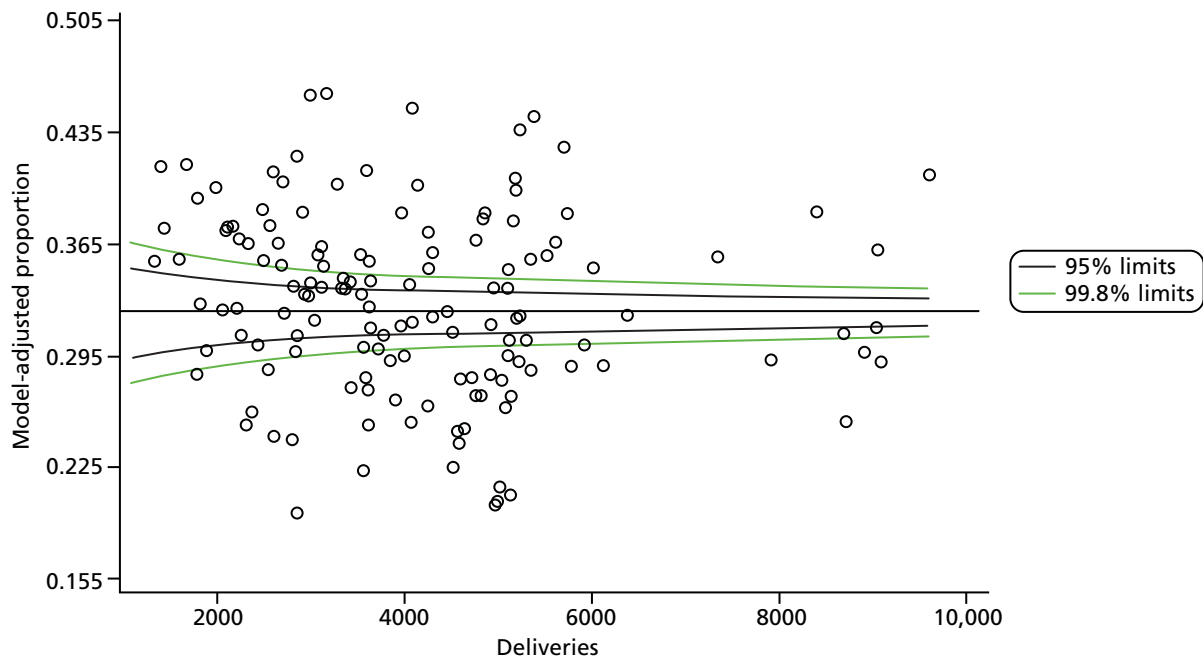
FIGURE 8 Funnel plot: delivery with bodily integrity intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
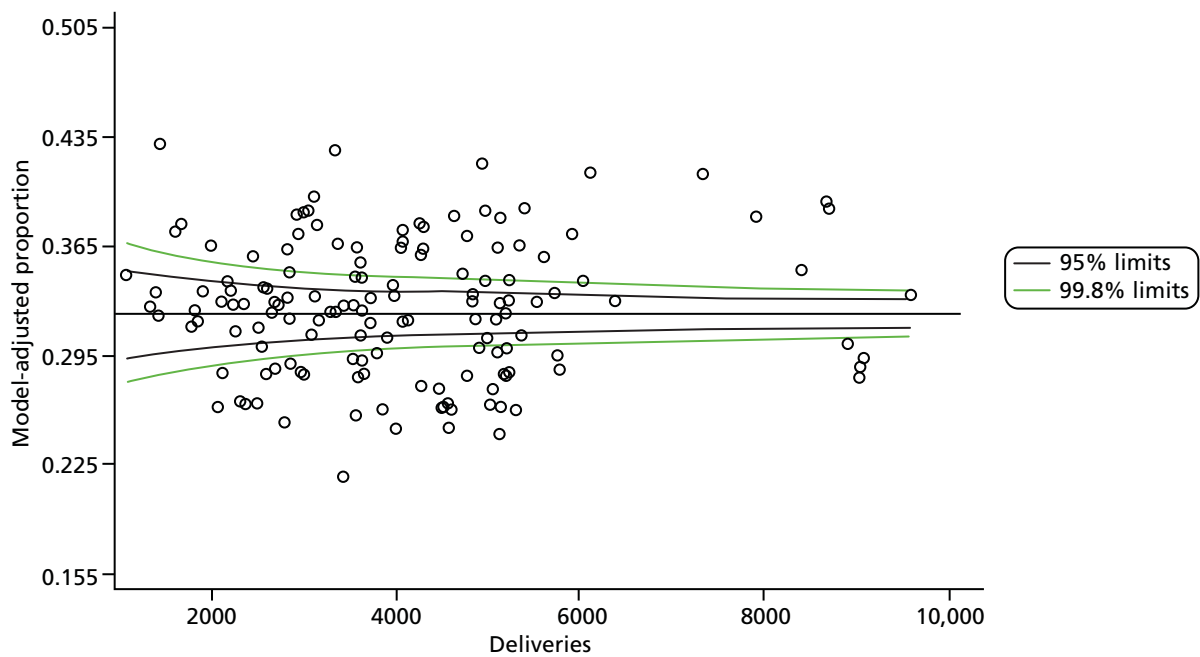
FIGURE 9 Funnel plot: delivery with bodily integrity full model.
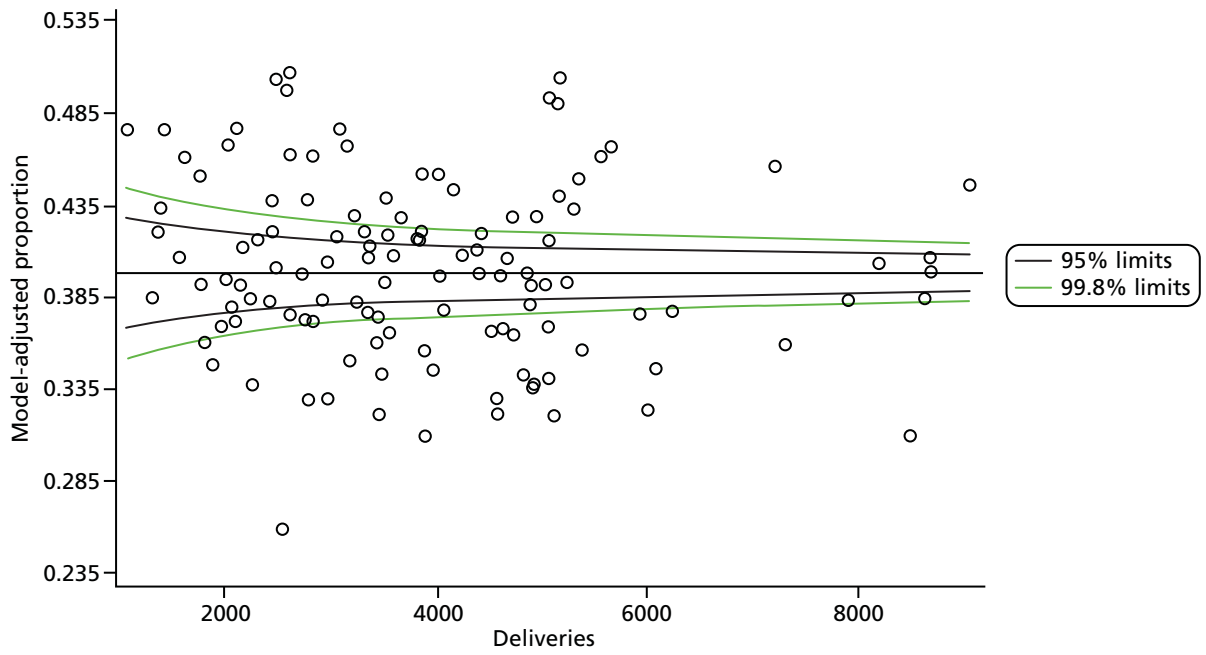
FIGURE 10 Funnel plot: normal birth intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
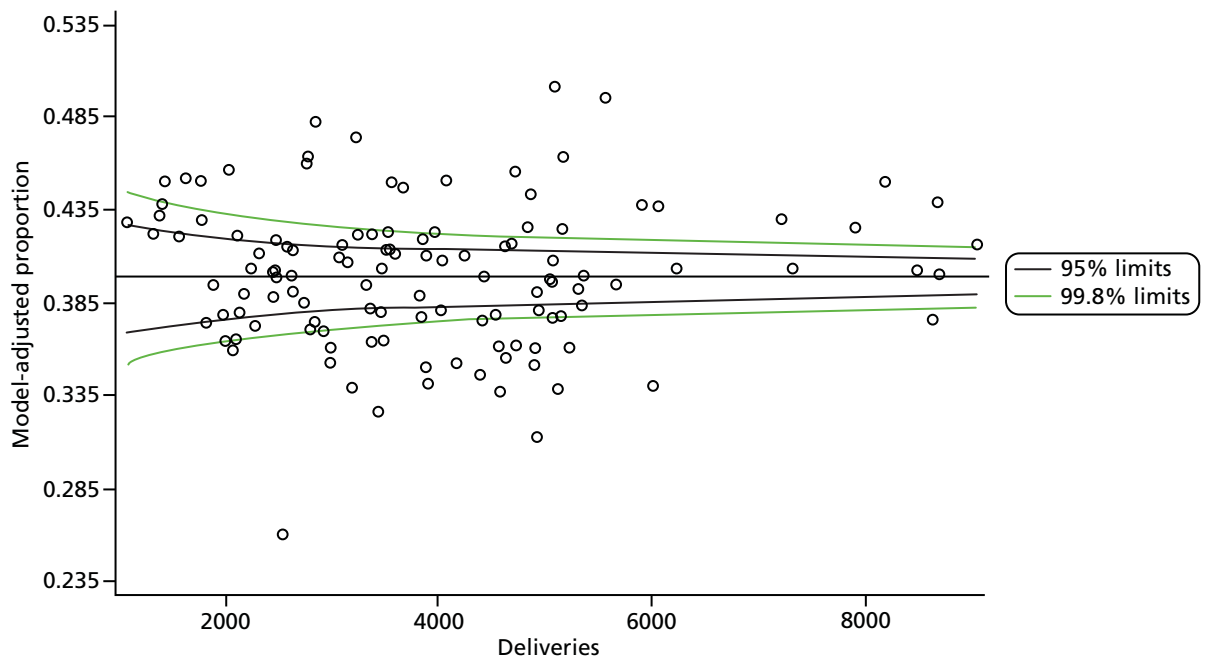
FIGURE 11 Funnel plot: normal birth full model.
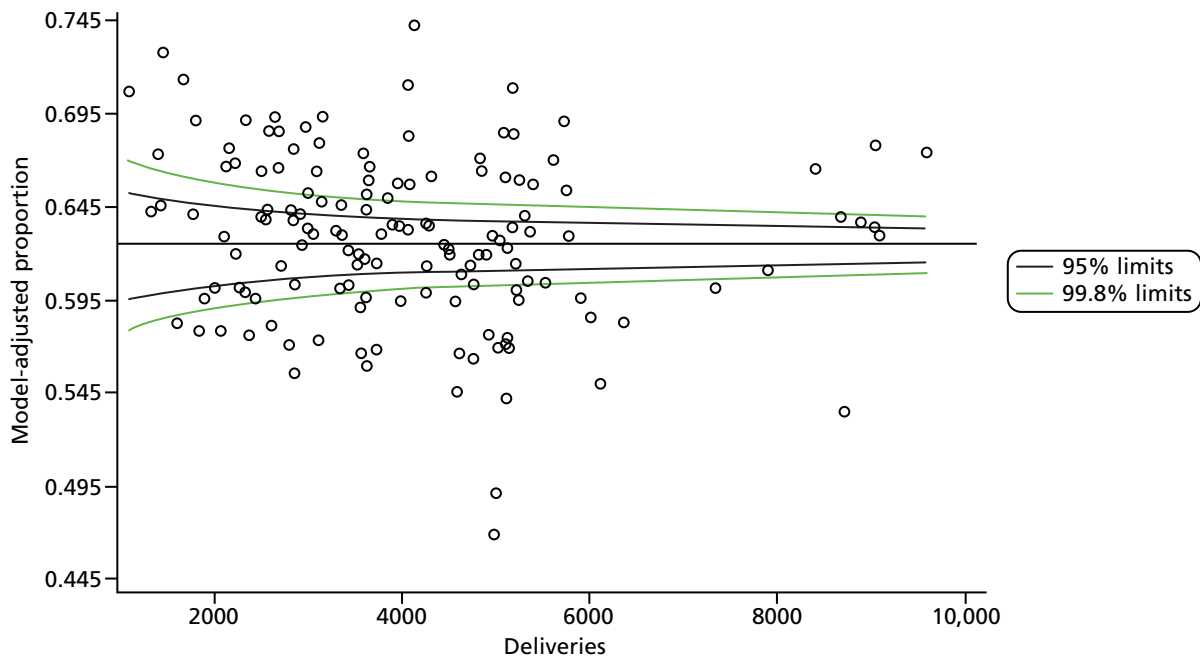
FIGURE 12 Funnel plot: spontaneous vaginal delivery intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
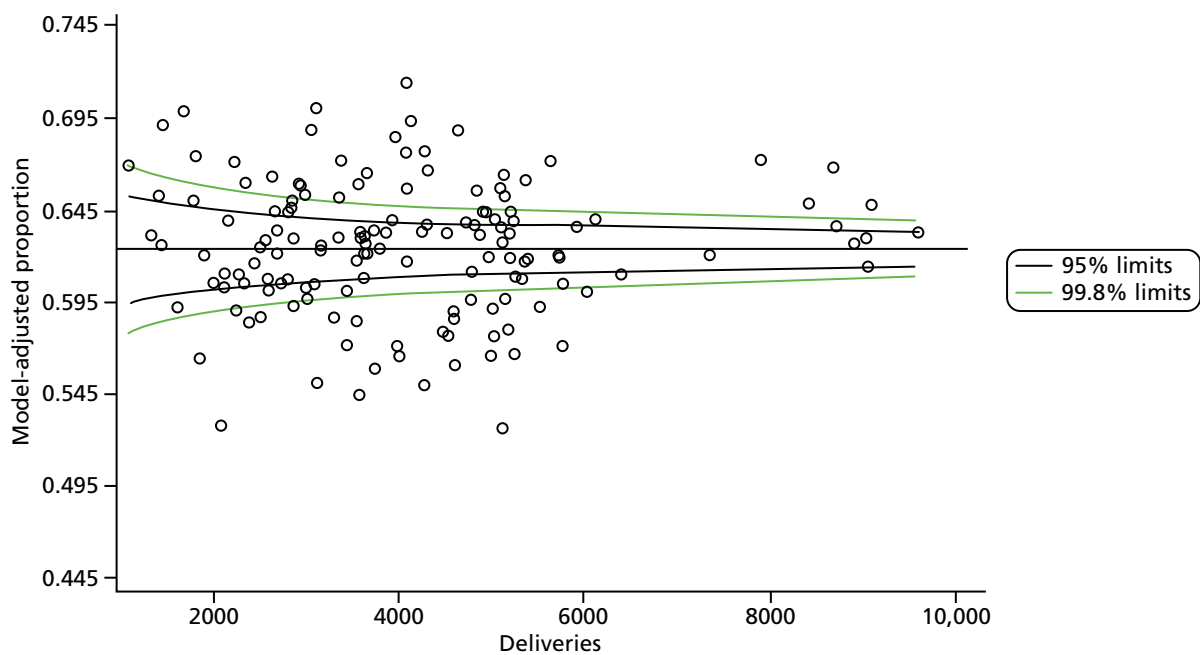
FIGURE 13 Funnel plot: spontaneous vaginal delivery full model.
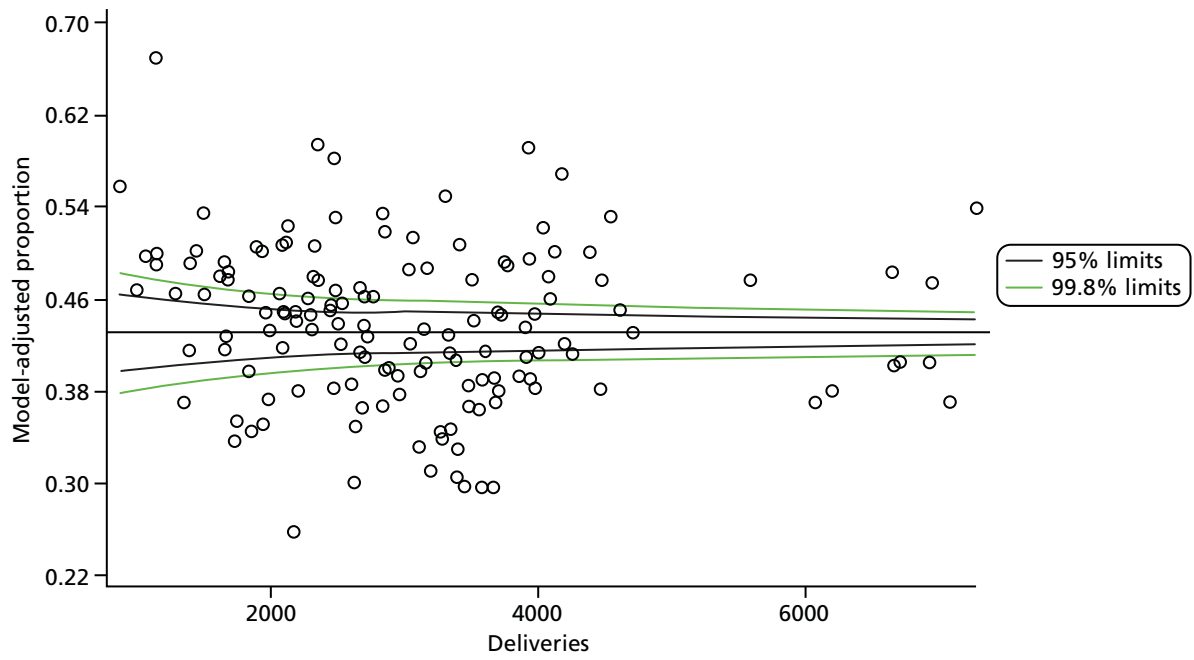
FIGURE 14 Funnel plot: intact perineum intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
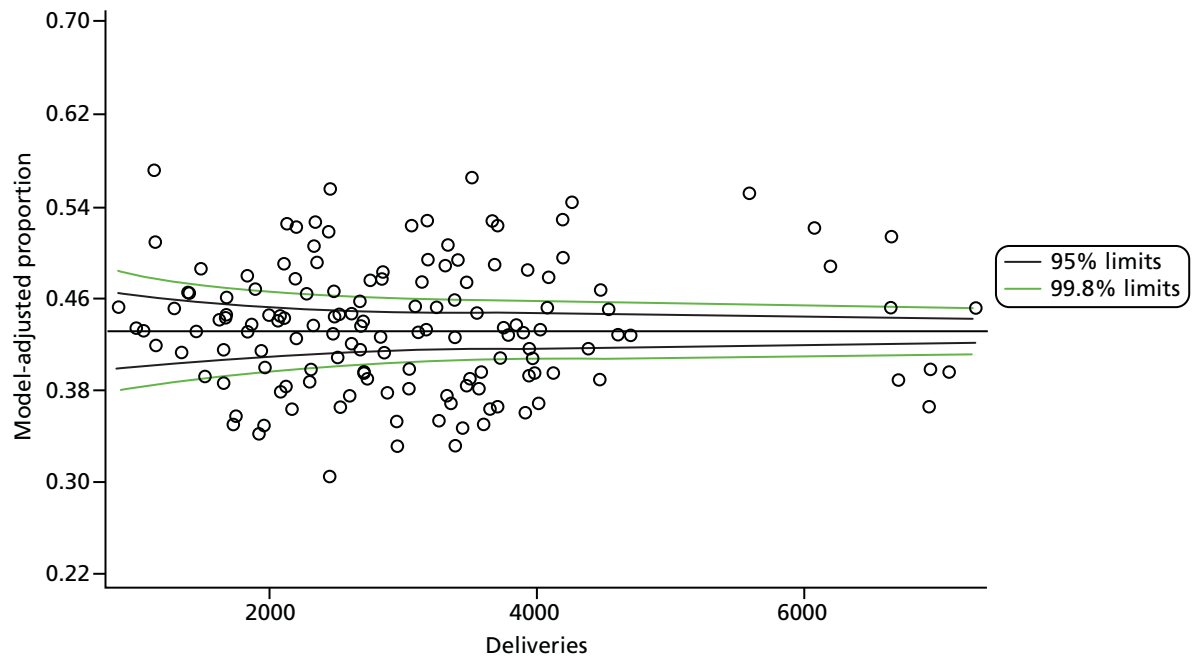
FIGURE 15 Funnel plot: intact perineum full model.
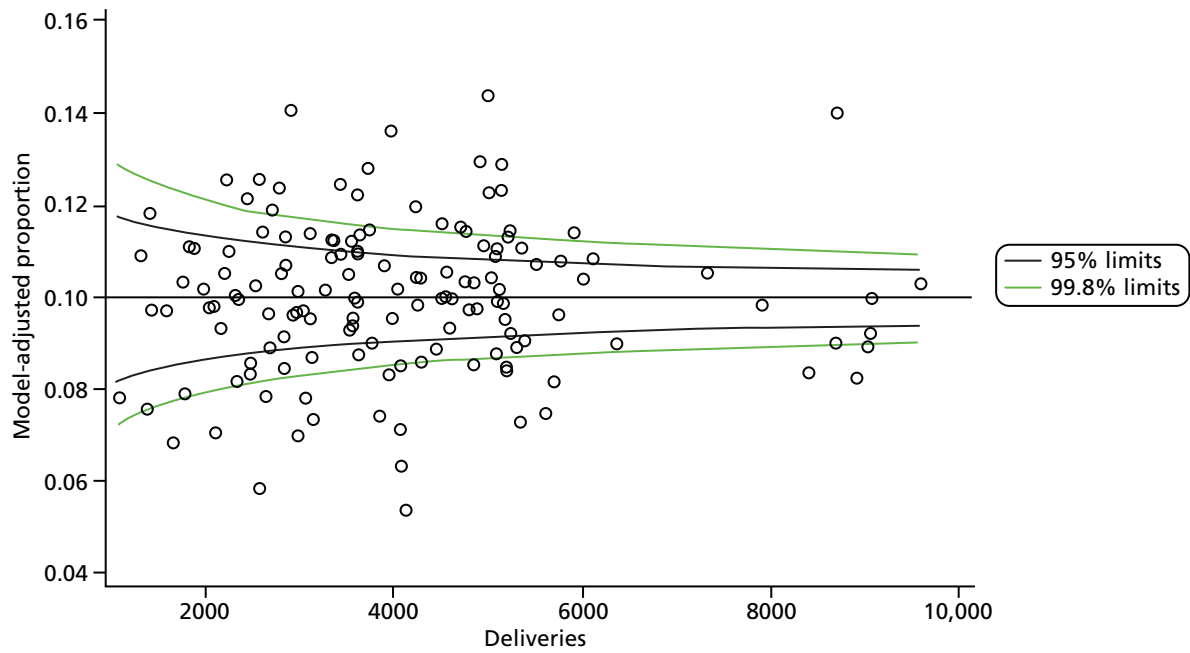
FIGURE 16 Funnel plot: elective caesarean intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
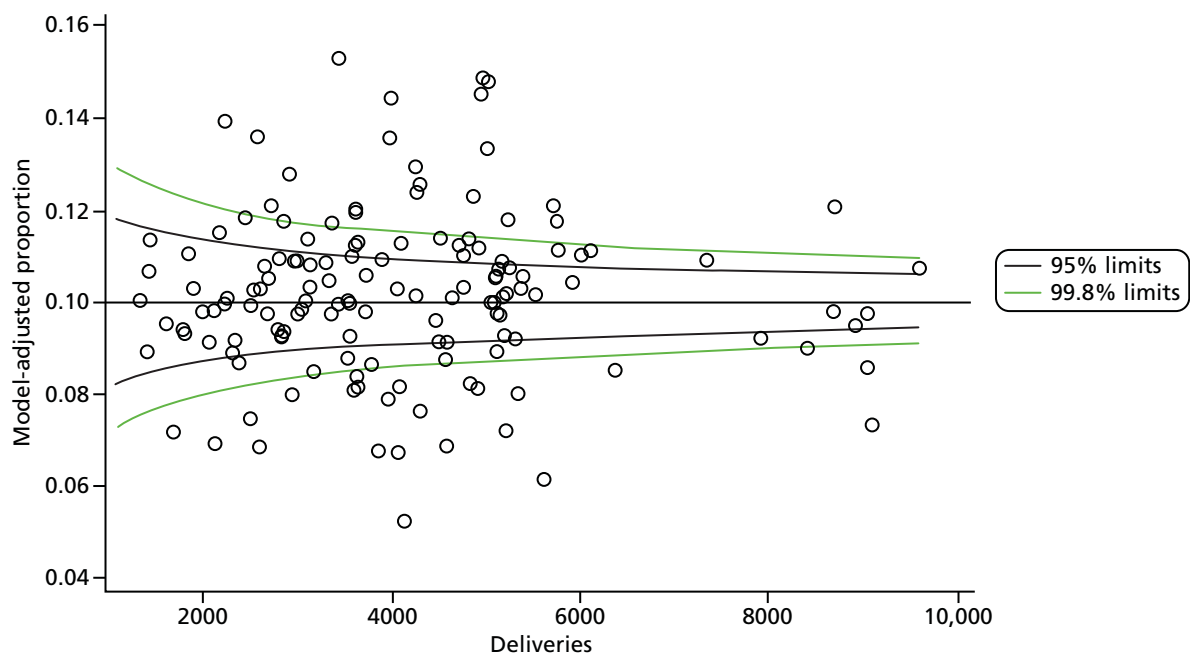
FIGURE 17 Funnel plot: elective caesarean full model.
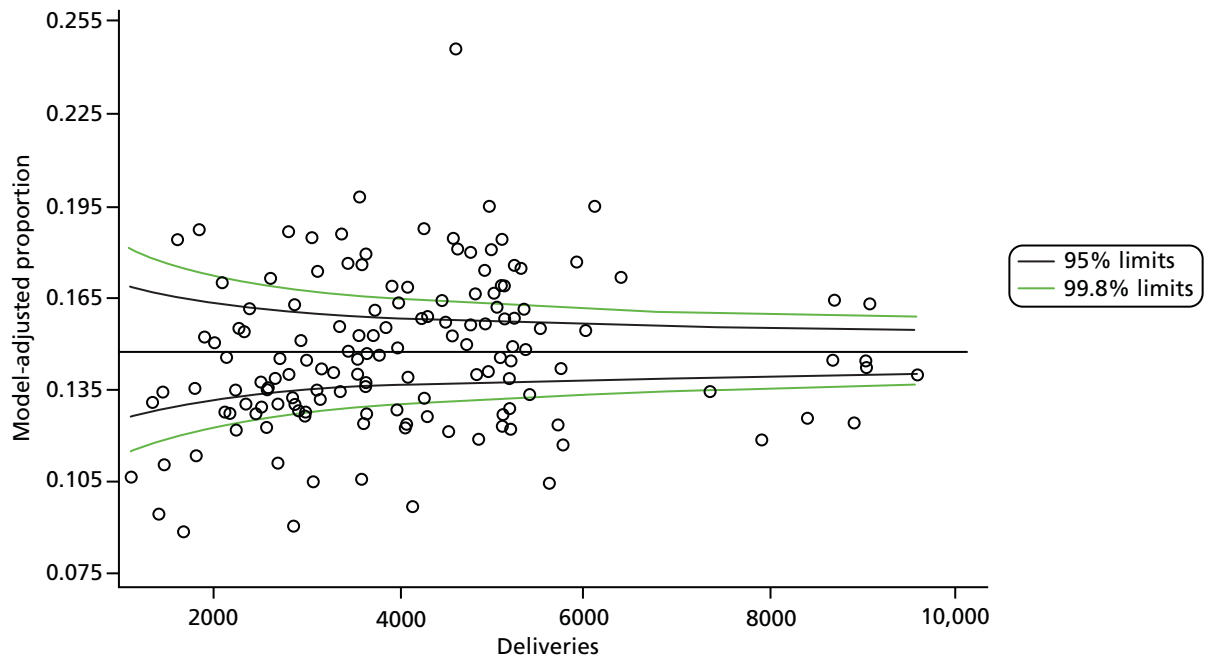
FIGURE 18 Funnel plot: emergency caesarean intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
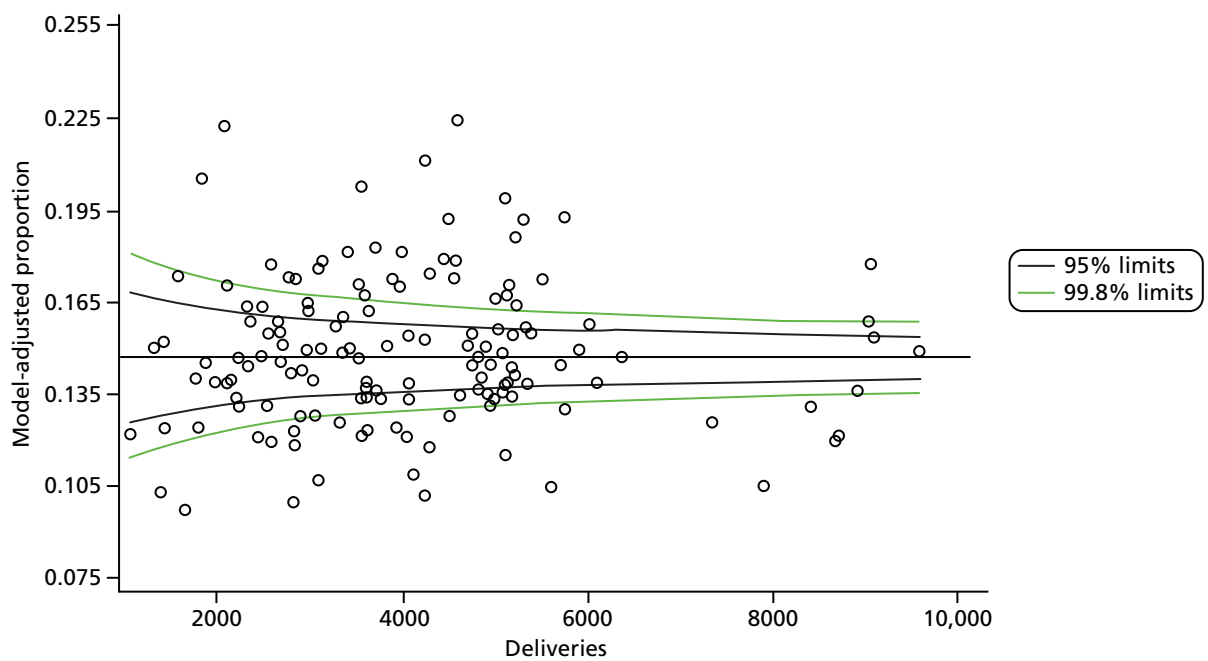
FIGURE 19 Funnel plot: emergency caesarean full model.
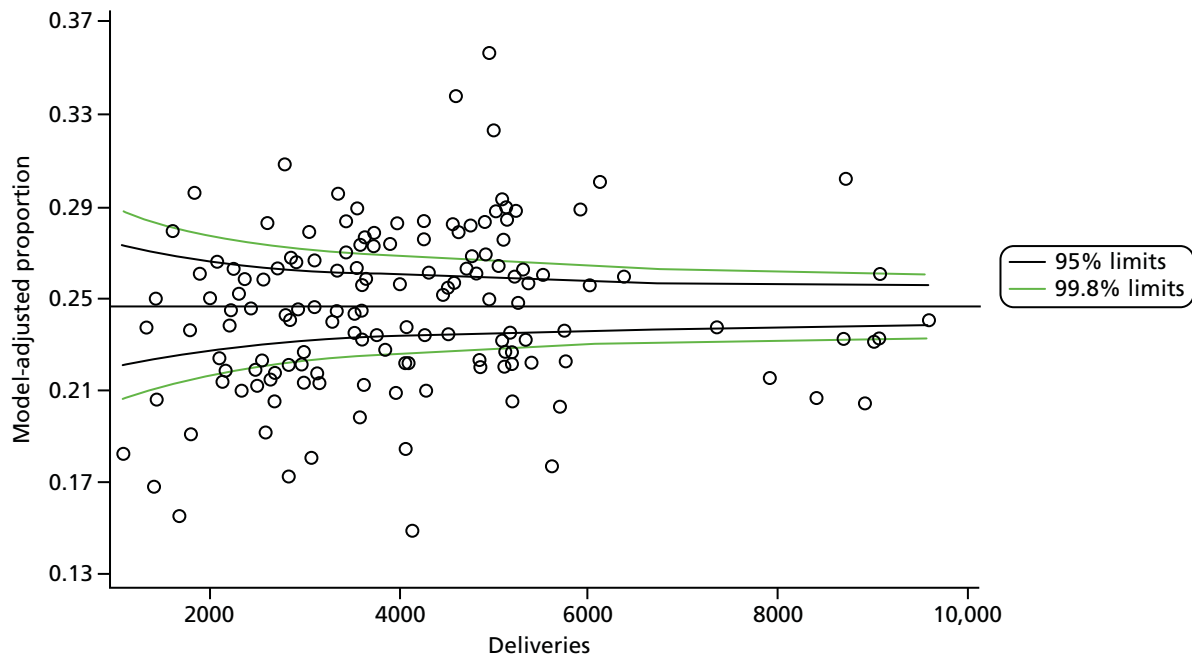
FIGURE 20 Funnel plot: caesarean intercept-only model. Note that 'model-adjusted proportion' is equivalent to the unadjusted proportion in the intercept-only model.
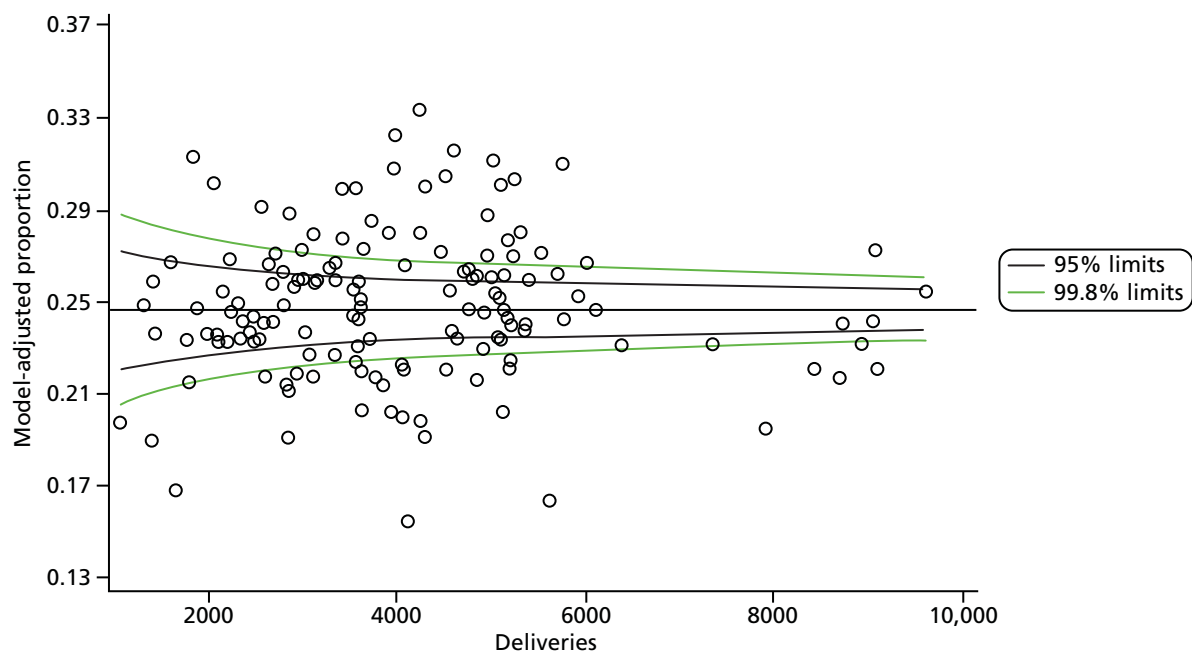
FIGURE 21 Funnel plot: caesarean full model.

Appendix 4 Multilevel models

The full reports of the models informing the results are below.

Healthy mother: multilevel model

Variable	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Intercept	−0.003	0.275	−0.011	0.9910				0.112	0.290	0.388	0.6987			
Mother's age group (years)														
≤19	1.320	0.092	14.422	<0.0001	3.744	3.129	4.479	1.330	0.092	14.502	<0.0001	3.781	3.159	4.526
20–24	0.999	0.091	11.032	<0.0001	2.715	2.273	3.242	1.008	0.091	11.115	<0.0001	2.741	2.294	3.274
25–29	0.629	0.090	6.962	<0.0001	1.62876	1.571	2.239	0.638	0.091	7.053	<0.0001	1.894	1.586	2.261
30–34	0.363	0.090	4.025	<0.0001	1.438	1.205	1.717	0.373	0.090	4.121	<0.0001	1.452	1.216	1.734
35–39	0.189	0.090	2.087	0.0369	1.208	1.012	1.442	0.199	0.091	2.189	0.0286	1.220	1.021	1.457
40–44	0.035	0.092	0.380	0.7043	1.036	0.864	1.241	0.044	0.092	0.475	0.6348	1.045	0.872	1.253
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−2.631	0.017	−154.267	<0.0001	0.072	0.070	0.074	−2.631	0.017	−154.270	<0.0001	0.072	0.070	0.074
1	−1.359	0.016	−85.732	<0.0001	0.257	0.249	0.265	−1.359	0.016	−85.734	<0.0001	0.257	0.249	0.265
2	−0.726	0.017	−43.835	<0.0001	0.484	0.468	0.500	−0.727	0.017	−43.844	<0.0001	0.484	0.468	0.500
3	−0.362	0.019	−19.084	<0.0001	0.696	0.671	0.722	−0.362	0.019	−19.083	<0.0001	0.696	0.671	0.722
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	−0.984	0.008	−125.860	<0.0001	0.374	0.368	0.380	−0.984	0.008	−125.860	<0.0001	0.374	0.368	0.380

| Variable | Model 1: FTE staffing variables | | | | | | | Model 2: staffing ratio variables | | | | | | |
Ethnicity	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Not given/not known/not stated	0.128	0.031	4.086	<0.0001	1.137	1.069	1.209	0.128	0.031	4.072	<0.0001	1.137	1.069	1.209
English/Welsh/Scottish/Northern Irish/British	0.145	0.025	5.777	<0.0001	1.156	1.101	1.215	0.145	0.025	5.783	<0.0001	1.157	1.101	1.215
Irish	−0.123	0.066	−1.878	0.0603	0.884	0.778	1.005	−0.124	0.066	−1.897	0.0578	0.883	0.777	1.004
Any other white background	0.107	0.028	3.900	<0.0001	1.113	1.055	1.175	0.107	0.028	3.901	<0.0001	1.113	1.055	1.175
White and black Caribbean	0.575	0.056	10.218	<0.0001	1.776	1.591	1.983	0.575	0.056	10.218	<0.0001	1.776	1.591	1.983
White and black African	0.146	0.070	2.100	0.0358	1.157	1.010	1.326	0.142	0.070	2.038	0.0416	1.152	1.005	1.321
White and Asian	−0.093	0.083	−1.113	0.2658	0.911	0.774	1.073	−0.094	0.083	−1.130	0.2585	0.910	0.773	1.072
Any other mixed/multiple ethnic background	0.162	0.057	2.835	0.0046	1.175	1.051	1.314	0.161	0.057	2.821	0.0048	1.174	1.050	1.313
Indian	−0.503	0.035	−14.218	<0.0001	0.605	0.564	0.648	−0.502	0.035	−14.198	<0.0001	0.605	0.565	0.649
Pakistani	−0.056	0.031	−1.796	0.0724	0.945	0.889	1.005	−0.055	0.031	−1.768	0.0771	0.946	0.890	1.006
Bangladeshi	−0.194	0.040	−4.820	<0.0001	0.823	0.761	0.891	−0.193	0.040	−4.787	<0.0001	0.825	0.762	0.892
Chinese	−0.372	0.058	−6.451	<0.0001	0.690	0.616	0.772	−0.372	0.058	−6.457	<0.0001	0.689	0.616	0.772
Any other Asian background	−0.200	0.038	−5.310	<0.0001	0.819	0.761	0.882	−0.199	0.038	−5.286	<0.0001	0.820	0.762	0.883
African	0.027	0.032	0.845	0.3980	1.027	0.965	1.093	0.028	0.032	0.872	0.3832	1.028	0.966	1.094
Caribbean	0.507	0.042	12.052	<0.0001	1.660	1.528	1.802	0.508	0.042	12.086	<0.0001	1.662	1.531	1.805
Any other black/African/Caribbean background	0.192	0.045	4.278	<0.0001	1.212	1.110	1.323	0.192	0.045	4.277	<0.0001	1.212	1.110	1.323
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

Variable	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
IMD	*β*	*SE(β)*	*t-statistic*	*Pr>t*	*OR*	*OR_L95%*	*OR_U95%*	*β*	*SE(β)*	*t-statistic*	*Pr>t*	*OR*	*OR_L95%*	*OR_U95%*
1 = most deprived	0.324	0.014	22.336	<0.0001	1.382	1.344	1.422	0.324	0.014	22.339	<0.0001	1.382	1.344	1.422
2	0.236	0.014	16.680	<0.0001	1.266	1.231	1.301	0.236	0.014	16.681	<0.0001	1.266	1.231	1.301
3	0.149	0.014	10.485	<0.0001	1.161	1.129	1.194	0.149	0.014	10.489	<0.0001	1.161	1.129	1.194
4	0.096	0.015	6.648	<0.0001	1.101	1.070	1.133	0.097	0.015	6.656	<0.0001	1.101	1.071	1.133
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	*β*	*SE(β)*	*t-statistic*	*Pr>t*	*OR*	*OR_L95%*	*OR_U95%*	*β*	*SE(β)*	*t-statistic*	*Pr>t*	*OR*	*OR_L95%*	*OR_U95%*
No information/other postcode														
Urban ≥ 10,000 – sparse	−0.146	0.096	−1.516	0.1296	0.864	0.716	1.044	−0.139	0.096	−1.442	0.1492	0.871	0.721	1.051
Town and fringe – sparse	0.006	0.070	0.079	0.9372	1.006	0.876	1.154	0.008	0.070	0.110	0.9125	1.008	0.878	1.156
Village – sparse	0.073	0.073	1.004	0.3155	1.076	0.933	1.241	0.069	0.073	0.947	0.3437	1.071	0.929	1.236
Hamlet and isolated dwelling – sparse	−0.100	0.098	−1.024	0.3060	0.905	0.747	1.096	−0.111	0.098	−1.139	0.2545	0.895	0.739	1.084
Urban ≥ 10,000 – less sparse	0.059	0.030	1.957	0.0504	1.060	1.000	1.124	0.056	0.030	1.882	0.0598	1.058	0.998	1.122
Town and fringe – less sparse	0.109	0.032	3.361	0.0008	1.115	1.046	1.188	0.107	0.032	3.299	0.0010	1.113	1.044	1.185
Village – less sparse	0.092	0.034	2.701	0.0069	1.096	1.026	1.172	0.089	0.034	2.613	0.0090	1.093	1.023	1.169
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

Variable	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
SHA	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	0.215	0.119	1.805	0.0711	1.240	0.982	1.567	0.215	0.120	1.792	0.0731	1.239	0.980	1.567
North West	0.032	0.092	0.347	0.7286	1.033	0.862	1.237	0.044	0.093	0.477	0.6330	1.045	0.871	1.254
Yorkshire and Humber	0.205	0.098	2.096	0.0361	1.228	1.013	1.488	0.212	0.099	2.150	0.0316	1.236	1.019	1.499
East Midlands	0.178	0.115	1.553	0.1204	1.195	0.954	1.496	0.176	0.115	1.535	0.1247	1.193	0.952	1.495
West Midlands	0.063	0.109	0.584	0.5595	1.065	0.861	1.318	0.059	0.109	0.546	0.5853	1.061	0.857	1.314
East of England	0.093	0.109	0.854	0.3930	1.097	0.887	1.359	0.084	0.109	0.772	0.4402	1.088	0.878	1.348
London	−0.116	0.101	−1.140	0.2544	0.891	0.730	1.087	−0.121	0.102	−1.187	0.2353	0.886	0.726	1.082
South East Coast	−0.030	0.121	−0.247	0.8048	0.970	0.765	1.231	−0.031	0.122	−0.256	0.7980	0.969	0.763	1.231
South Central	−0.084	0.115	−0.732	0.4642	0.919	0.734	1.152	−0.101	0.115	−0.878	0.3801	0.904	0.722	1.132
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	−0.028	0.015	−1.880	0.0602	0.972	0.944	1.001	−0.033	0.015	−2.178	0.0294	0.968	0.940	0.997
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	0.057	0.057	1.002	0.3164	1.058	0.947	1.183	0.054	0.057	0.940	0.3470	1.055	0.943	1.180
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
OU	0.006	0.071	0.083	0.9336	1.006	0.875	1.157	0.004	0.072	0.063	0.9499	1.005	0.873	1.156
OU/AMU	−0.052	0.081	−0.635	0.5252	0.950	0.810	1.114	−0.047	0.081	−0.578	0.5631	0.954	0.813	1.119
OU/AMU/FMU	0.003	0.098	0.032	0.9748	1.003	0.828	1.215	0.024	0.098	0.240	0.8101	1.024	0.845	1.241
OU/FMU	0.000				1.000			0.000				1.000		

Variable	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
Staffing variables	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
FTE doctors per 100 maternities	0.069	0.134	0.510	0.6099	1.071	0.823	1.394							
FTE midwives per 100 maternities	0.084	0.062	1.353	0.1759	1.088	0.963	1.230							
FTE support workers per 100 maternities	−0.114	0.071	−1.603	0.1089	0.892	0.776	1.026							
FTE staff per 100 maternities								0.028	0.038	0.743	0.4575	1.028	0.955	1.107
Doctor-to-midwife ratio								0.274	0.394	0.697	0.4860	1.316	0.608	2.846
Support worker-to-midwife ratio								−0.334	0.235	−1.421	0.1552	0.716	0.452	1.135
Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.283	0.011						0.283	0.011					
Mother level	0.308	0.021						0.308	0.021					
Mother level, sociodemographics	0.256	0.018						0.256	0.018					
Mother level, sociodemographics, trust level	0.242	0.017						0.242	0.017					
Mother level, sociodemographics, trust level, staff 1	0.238	0.016												
Mother level, sociodemographics, trust level, staff 2								0.239	0.016					

Variable	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
Global tests (df)	χ^2	Pr > χ^2	χ^2	Pr > χ^2
Mother's age group (6 df)	6495.901	< 0.0001	6499.150	< 0.0001
Parity (4 df)	40,071.673	< 0.0001	40,071.748	< 0.0001
Clinical risk (1 df)	15,840.786	< 0.0001	15,840.642	< 0.0001
Ethnicity (16 df)	1190.077	< 0.0001	1188.515	< 0.0001
IMD (4 df)	582.919	< 0.0001	582.912	< 0.0001
Rural/urban classification (7 df)	27.018	0.0003	26.419	0.0004
SHA (9 df)	19.979	0.0180	20.781	0.0137
ONS maternities (1 df)	3.533	0.0602	4.744	0.0294
University trust (1 df)	1.004	0.3164	0.884	0.3470
Configuration (3 df)	0.966	0.8095	0.972	0.8080
FTE doctors per 100 maternities (1 df)	0.260	0.6099		
FTE midwives per 100 maternities (1 df)	1.832	0.1759		
FTE support workers per 100 maternities (1 df)	2.570	0.1089		
FTE staff per 100 maternities (1 df)			0.552	0.4575
Doctor-to-midwife ratio (1 df)			0.485	0.4860
Support worker-to-midwife ratio (1 df)			2.020	0.1552

Variable	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
AUC	*AUC*	*SE(AUC)*	*AUC$_{L95\%}$*	*AUC$_{U95\%}$*	*AUC*	*SE(AUC)*	*AUC$_{L95\%}$*	*AUC$_{U95\%}$*
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.735	0.001	0.733	0.737	0.735	0.001	0.733	0.737
Mother level, sociodemographics	0.741	0.001	0.740	0.743	0.741	0.001	0.740	0.743
Mother level, sociodemographics, trust level	0.742	0.001	0.740	0.744	0.742	0.001	0.740	0.744
Mother level, sociodemographics, trust level, staff 1	0.742	0.001	0.741	0.744				
Mother level, sociodemographics, trust level, staff 2					0.742	0.001	0.740	0.744

AUC$_{L95\%}$, lower 95% confidence limit of the AUC; AUC$_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient; SE(AUC), SE of area under the curve; SE(AUC), SE of area under the curve.
Pr > χ^2 refers to the probability of exceeding the χ^2-statistic.
Pr > *t* refers to the probability of exceeding the *t*-statistic.

Healthy baby: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Intercept	2.817	0.272	10.350	<0.0001				2.828	0.285	9.918	<0.0001			
Mother's age group (years)														
≤19	0.092	0.086	1.071	0.2841	1.097	0.926	1.299	0.092	0.086	1.070	0.2847	1.097	0.926	1.299
20–24	0.116	0.085	1.375	0.1692	1.123	0.952	1.326	0.116	0.085	1.376	0.1690	1.123	0.952	1.326
25–29	0.193	0.084	2.290	0.0220	1.213	1.028	1.431	0.193	0.084	2.288	0.0222	1.213	1.028	1.431
30–34	0.208	0.084	2.463	0.0138	1.231	1.043	1.452	0.207	0.084	2.461	0.0138	1.230	1.043	1.451
35–39	0.212	0.085	2.504	0.0123	1.236	1.047	1.458	0.212	0.085	2.504	0.0123	1.236	1.047	1.458
40–44	0.267	0.087	3.078	0.0021	1.306	1.102	1.549	0.267	0.087	3.074	0.0021	1.306	1.102	1.548
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−0.317	0.020	−15.732	<0.0001	0.728	0.700	0.757	−0.317	0.020	−15.727	<0.0001	0.728	0.700	0.758
1	0.185	0.020	9.126	<0.0001	1.203	1.156	1.252	0.185	0.020	9.135	<0.0001	1.203	1.157	1.252
2	0.154	0.022	7.037	<0.0001	1.166	1.117	1.217	0.154	0.022	7.055	<0.0001	1.167	1.118	1.218
3	0.080	0.025	3.173	0.0015	1.084	1.031	1.139	0.080	0.025	3.163	0.0016	1.083	1.031	1.139
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	−1.986	0.012	−163.455	<0.0001	0.137	0.134	0.141	−1.986	0.012	−163.462	<0.0001	0.137	0.134	0.141

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Not given/not known/not stated	−0.058	0.038	−1.552	0.1208	0.943	0.876	1.015	−0.059	0.038	−1.574	0.1155	0.943	0.876	1.015
English/Welsh/Scottish/Northern Irish/British	−0.110	0.030	−3.629	0.0003	0.896	0.845	0.951	−0.109	0.030	−3.626	0.0003	0.896	0.845	0.951
Irish	−0.166	0.072	−2.301	0.0214	0.847	0.736	0.976	−0.166	0.072	−2.306	0.0211	0.847	0.735	0.975
Any other white background	−0.016	0.033	−0.497	0.6193	0.984	0.922	1.050	−0.016	0.033	−0.469	0.6390	0.985	0.923	1.051
White and black Caribbean	−0.165	0.071	−2.320	0.0203	0.848	0.738	0.975	−0.165	0.071	−2.325	0.0201	0.848	0.738	0.974
White and black African	0.080	0.089	0.892	0.3722	1.083	0.909	1.290	0.080	0.089	0.895	0.3706	1.083	0.909	1.291
White and Asian	−0.151	0.095	−1.601	0.1093	0.860	0.714	1.034	−0.151	0.095	−1.601	0.1093	0.860	0.714	1.034
Any other mixed/multiple ethnic background	−0.046	0.069	−0.659	0.5098	0.955	0.834	1.094	−0.046	0.069	−0.663	0.5074	0.955	0.834	1.094
Indian	−0.136	0.039	−3.511	0.0004	0.873	0.809	0.942	−0.136	0.039	−3.517	0.0004	0.873	0.809	0.942
Pakistani	−0.118	0.038	−3.121	0.0018	0.888	0.825	0.957	−0.118	0.038	−3.106	0.0019	0.889	0.825	0.957
Bangladeshi	−0.161	0.048	−3.344	0.0008	0.851	0.774	0.935	−0.161	0.048	−3.346	0.0008	0.851	0.774	0.935
Chinese	0.173	0.069	2.509	0.0121	1.189	1.039	1.361	0.173	0.069	2.506	0.0122	1.188	1.038	1.360
Any other Asian background	−0.028	0.044	−0.633	0.5268	0.973	0.892	1.060	−0.029	0.044	−0.670	0.5030	0.971	0.891	1.058
African	−0.017	0.038	−0.448	0.6542	0.983	0.913	1.059	−0.017	0.038	−0.455	0.6495	0.983	0.912	1.059
Caribbean	−0.138	0.052	−2.626	0.0086	0.872	0.786	0.966	−0.137	0.052	−2.625	0.0087	0.872	0.787	0.966
Any other black/African/Caribbean background	−0.180	0.053	−3.390	0.0007	0.835	0.753	0.927	−0.181	0.053	−3.404	0.0007	0.835	0.752	0.926
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
IMD	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
1 = most deprived	−0.158	0.017	−9.206	<0.0001	0.854	0.826	0.883	−0.158	0.017	−9.213	<0.0001	0.854	0.826	0.883
2	−0.096	0.017	−5.775	<0.0001	0.909	0.879	0.939	−0.096	0.017	−5.778	<0.0001	0.908	0.879	0.939
3	−0.059	0.017	−3.563	0.0004	0.942	0.912	0.974	−0.059	0.017	−3.544	0.0004	0.943	0.912	0.974
4	−0.032	0.017	−1.905	0.0567	0.968	0.937	1.001	−0.032	0.017	−1.913	0.0557	0.968	0.937	1.001
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
No information/ other postcode														
Urban ≥10,000 – sparse	−0.227	0.115	−1.979	0.0478	0.797	0.637	0.998	−0.226	0.115	−1.977	0.0481	0.797	0.637	0.998
Town and fringe – sparse	0.049	0.085	0.580	0.5620	1.051	0.889	1.241	0.049	0.085	0.579	0.5624	1.051	0.889	1.241
Village – sparse	0.078	0.090	0.862	0.3885	1.081	0.906	1.289	0.078	0.090	0.864	0.3877	1.081	0.906	1.290
Hamlet and isolated dwelling – sparse	−0.121	0.112	−1.079	0.2807	0.886	0.711	1.104	−0.119	0.112	−1.055	0.2915	0.888	0.713	1.107
Urban ≥10,000 – less sparse	0.060	0.035	1.702	0.0888	1.061	0.991	1.137	0.061	0.035	1.735	0.0828	1.063	0.992	1.138
Town and fringe – less sparse	0.048	0.038	1.269	0.2045	1.050	0.974	1.131	0.049	0.038	1.284	0.1991	1.050	0.975	1.132
Village – less sparse	0.099	0.040	2.442	0.0146	1.104	1.020	1.194	0.100	0.040	2.468	0.0136	1.105	1.021	1.196
Hamlet and isolated dwelling– less sparse	0.000				1.000			0.000				1.000		

SHA	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	−0.089	0.118	−0.754	0.4511	0.915	0.725	1.154	−0.087	0.118	−0.740	0.4591	0.916	0.727	1.155
North West	0.234	0.092	2.555	0.0106	1.264	1.056	1.512	0.236	0.091	2.581	0.0098	1.266	1.058	1.514
Yorkshire and Humber	0.103	0.097	1.063	0.2877	1.109	0.916	1.342	0.104	0.097	1.069	0.2849	1.109	0.917	1.342
East Midlands	0.209	0.114	1.839	0.0658	1.232	0.986	1.539	0.211	0.113	1.865	0.0622	1.235	0.989	1.541
West Midlands	0.144	0.108	1.335	0.1818	1.154	0.935	1.425	0.145	0.107	1.353	0.1760	1.156	0.937	1.427
East of England	0.169	0.108	1.563	0.1181	1.184	0.958	1.462	0.168	0.107	1.564	0.1179	1.183	0.958	1.460
London	0.263	0.100	2.618	0.0088	1.301	1.068	1.583	0.263	0.100	2.622	0.0087	1.300	1.069	1.582
South East Coast	0.173	0.120	1.441	0.1495	1.189	0.940	1.505	0.176	0.120	1.471	0.1413	1.193	0.943	1.509
South Central	0.098	0.114	0.857	0.3913	1.102	0.882	1.378	0.102	0.113	0.907	0.3642	1.108	0.888	1.382
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	−0.002	0.015	−0.140	0.8890	0.998	0.969	1.027	−0.003	0.015	−0.199	0.8426	0.997	0.969	1.026
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	0.125	0.056	2.237	0.0253	1.134	1.016	1.265	0.121	0.056	2.170	0.0300	1.129	1.012	1.260
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
OU	−0.040	0.071	−0.560	0.5756	0.961	0.837	1.104	−0.040	0.071	−0.561	0.5745	0.961	0.837	1.104
OU/AMU	−0.126	0.081	−1.558	0.1191	0.882	0.753	1.033	−0.126	0.080	−1.567	0.1172	0.882	0.754	1.032
OU/AMU/FMU	−0.167	0.097	−1.719	0.0855	0.846	0.700	1.024	−0.168	0.097	−1.742	0.0816	0.845	0.700	1.021
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
FTE doctors per 100 maternities	0.149	0.133	1.116	0.2642	1.160	0.894	1.507							
FTE midwives per 100 maternities	0.028	0.062	0.460	0.6456	1.029	0.912	1.161							
FTE support workers per 100 maternities	−0.034	0.071	−0.481	0.6308	0.967	0.842	1.110							
FTE staff per 100 maternities								0.036	0.037	0.963	0.3358	1.036	0.964	1.114
Doctor-to-midwife ratio								0.310	0.388	0.800	0.4239	1.363	0.638	2.914
Support worker-to-midwife ratio								−0.277	0.231	−1.201	0.2296	0.758	0.482	1.191

Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.227	0.011						0.227	0.011					
Mother level	0.268	0.019						0.268	0.019					
Mother level, sociodemographics	0.250	0.017						0.250	0.017					
Mother level, sociodemographics, trust level	0.235	0.016						0.235	0.016					
Mother level, sociodemographics, trust level, staff 1	0.233	0.016												
Mother level, sociodemographics, trust level, staff 2								0.233	0.016					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr > \chi^2$	χ^2	
Mother's age group (6 df)	85.881	< 0.0001	85.591	< 0.0001
Parity (4 df)	2461.647	< 0.0001	2462.493	< 0.0001
Clinical risk (1 df)	26,717.623	< 0.0001	26,719.786	< 0.0001
Ethnicity (16 df)	89.968	< 0.0001	90.138	< 0.0001
IMD (4 df)	103.297	< 0.0001	103.507	< 0.0001
Rural/urban classification (7 df)	16.186	0.0235	16.312	0.0224
SHA (9 df)	15.450	0.0793	15.567	0.0765
ONS maternities (1 df)	0.019	0.8890	0.039	0.8426
University trust (1 df)	5.005	0.0253	4.708	0.0300
Configuration (3 df)	4.507	0.2116	4.619	0.2019
FTE doctors per 100 maternities (1 df)	1.246	0.2642		
FTE midwives per 100 maternities (1 df)	0.211	0.6456		
FTE support workers per 100 maternities (1 df)	0.231	0.6308		
FTE staff per 100 maternities (1 df)			0.926	0.3358
Doctor-to-midwife ratio (1 df)			0.640	0.4239
Support worker-to-midwife ratio (1 df)			1.443	0.2296

AUC	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.720	0.001	0.718	0.722	0.720	0.001	0.718	0.722
Mother level, sociodemographics	0.724	0.001	0.722	0.726	0.724	0.001	0.722	0.726
Mother level, sociodemographics, trust level	0.726	0.001	0.724	0.728	0.726	0.001	0.724	0.728
Mother level, sociodemographics, trust level, staff 1	0.726	0.001	0.724	0.728				
Mother level, sociodemographics, trust level, staff 2					0.726	0.001	0.724	0.728

AUC$_{L95\%}$, lower 95% confidence limit of the AUC; AUC$_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.
Pr > χ^2 refers to the probability of exceeding the χ^2-statistic.
Pr > t refers to the probability of exceeding the t-statistic.

Healthy mother/healthy baby dyad: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Intercept	−0.107	0.276	−0.387	0.7000				−0.015	0.289	−0.053	0.9575			
Mother's age group (years)														
≤19	1.216	0.097	12.583	<0.0001	3.375	2.793	4.079	1.223	0.097	12.633	<0.0001	3.398	2.811	4.108
20–24	0.905	0.096	9.463	<0.0001	2.472	2.049	2.981	0.912	0.096	9.518	<0.0001	2.488	2.062	3.002
25–29	0.574	0.095	6.015	<0.0001	1.775	1.473	2.140	0.581	0.096	6.075	<0.0001	1.787	1.482	2.155
30–34	0.319	0.095	3.348	0.0008	1.376	1.142	1.659	0.326	0.096	3.413	0.0006	1.386	1.149	1.671
35–39	0.158	0.096	1.654	0.0982	1.171	0.971	1.413	0.165	0.096	1.719	0.0856	1.179	0.977	1.422
40–44	0.042	0.097	0.427	0.6695	1.042	0.861	1.262	0.049	0.098	0.501	0.6163	1.050	0.867	1.272
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−2.596	0.018	−147.515	<0.0001	0.075	0.072	0.077	−2.595	0.018	−147.489	<0.0001	0.075	0.072	0.077
1	−1.277	0.016	−78.648	<0.0001	0.279	0.270	0.288	−1.276	0.016	−78.616	<0.0001	0.279	0.270	0.288
2	−0.669	0.017	−39.482	<0.0001	0.512	0.495	0.529	−0.669	0.017	−39.447	<0.0001	0.512	0.496	0.530
3	−0.319	0.019	−16.440	<0.0001	0.727	0.700	0.755	−0.318	0.019	−16.418	<0.0001	0.727	0.700	0.756
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	−1.267	0.008	−153.090	<0.0001	0.282	0.277	0.286	−1.267	0.008	−153.088	<0.0001	0.282	0.277	0.286

Ethnicity	Model 1: FTE staffing variables								Model 2: staffing ratio variables							
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$		β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Not given/not known/not stated	0.104	0.033	3.174	0.0015	1.110	1.041	1.184		0.105	0.033	3.192	0.0014	1.111	1.041	1.184
English/Welsh/Scottish/Northern Irish/British	0.118	0.026	4.509	<0.0001	1.126	1.069	1.185		0.118	0.026	4.501	<0.0001	1.125	1.069	1.185
Irish	−0.134	0.069	−1.952	0.0509	0.875	0.764	1.001		−0.137	0.069	−1.991	0.0465	0.872	0.762	0.998
Any other white background	0.086	0.029	2.986	0.0028	1.090	1.030	1.153		0.085	0.029	2.968	0.0030	1.089	1.029	1.152
White and black Caribbean	0.486	0.058	8.338	<0.0001	1.626	1.451	1.823		0.484	0.058	8.298	<0.0001	1.622	1.447	1.819
White and black African	0.190	0.072	2.634	0.0084	1.209	1.050	1.392		0.185	0.072	2.571	0.0101	1.204	1.045	1.386
White and Asian	−0.091	0.087	−1.044	0.2967	0.913	0.770	1.083		−0.093	0.087	−1.067	0.2860	0.911	0.768	1.081
Any other mixed/multiple ethnic background	0.170	0.059	2.862	0.0042	1.185	1.055	1.331		0.173	0.059	2.919	0.0035	1.189	1.058	1.335
Indian	−0.516	0.037	−13.881	<0.0001	0.597	0.555	0.642		−0.516	0.037	−13.883	<0.0001	0.597	0.555	0.642
Pakistani	−0.054	0.033	−1.651	0.0987	0.947	0.889	1.010		−0.054	0.033	−1.651	0.0988	0.947	0.889	1.010
Bangladeshi	−0.197	0.042	−4.671	<0.0001	0.821	0.756	0.892		−0.195	0.042	−4.634	<0.0001	0.822	0.757	0.893
Chinese	−0.350	0.060	−5.859	<0.0001	0.705	0.627	0.792		−0.347	0.060	−5.807	<0.0001	0.707	0.629	0.795
Any other Asian background	−0.213	0.039	−5.411	<0.0001	0.808	0.748	0.873		−0.213	0.039	−5.410	<0.0001	0.808	0.748	0.873
African	0.067	0.033	2.039	0.0414	1.070	1.003	1.141		0.067	0.033	2.031	0.0423	1.069	1.002	1.141
Caribbean	0.505	0.044	11.603	<0.0001	1.657	1.522	1.805		0.502	0.044	11.530	<0.0001	1.652	1.517	1.799
Any other black/African/Caribbean background	0.162	0.047	3.442	0.0006	1.176	1.072	1.289		0.164	0.047	3.494	0.0005	1.179	1.075	1.292
Any other ethnic group, please describe	0.000				1.000				0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
IMD	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**
1 = most deprived	0.280	0.015	18.504	<0.0001	1.323	1.285	1.363	0.280	0.015	18.489	<0.0001	1.323	1.284	1.363
2	0.213	0.015	14.403	<0.0001	1.237	1.202	1.273	0.213	0.015	14.388	<0.0001	1.237	1.202	1.273
3	0.132	0.015	8.881	<0.0001	1.141	1.109	1.175	0.132	0.015	8.868	<0.0001	1.141	1.108	1.175
4	0.077	0.015	5.038	<0.0001	1.080	1.048	1.112	0.076	0.015	5.034	<0.0001	1.079	1.048	1.112
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**
No information/other postcode														
Urban ≥ 10,000 – sparse	–0.177	0.101	–1.756	0.0791	0.838	0.688	1.021	–0.170	0.100	–1.694	0.0903	0.844	0.693	1.027
Town and fringe – sparse	0.037	0.073	0.510	0.6098	1.038	0.899	1.198	0.036	0.073	0.498	0.6187	1.037	0.899	1.197
Village – sparse	0.075	0.076	0.982	0.3262	1.077	0.928	1.251	0.069	0.076	0.902	0.3671	1.071	0.923	1.243
Hamlet and isolated dwelling – sparse	–0.113	0.103	–1.099	0.2717	0.893	0.730	1.093	–0.127	0.103	–1.233	0.2177	0.881	0.720	1.078
Urban ≥ 10,000 – less sparse	0.073	0.031	2.326	0.0200	1.076	1.012	1.144	0.069	0.031	2.203	0.0276	1.072	1.008	1.140
Town and fringe – less sparse	0.114	0.034	3.356	0.0008	1.120	1.048	1.197	0.109	0.034	3.231	0.0012	1.116	1.044	1.192
Village – less sparse	0.109	0.036	3.066	0.0022	1.116	1.040	1.196	0.105	0.036	2.947	0.0032	1.111	1.036	1.191
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
SHA														
North East	0.191	0.119	1.614	0.1066	1.211	0.960	1.528	0.189	0.119	1.590	0.1119	1.208	0.957	1.524
North West	0.058	0.092	0.631	0.5277	1.060	0.885	1.269	0.072	0.092	0.779	0.4358	1.074	0.897	1.287
Yorkshire and Humber	0.205	0.097	2.106	0.0352	1.228	1.014	1.486	0.212	0.098	2.174	0.0297	1.237	1.021	1.497
East Midlands	0.225	0.114	1.977	0.0480	1.253	1.002	1.566	0.223	0.114	1.958	0.0502	1.250	1.000	1.562
West Midlands	0.066	0.108	0.610	0.5416	1.068	0.864	1.320	0.064	0.108	0.596	0.5509	1.067	0.863	1.318
East of England	0.109	0.108	1.008	0.3135	1.115	0.902	1.379	0.105	0.108	0.970	0.3323	1.111	0.898	1.373
London	−0.098	0.101	−0.968	0.3329	0.907	0.744	1.105	−0.102	0.101	−1.008	0.3133	0.903	0.741	1.101
South East Coast	−0.019	0.121	−0.154	0.8779	0.982	0.775	1.244	−0.018	0.121	−0.150	0.8809	0.982	0.775	1.245
South Central	−0.051	0.114	−0.442	0.6585	0.951	0.760	1.189	−0.065	0.114	−0.568	0.5702	0.937	0.750	1.172
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
ONS maternities (thousands)	−0.025	0.015	−1.661	0.0966	0.976	0.948	1.004	−0.029	0.015	−1.946	0.0516	0.972	0.944	1.000
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Yes	0.000				1.000			0.000				1.000		
No	0.076	0.056	1.342	0.1796	1.079	0.966	1.205	0.072	0.057	1.280	0.2005	1.075	0.962	1.201
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
OU	−0.001	0.071	−0.016	0.9871	0.999	0.869	1.148	−0.003	0.071	−0.044	0.9650	0.997	0.868	1.146
OU/AMU	−0.077	0.081	−0.947	0.3437	0.926	0.791	1.085	−0.074	0.081	−0.915	0.3604	0.929	0.793	1.088
OU/AMU/FMU	−0.044	0.097	−0.451	0.6517	0.957	0.791	1.158	−0.026	0.097	−0.265	0.7909	0.975	0.805	1.179
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
FTE doctors per 100 maternities	0.106	0.134	0.792	0.4281	1.112	0.855	1.445							
FTE midwives per 100 maternities	0.089	0.062	1.427	0.1536	1.093	0.967	1.234							
FTE support workers per 100 maternities	−0.108	0.071	−1.529	0.1262	0.897	0.781	1.031							
FTE staff per 100 maternities								0.041	0.037	1.094	0.2738	1.042	0.968	1.121
Doctor-to-midwife ratio								0.363	0.390	0.930	0.3526	1.437	0.669	3.088
Support worker-to-midwife ratio								−0.367	0.233	−1.575	0.1152	0.693	0.439	1.094
Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.275	0.011						0.275	0.011					
Mother level	0.306	0.021						0.306	0.021					
Mother level, sociodemographics	0.257	0.018						0.257	0.018					
Mother level, sociodemographics, trust level	0.241	0.017						0.241	0.017					
Mother level, sociodemographics, trust level, staff 1	0.236	0.016												
Mother level, sociodemographics, trust level, staff 2								0.236	0.016					

146

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr > \chi^2$	χ^2	$Pr > \chi^2$
Mother's age group (6 df)	5099.826	< 0.0001	5101.456	< 0.0001
Parity (4 df)	36,597.737	< 0.0001	36,595.628	< 0.0001
Clinical risk (1 df)	23,436.464	< 0.0001	23,435.882	< 0.0001
Ethnicity (16 df)	999.337	< 0.0001	995.113	< 0.0001
IMD (4 df)	413.126	< 0.0001	412.451	< 0.0001
Rural/urban classification (7 df)	26.818	0.0004	25.779	0.0006
SHA (9 df)	18.926	0.0258	19.770	0.0194
ONS maternities (1 df)	2.761	0.0966	3.789	0.0516
University trust (1 df)	1.801	0.1796	1.639	0.2005
Configuration (3 df)	1.653	0.6474	1.459	0.6918
FTE doctors per 100 maternities (1 df)	0.628	0.4281		
FTE midwives per 100 maternities (1 df)	2.037	0.1536		
FTE support workers per 100 maternities (1 df)	2.339	0.1262		
FTE staff per 100 maternities (1 df)			1.198	0.2738
Doctor-to-midwife ratio (1 df)			0.864	0.3526
Support worker-to-midwife ratio (1 df)			2.482	0.1152

AUC	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.747	0.001	0.745	0.749	0.747	0.001	0.745	0.749
Mother level, sociodemographics	0.752	0.001	0.750	0.754	0.752	0.001	0.750	0.754
Mother level, sociodemographics, trust level	0.753	0.001	0.751	0.755	0.753	0.001	0.751	0.755
Mother level, sociodemographics, trust level, staff 1	0.753	0.001	0.751	0.755				
Mother level, sociodemographics, trust level, staff 2					0.753	0.001	0.751	0.755

AUC$_{L95\%}$, lower 95% confidence limit of the AUC; AUC$_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(β), SE of beta coefficient; SE(AUC), SE of area under the curve; SE(σ), SE of sigma coefficient.
Pr > χ^2 refers to the probability of exceeding the χ^2-statistic.
Pr > t refers to the probability of exceeding the t-statistic.

Delivery with bodily integrity: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Intercept	0.178	0.226	0.787	0.4326				0.208	0.247	0.842	0.4013			
Mother's age group (years)														
≤19	1.292	0.073	17.615	<0.0001	3.638	3.151	4.201	1.295	0.073	17.652	<0.0001	3.651	3.162	4.215
20–24	0.971	0.072	13.394	<0.0001	2.641	2.291	3.044	0.974	0.073	13.429	<0.0001	2.649	2.298	3.053
25–29	0.598	0.072	8.262	<0.0001	1.818	1.578	2.095	0.601	0.072	8.300	<0.0001	1.823	1.582	2.101
30–34	0.343	0.072	4.742	<0.0001	1.409	1.223	1.623	0.346	0.072	4.786	<0.0001	1.414	1.227	1.629
35–39	0.173	0.072	2.385	0.0171	1.189	1.031	1.370	0.176	0.072	2.429	0.0151	1.193	1.035	1.375
40–44	0.054	0.074	0.731	0.4645	1.056	0.913	1.220	0.058	0.074	0.785	0.4327	1.060	0.917	1.225
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−2.616	0.014	−181.305	<0.0001	0.073	0.071	0.075	−2.616	0.014	−181.320	<0.0001	0.073	0.071	0.075
1	−1.483	0.014	−108.524	<0.0001	0.227	0.221	0.233	−1.483	0.014	−108.544	<0.0001	0.227	0.221	0.233
2	−0.823	0.014	−57.529	<0.0001	0.439	0.427	0.451	−0.824	0.014	−57.547	<0.0001	0.439	0.427	0.451
3	−0.417	0.016	−25.390	<0.0001	0.659	0.638	0.681	−0.417	0.016	−25.420	<0.0001	0.659	0.638	0.680
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	−0.845	0.006	−132.176	<0.0001	0.430	0.424	0.435	−0.844	0.006	−132.154	<0.0001	0.430	0.424	0.435

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr > t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr > t	OR	OR_L95%	OR_U95%
Not given/not known/not stated	0.029	0.025	1.154	0.2485	1.029	0.980	1.080	0.029	0.025	1.187	0.2353	1.030	0.981	1.081
English/Welsh/Scottish/Northern Irish/British	0.111	0.020	5.643	<0.0001	1.117	1.075	1.161	0.111	0.020	5.625	<0.0001	1.117	1.075	1.161
Irish	−0.083	0.051	−1.629	0.1033	0.920	0.832	1.017	−0.083	0.051	−1.633	0.1024	0.920	0.832	1.017
Any other white background	0.059	0.022	2.713	0.0067	1.061	1.017	1.107	0.058	0.022	2.667	0.0076	1.060	1.016	1.106
White and black Caribbean	0.532	0.046	11.606	<0.0001	1.703	1.556	1.863	0.531	0.046	11.589	<0.0001	1.701	1.555	1.861
White and black African	0.132	0.057	2.317	0.0205	1.141	1.021	1.277	0.135	0.057	2.365	0.0180	1.145	1.023	1.280
White and Asian	−0.042	0.066	−0.632	0.5271	0.959	0.842	1.092	−0.042	0.066	−0.638	0.5232	0.959	0.842	1.092
Any other mixed/multiple ethnic background	0.176	0.045	3.882	0.0001	1.193	1.091	1.304	0.180	0.045	3.961	<0.0001	1.197	1.095	1.308
Indian	−0.521	0.028	−18.609	<0.0001	0.594	0.562	0.628	−0.521	0.028	−18.633	<0.0001	0.594	0.562	0.627
Pakistani	−0.110	0.025	−4.476	<0.0001	0.896	0.854	0.940	−0.111	0.025	−4.525	<0.0001	0.895	0.853	0.939
Bangladeshi	−0.233	0.033	−7.009	<0.0001	0.792	0.742	0.845	−0.233	0.033	−7.011	<0.0001	0.792	0.742	0.845
Chinese	−0.467	0.046	−10.054	<0.0001	0.627	0.572	0.687	−0.466	0.046	−10.032	<0.0001	0.628	0.573	0.687
Any other Asian background	−0.251	0.030	−8.421	<0.0001	0.778	0.734	0.825	−0.252	0.030	−8.467	<0.0001	0.777	0.733	0.824
African	0.028	0.025	1.138	0.2550	1.029	0.980	1.081	0.028	0.025	1.122	0.2620	1.028	0.979	1.080
Caribbean	0.587	0.034	17.190	<0.0001	1.799	1.683	1.924	0.589	0.034	17.238	<0.0001	1.802	1.685	1.927
Any other black/African/Caribbean background	0.244	0.036	6.749	<0.0001	1.276	1.189	1.370	0.243	0.036	6.718	<0.0001	1.275	1.188	1.368
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

| | Model 1: FTE staffing variables | | | | | | | Model 2: staffing ratio variables | | | | | | |
IMD	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
1 = most deprived	0.376	0.012	31.955	<0.0001	1.457	1.424	1.491	0.376	0.012	31.959	<0.0001	1.457	1.424	1.491
2	0.267	0.011	23.386	<0.0001	1.306	1.277	1.336	0.267	0.011	23.401	<0.0001	1.307	1.278	1.336
3	0.175	0.011	15.273	<0.0001	1.191	1.165	1.218	0.175	0.011	15.297	<0.0001	1.191	1.165	1.218
4	0.098	0.012	8.376	<0.0001	1.103	1.078	1.128	0.098	0.012	8.415	<0.0001	1.103	1.078	1.129
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
No information/ other postcode														
Urban ≥10,000 – sparse	0.011	0.077	0.138	0.8899	1.011	0.869	1.176	0.010	0.077	0.133	0.8938	1.010	0.869	1.175
Town and fringe – sparse	0.107	0.057	1.865	0.0622	1.113	0.995	1.245	0.101	0.057	1.762	0.0781	1.106	0.989	1.238
Village – sparse	0.170	0.060	2.809	0.0050	1.185	1.053	1.334	0.167	0.060	2.758	0.0058	1.181	1.049	1.330
Hamlet and isolated dwelling – sparse	–0.050	0.081	–0.620	0.5352	0.951	0.812	1.114	–0.057	0.081	–0.706	0.4804	0.945	0.807	1.106
Urban ≥10,000 – less sparse	0.042	0.024	1.712	0.0869	1.042	0.994	1.093	0.042	0.024	1.721	0.0853	1.043	0.994	1.093
Town and fringe – less sparse	0.068	0.026	2.579	0.0099	1.070	1.016	1.127	0.068	0.026	2.581	0.0099	1.071	1.017	1.127
Village – less sparse	0.057	0.028	2.067	0.0387	1.059	1.003	1.118	0.057	0.028	2.074	0.0381	1.059	1.003	1.118
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

SHA	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	0.078	0.106	0.730	0.4656	1.081	0.877	1.331	0.076	0.106	0.715	0.4747	1.079	0.876	1.329
North West	−0.056	0.082	−0.678	0.4979	0.946	0.805	1.111	−0.047	0.082	−0.575	0.5650	0.954	0.812	1.121
Yorkshire and Humber	0.092	0.089	1.031	0.3025	1.096	0.921	1.305	0.096	0.089	1.079	0.2807	1.101	0.925	1.310
East Midlands	0.102	0.107	0.961	0.3366	1.108	0.899	1.365	0.101	0.106	0.947	0.3435	1.106	0.898	1.362
West Midlands	−0.003	0.089	−0.037	0.9708	0.997	0.837	1.187	−0.002	0.089	−0.027	0.9787	0.998	0.837	1.188
East of England	−0.005	0.092	−0.051	0.9594	0.995	0.832	1.191	−0.009	0.091	−0.095	0.9244	0.991	0.829	1.186
London	−0.180	0.091	−1.978	0.0479	0.835	0.699	0.998	−0.176	0.091	−1.934	0.0531	0.839	0.702	1.002
South East Coast	−0.169	0.098	−1.724	0.0847	0.845	0.697	1.023	−0.169	0.098	−1.732	0.0834	0.844	0.697	1.023
South Central	−0.095	0.099	−0.966	0.3340	0.909	0.749	1.103	−0.104	0.098	−1.056	0.2909	0.901	0.744	1.093
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	−0.025	0.012	−2.042	0.0411	0.975	0.952	0.999	−0.026	0.012	−2.177	0.0295	0.974	0.951	0.997
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	0.074	0.049	1.524	0.1276	1.077	0.979	1.185	0.073	0.049	1.499	0.1339	1.075	0.978	1.183
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
OU	0.009	0.065	0.141	0.8882	1.009	0.889	1.145	0.015	0.065	0.235	0.8146	1.015	0.894	1.153
OU/AMU	−0.022	0.071	−0.315	0.7531	0.978	0.850	1.125	−0.018	0.072	−0.255	0.7990	0.982	0.854	1.130
OU/AMU/FMU	0.048	0.084	0.576	0.5645	1.050	0.890	1.237	0.061	0.084	0.717	0.4731	1.062	0.900	1.253
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	$SE(\beta)$	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	$SE(\beta)$	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
FTE doctors per 100 maternities	0.096	0.115	0.835	0.4035	1.100	0.879	1.378							
FTE midwives per 100 maternities	0.105	0.051	2.055	0.0399	1.110	1.005	1.227							
FTE support workers per 100 maternities	0.000	0.063	−0.006	0.9951	1.000	0.884	1.131							
FTE staff per 100 maternities								0.076	0.031	2.471	0.0135	1.079	1.016	1.147
Doctor-to-midwife ratio								0.139	0.327	0.426	0.6702	1.149	0.606	2.180
Support worker-to-midwife ratio								−0.123	0.189	−0.653	0.5137	0.884	0.610	1.280

Random variation (trust level)	σ	$SE(\sigma)$						σ	$SE(\sigma)$					
Intercept only	0.283	0.006						0.283	0.006					
Mother level	0.297	0.018						0.297	0.018					
Mother level, sociodemographics	0.244	0.015						0.244	0.015					
Mother level, sociodemographics, trust level	0.233	0.014						0.233	0.014					
Mother level, sociodemographics, trust level, staff 1	0.230	0.014												
Mother level, sociodemographics, trust level, staff 2								0.230	0.014					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr > \chi^2$	χ^2	$Pr > \chi^2$
Mothers' age group (6 df)	9408.646	<0.0001	9403.169	<0.0001
Parity (4 df)	56,739.596	<0.0001	56,741.230	<0.0001
Clinical risk (1 df)	17,470.422	<0.0001	17,464.730	<0.0001
Ethnicity (16 df)	1976.990	<0.0001	1981.166	<0.0001
IMD (4 df)	1199.651	<0.0001	1198.775	<0.0001
Rural/urban classification (7 df)	16.701	0.0194	16.431	0.0215
SHA (9 df)	18.824	0.0267	18.652	0.0283
ONS maternities (1 df)	4.172	0.0411	4.739	0.0295
University trust (1 df)	2.321	0.1276	2.246	0.1339
Configuration (3 df)	1.030	0.7940	1.286	0.7324
FTE doctors per 100 maternities (1 df)	0.698	0.4035		
FTE midwives per 100 maternities (1 df)	4.223	0.0399		
FTE support workers per 100 maternities (1 df)	0.000	0.9951		
FTE staff per 100 maternities (1 df)			6.106	0.0135
Doctor-to-midwife ratio (1 df)			0.181	0.6702
Support worker-to-midwife ratio (1 df)			0.426	0.5137

AUC	Model 1: FTE staffing variables					Model 2: staffing ratio variables				
	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$		AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$	
Intercept only	0.500	0.000	0.500	0.500		0.500	0.000	0.500	0.500	
Mother level	0.724	0.001	0.723	0.726		0.724	0.001	0.723	0.726	
Mother level, sociodemographics	0.732	0.001	0.730	0.733		0.732	0.001	0.730	0.733	
Mother level, sociodemographics, trust level	0.732	0.001	0.731	0.734		0.732	0.001	0.731	0.734	
Mother level, sociodemographics, trust level, staff 1	0.733	0.001	0.731	0.734						
Mother level, sociodemographics, trust level, staff 2						0.733	0.001	0.731	0.734	

AUC$_{L95\%}$, lower 95% confidence limit of the AUC; AUC$_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.
Pr > χ^2 refers to the probability of exceeding the χ^2-statistic.
Pr > t refers to the probability of exceeding the t-statistic.

Normal birth: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Intercept	0.509	0.222	2.296	0.0238				0.540	0.247	2.184	0.0313			
Mother's age group (years)														
≤19	1.415	0.090	15.641	< 0.0001	4.116	3.448	4.915	1.416	0.090	15.649	< 0.0001	4.121	3.452	4.921
20–24	1.179	0.090	13.156	< 0.0001	3.251	2.727	3.875	1.180	0.090	13.166	< 0.0001	3.255	2.731	3.880
25–29	0.986	0.089	11.029	< 0.0001	2.682	2.250	3.196	0.988	0.089	11.040	< 0.0001	2.685	2.253	3.200
30–34	0.830	0.089	9.289	< 0.0001	2.294	1.926	2.734	0.832	0.089	9.301	< 0.0001	2.297	1.928	2.738
35–39	0.685	0.090	7.653	< 0.0001	1.985	1.665	2.365	0.687	0.090	7.667	< 0.0001	1.987	1.667	2.369
40–44	0.419	0.091	4.604	< 0.0001	1.521	1.272	1.818	0.421	0.091	4.618	< 0.0001	1.523	1.274	1.821
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−1.751	0.016	−108.775	< 0.0001	0.174	0.168	0.179	−1.751	0.016	−108.755	< 0.0001	0.174	0.168	0.179
1	−0.607	0.015	−39.244	< 0.0001	0.545	0.529	0.562	−0.607	0.015	−39.228	< 0.0001	0.545	0.529	0.562
2	−0.351	0.016	−21.290	< 0.0001	0.704	0.682	0.727	−0.350	0.016	−21.274	< 0.0001	0.705	0.682	0.728
3	−0.199	0.019	−10.438	< 0.0001	0.820	0.790	0.851	−0.199	0.019	−10.433	< 0.0001	0.820	0.790	0.851
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	−1.822	0.007	−251.057	< 0.0001	0.162	0.159	0.164	−1.822	0.007	−251.052	< 0.0001	0.162	0.159	0.164

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Not given/not known/not stated	0.010	0.027	0.361	0.7178	1.010	0.958	1.065	0.010	0.027	0.373	0.7089	1.010	0.958	1.065
English/Welsh/Scottish/Northern Irish/British	−0.108	0.022	−4.872	<0.0001	0.898	0.860	0.938	−0.107	0.022	−4.865	<0.0001	0.898	0.860	0.938
Irish	−0.314	0.056	−5.628	<0.0001	0.731	0.655	0.815	−0.314	0.056	−5.630	<0.0001	0.730	0.655	0.815
Any other white background	−0.025	0.024	−1.012	0.3114	0.976	0.930	1.023	−0.024	0.024	−1.010	0.3127	0.976	0.931	1.023
White and black Caribbean	0.159	0.053	2.976	0.0029	1.172	1.056	1.301	0.159	0.053	2.978	0.0029	1.172	1.056	1.301
White and black African	0.129	0.064	2.019	0.0435	1.138	1.004	1.290	0.129	0.064	2.024	0.0430	1.138	1.004	1.290
White and Asian	−0.106	0.073	−1.455	0.1458	0.900	0.780	1.037	−0.106	0.073	−1.455	0.1457	0.900	0.780	1.037
Any other mixed/multiple ethnic background	−0.018	0.051	−0.360	0.7187	0.982	0.889	1.084	−0.018	0.051	−0.359	0.7194	0.982	0.889	1.084
Indian	−0.205	0.029	−7.031	<0.0001	0.815	0.770	0.863	−0.205	0.029	−7.044	<0.0001	0.815	0.770	0.863
Pakistani	0.081	0.028	2.916	0.0035	1.084	1.027	1.145	0.081	0.028	2.913	0.0036	1.084	1.027	1.145
Bangladeshi	0.107	0.036	2.984	0.0028	1.113	1.037	1.193	0.107	0.036	2.979	0.0029	1.112	1.037	1.193
Chinese	0.182	0.045	4.048	<0.0001	1.199	1.098	1.310	0.182	0.045	4.052	<0.0001	1.200	1.099	1.310
Any other Asian background	−0.042	0.032	−1.316	0.1883	0.959	0.900	1.021	−0.042	0.032	−1.310	0.1902	0.959	0.900	1.021
African	0.036	0.028	1.284	0.1990	1.037	0.981	1.096	0.036	0.028	1.287	0.1981	1.037	0.981	1.096
Caribbean	0.223	0.040	5.547	<0.0001	1.249	1.155	1.352	0.223	0.040	5.553	<0.0001	1.250	1.155	1.352
Any other black/African/Caribbean background	0.099	0.044	2.262	0.0237	1.104	1.013	1.202	0.099	0.044	2.260	0.0238	1.104	1.013	1.202
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

continued

IMD	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
1 = most deprived	0.118	0.013	9.277	< 0.0001	1.125	1.098	1.154	0.118	0.013	9.284	< 0.0001	1.126	1.098	1.154
2	0.068	0.012	5.522	< 0.0001	1.071	1.045	1.097	0.068	0.012	5.525	< 0.0001	1.071	1.045	1.097
3	0.022	0.012	1.814	0.0697	1.023	0.998	1.048	0.022	0.012	1.822	0.0684	1.023	0.998	1.048
4	0.010	0.012	0.790	0.4293	1.010	0.986	1.035	0.010	0.012	0.797	0.4256	1.010	0.986	1.035
5 = least deprived	0.000				1.000			0.000				1.000		

Rural/urban classification	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
No information/other postcode														
Urban ≥ 10,000 – sparse	−0.083	0.084	−0.982	0.3263	0.921	0.780	1.086	−0.084	0.084	−0.997	0.3187	0.919	0.779	1.085
Town and fringe – sparse	0.108	0.063	1.732	0.0833	1.114	0.986	1.260	0.108	0.063	1.726	0.0843	1.114	0.985	1.259
Village – sparse	0.199	0.066	3.034	0.0024	1.221	1.073	1.388	0.199	0.066	3.030	0.0024	1.220	1.073	1.388
Hamlet and isolated dwelling – sparse	−0.049	0.089	−0.548	0.5835	0.953	0.800	1.133	−0.050	0.089	−0.566	0.5712	0.951	0.799	1.132
Urban ≥ 10,000 – less sparse	−0.024	0.026	−0.925	0.3552	0.976	0.927	1.028	−0.024	0.026	−0.922	0.3566	0.976	0.927	1.028
Town and fringe – less sparse	−0.001	0.029	−0.047	0.9627	0.999	0.944	1.056	−0.001	0.029	−0.039	0.9693	0.999	0.945	1.056
Village – less sparse	0.006	0.030	0.201	0.8409	1.006	0.949	1.067	0.006	0.030	0.203	0.8395	1.006	0.949	1.067
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
SHA	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	0.072	0.099	0.726	0.4679	1.074	0.885	1.304	0.066	0.099	0.666	0.5053	1.068	0.880	1.297
North West	0.112	0.076	1.481	0.1387	1.119	0.964	1.298	0.117	0.076	1.540	0.1236	1.124	0.969	1.305
Yorkshire and Humber	0.130	0.082	1.592	0.1115	1.139	0.970	1.338	0.133	0.082	1.612	0.1069	1.142	0.972	1.342
East Midlands	0.236	0.095	2.468	0.0136	1.266	1.050	1.526	0.235	0.096	2.460	0.0139	1.265	1.049	1.526
West Midlands	0.179	0.087	2.051	0.0403	1.196	1.008	1.418	0.180	0.087	2.056	0.0398	1.197	1.008	1.421
East of England	0.128	0.086	1.491	0.1359	1.136	0.961	1.344	0.128	0.086	1.490	0.1361	1.136	0.961	1.344
London	−0.033	0.088	−0.379	0.7045	0.967	0.814	1.149	−0.034	0.088	−0.386	0.6998	0.967	0.814	1.148
South East Coast	0.067	0.099	0.680	0.4968	1.069	0.881	1.297	0.070	0.099	0.704	0.4817	1.072	0.883	1.302
South Central	0.184	0.100	1.844	0.0652	1.202	0.988	1.462	0.177	0.100	1.779	0.0752	1.194	0.982	1.452
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	−0.017	0.011	−1.453	0.1461	0.984	0.962	1.006	−0.017	0.011	−1.534	0.1250	0.983	0.961	1.005
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	0.042	0.046	0.925	0.3549	1.043	0.954	1.140	0.045	0.046	0.988	0.3232	1.046	0.957	1.144
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
OU	−0.122	0.058	−2.094	0.0362	0.885	0.789	0.992	−0.118	0.059	−1.997	0.0459	0.889	0.791	0.998
OU/AMU	−0.076	0.067	−1.132	0.2578	0.927	0.813	1.057	−0.071	0.067	−1.046	0.2955	0.932	0.817	1.064
OU/AMU/FMU	−0.120	0.081	−1.487	0.1370	0.887	0.758	1.039	−0.110	0.081	−1.352	0.1764	0.896	0.764	1.051
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
FTE doctors per 100 maternities	−0.075	0.113	−0.658	0.5107	0.928	0.743	1.159							
FTE midwives per 100 maternities	0.060	0.048	1.268	0.2048	1.062	0.968	1.166							
FTE support workers per 100 maternities	0.010	0.060	0.175	0.8614	1.011	0.898	1.137							
FTE staff per 100 maternities								0.028	0.029	0.949	0.3424	1.028	0.971	1.088
Doctor-to-midwife ratio								−0.164	0.326	−0.503	0.6150	0.849	0.448	1.608
Support worker-to-midwife ratio								0.030	0.177	0.170	0.8649	1.031	0.729	1.457

Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.270	0.012						0.270	0.012					
Mother level	0.225	0.015						0.225	0.015					
Mother level, sociodemographics	0.209	0.014						0.209	0.014					
Mother level, sociodemographics, trust level	0.199	0.013						0.199	0.013					
Mother level, sociodemographics, trust level, staff 1	0.198	0.013												
Mother level, sociodemographics, trust level, staff 2								0.198	0.013					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr > \chi^2$	χ^2	$Pr > \chi^2$
Mother's age group (6 df)	3455.872	< 0.0001	3454.600	< 0.0001
Parity (4 df)	28,952.972	< 0.0001	28,948.399	< 0.0001
Clinical risk (1 df)	63,029.595	< 0.0001	63,027.235	< 0.0001
Ethnicity (16 df)	452.266	< 0.0001	452.348	< 0.0001
IMD (4 df)	126.536	< 0.0001	126.569	< 0.0001
Rural/urban classification (7 df)	23.057	0.0017	23.067	0.0017
SHA (9 df)	15.832	0.0705	15.752	0.0723
ONS maternities (1 df)	2.112	0.1461	2.354	0.1250
University trust (1 df)	0.856	0.3549	0.976	0.3232
Configuration (3 df)	4.877	0.1810	4.403	0.2211
FTE doctors per 100 maternities (1 df)	0.433	0.5107		
FTE midwives per 100 maternities (1 df)	1.608	0.2048		
FTE support workers per 100 maternities (1 df)	0.030	0.8614		
FTE staff per 100 maternities (1 df)			0.901	0.3424
Doctor-to-midwife ratio (1 df)			0.253	0.6150
Support worker-to-midwife ratio (1 df)			0.029	0.8649

AUC	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.754	0.001	0.752	0.755	0.754	0.001	0.752	0.755
Mother level, sociodemographics	0.756	0.001	0.754	0.757	0.756	0.001	0.754	0.757
Mother level, sociodemographics, trust level	0.756	0.001	0.755	0.758	0.756	0.001	0.755	0.758
Mother level, sociodemographics, trust level, staff 1	0.756	0.001	0.755	0.758				
Mother level, sociodemographics, trust level, staff 2					0.756	0.001	0.755	0.758

AUC$_{L95\%}$, lower 95% confidence limit of the AUC; AUC$_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.
Pr > χ² refers to the probability of exceeding the χ²-statistic.
Pr > t refers to the probability of exceeding the t-statistic.

Spontaneous vaginal delivery: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Intercept	1.250	0.181	6.908	<0.0001				1.305	0.197	6.609	<0.0001			
Mother's age group (years)														
≤19	1.771	0.065	27.264	<0.0001	5.877	5.175	6.676	1.772	0.065	27.270	<0.0001	5.881	5.178	6.679
20–24	1.406	0.064	22.008	<0.0001	4.079	3.599	4.623	1.407	0.064	22.018	<0.0001	4.082	3.602	4.626
25–29	1.088	0.064	17.092	<0.0001	2.969	2.621	3.364	1.089	0.064	17.104	<0.0001	2.972	2.623	3.367
30–34	0.842	0.064	13.236	<0.0001	2.321	2.049	2.629	0.843	0.064	13.249	<0.0001	2.323	2.051	2.631
35–39	0.602	0.064	9.448	<0.0001	1.826	1.612	2.069	0.603	0.064	9.462	<0.0001	1.828	1.613	2.071
40–44	0.388	0.065	5.965	<0.0001	1.474	1.298	1.674	0.389	0.065	5.973	<0.0001	1.475	1.298	1.675
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−1.610	0.015	−108.540	<0.0001	0.200	0.194	0.206	−1.610	0.015	−108.531	<0.0001	0.200	0.194	0.206
1	−0.670	0.015	−45.726	<0.0001	0.512	0.497	0.527	−0.669	0.015	−45.715	<0.0001	0.512	0.497	0.527
2	−0.401	0.016	−25.593	<0.0001	0.670	0.649	0.690	−0.401	0.016	−25.579	<0.0001	0.670	0.649	0.691
3	−0.225	0.018	−12.377	<0.0001	0.798	0.770	0.827	−0.225	0.018	−12.385	<0.0001	0.798	0.770	0.827
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	−1.417	0.006	−222.673	<0.0001	0.243	0.240	0.246	−1.417	0.006	−222.679	<0.0001	0.243	0.240	0.246

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Not given/not known/not stated	0.044	0.023	1.948	0.0514	1.045	1.000	1.093	0.044	0.023	1.947	0.0515	1.045	1.000	1.093
English/Welsh/Scottish/Northern Irish/British	−0.001	0.018	−0.059	0.9527	0.999	0.964	1.035	−0.001	0.018	−0.064	0.9492	0.999	0.964	1.035
Irish	−0.161	0.045	−3.550	0.0004	0.851	0.778	0.930	−0.161	0.045	−3.541	0.0004	0.851	0.779	0.931
Any other white background	0.044	0.020	2.165	0.0304	1.045	1.004	1.087	0.044	0.020	2.169	0.0301	1.045	1.004	1.087
White and black Caribbean	0.265	0.048	5.470	<0.0001	1.304	1.185	1.433	0.266	0.048	5.479	<0.0001	1.304	1.186	1.434
White and black African	0.034	0.056	0.600	0.5484	1.034	0.927	1.154	0.032	0.056	0.575	0.5653	1.033	0.925	1.153
White and Asian	−0.135	0.061	−2.206	0.0274	0.874	0.775	0.985	−0.134	0.061	−2.195	0.0282	0.874	0.776	0.986
Any other mixed/multiple ethnic background	0.047	0.044	1.080	0.2799	1.048	0.962	1.142	0.047	0.044	1.079	0.2808	1.048	0.962	1.142
Indian	−0.187	0.024	−7.732	<0.0001	0.830	0.791	0.870	−0.187	0.024	−7.739	<0.0001	0.829	0.791	0.870
Pakistani	0.069	0.024	2.932	0.0034	1.072	1.023	1.123	0.069	0.024	2.928	0.0034	1.072	1.023	1.122
Bangladeshi	0.059	0.032	1.845	0.0650	1.061	0.996	1.130	0.059	0.032	1.838	0.0660	1.061	0.996	1.130
Chinese	0.111	0.040	2.807	0.0050	1.118	1.034	1.208	0.111	0.040	2.797	0.0052	1.117	1.034	1.207
Any other Asian background	−0.087	0.027	−3.262	0.0011	0.916	0.869	0.966	−0.088	0.027	−3.277	0.0011	0.916	0.869	0.965
African	−0.039	0.024	−1.651	0.0986	0.962	0.918	1.007	−0.039	0.024	−1.671	0.0948	0.961	0.918	1.007
Caribbean	0.188	0.034	5.456	<0.0001	1.207	1.128	1.291	0.188	0.034	5.453	<0.0001	1.207	1.128	1.291
Any other black/African/Caribbean background	0.125	0.035	3.540	0.0004	1.134	1.058	1.215	0.125	0.035	3.537	0.0004	1.133	1.057	1.215
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
IMD	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
1 = most deprived	0.096	0.011	8.638	<0.0001	1.100	1.077	1.124	0.096	0.011	8.640	<0.0001	1.100	1.077	1.124
2	0.059	0.011	5.551	<0.0001	1.060	1.039	1.083	0.059	0.011	5.556	<0.0001	1.061	1.039	1.083
3	0.029	0.010	2.811	0.0049	1.030	1.009	1.051	0.030	0.010	2.824	0.0047	1.030	1.009	1.051
4	0.018	0.011	1.732	0.0833	1.018	0.998	1.040	0.019	0.011	1.764	0.0778	1.019	0.998	1.040
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
No information/other postcode														
Urban ≥10,000 – sparse	−0.031	0.079	−0.387	0.6986	0.970	0.831	1.132	−0.031	0.079	−0.388	0.6977	0.970	0.831	1.132
Town and fringe – sparse	0.089	0.057	1.552	0.1207	1.093	0.977	1.224	0.090	0.057	1.559	0.1189	1.094	0.977	1.224
Village – sparse	0.218	0.061	3.573	0.0004	1.244	1.104	1.402	0.219	0.061	3.582	0.0003	1.245	1.104	1.403
Hamlet and isolated dwelling – sparse	−0.073	0.078	−0.936	0.3494	0.930	0.798	1.083	−0.072	0.078	−0.916	0.3594	0.931	0.799	1.085
Urban ≥10,000 – less sparse	0.010	0.023	0.444	0.6573	1.010	0.966	1.056	0.011	0.023	0.468	0.6397	1.011	0.967	1.056
Town and fringe – less sparse	0.052	0.025	2.122	0.0338	1.054	1.004	1.106	0.053	0.025	2.131	0.0331	1.054	1.004	1.106
Village – less sparse	0.053	0.026	2.047	0.0407	1.054	1.002	1.109	0.052	0.026	2.024	0.0430	1.054	1.002	1.109
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
SHA	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	−0.039	0.084	−0.467	0.6405	0.962	0.816	1.133	−0.043	0.084	−0.510	0.6102	0.958	0.813	1.129
North West	0.092	0.065	1.413	0.1576	1.096	0.965	1.244	0.090	0.065	1.389	0.1647	1.094	0.964	1.242
Yorkshire and Humber	0.047	0.070	0.678	0.4978	1.049	0.914	1.203	0.044	0.070	0.625	0.5322	1.045	0.911	1.199
East Midlands	0.062	0.084	0.735	0.4626	1.063	0.902	1.253	0.057	0.084	0.677	0.4984	1.058	0.898	1.247
West Midlands	0.028	0.070	0.397	0.6915	1.028	0.896	1.180	0.024	0.070	0.340	0.7337	1.024	0.892	1.175
East of England	−0.027	0.072	−0.373	0.7092	0.973	0.845	1.121	−0.030	0.072	−0.423	0.6724	0.970	0.842	1.117
London	−0.099	0.071	−1.391	0.1642	0.905	0.787	1.042	−0.103	0.071	−1.435	0.1512	0.902	0.785	1.038
South East Coast	−0.045	0.077	−0.582	0.5607	0.956	0.823	1.112	−0.049	0.077	−0.637	0.5242	0.952	0.819	1.107
South Central	−0.040	0.078	−0.519	0.6041	0.960	0.825	1.119	−0.044	0.077	−0.564	0.5731	0.957	0.823	1.114
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	−0.011	0.010	−1.199	0.230	0.989	0.970	1.007	−0.012	0.010	−1.290	0.197	0.988	0.970	1.006
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	0.086	0.038	2.260	0.024	1.090	1.012	1.175	0.086	0.038	2.251	0.024	1.090	1.011	1.174
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
OU	−0.030	0.051	−0.583	0.560	0.971	0.879	1.072	−0.031	0.051	−0.613	0.540	0.969	0.877	1.071
OU/AMU	−0.020	0.056	−0.351	0.726	0.980	0.878	1.094	−0.022	0.056	−0.392	0.695	0.978	0.876	1.092
OU/AMU/FMU	−0.012	0.066	−0.186	0.852	0.988	0.868	1.124	−0.015	0.066	−0.223	0.823	0.985	0.865	1.122
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	$SE(\beta)$	t-statistic	Pr>t	OR	$OR_{L95\%}$	$OR_{U95\%}$	β	$SE(\beta)$	t-statistic	Pr>t	OR	$OR_{L95\%}$	$OR_{U95\%}$
FTE doctors per 100 maternities	0.017	0.090	0.192	0.8480	1.017	0.853	1.214							
FTE midwives per 100 maternities	0.025	0.040	0.619	0.5362	1.025	0.948	1.109							
FTE support workers per 100 maternities	−0.039	0.049	−0.783	0.4334	0.962	0.873	1.060							
FTE staff per 100 maternities								0.009	0.024	0.367	0.7136	1.009	0.962	1.058
Doctor-to-midwife ratio								0.018	0.257	0.070	0.9441	1.018	0.615	1.685
Support worker-to-midwife ratio								−0.133	0.148	−0.895	0.3706	0.876	0.655	1.171
Random variation (trust level)	σ	$SE(\sigma)$						σ	$SE(\sigma)$					
Intercept only	0.226	0.006						0.226	0.006					
Mother level	0.200	0.012						0.200	0.012					
Mother level, sociodemographics	0.186	0.011						0.186	0.011					
Mother level, sociodemographics, trust level	0.180	0.011						0.180	0.011					
Mother level, sociodemographics, trust level, staff 1	0.179	0.011												
Mother level, sociodemographics, trust level, staff 2								0.179	0.011					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr>\chi^2$	χ^2	$Pr>\chi^2$
Mother's age group (6 df)	10,476.241	< 0.0001	10,473.595	< 0.0001
Parity (4 df)	30,418.259	< 0.0001	30,416.199	< 0.0001
Clinical risk (1 df)	49,583.344	< 0.0001	49,585.788	< 0.0001
Ethnicity (16 df)	323.687	< 0.0001	324.110	< 0.0001
IMD (4 df)	91.541	< 0.0001	91.300	< 0.0001
Rural/urban classification (7 df)	33.710	< 0.0001	33.225	< 0.0001
SHA (9 df)	12.780	0.1728	13.018	0.1618
ONS maternities (1 df)	1.438	0.2304	1.663	0.1971
University trust (1 df)	5.109	0.0238	5.067	0.0244
Configuration (3 df)	0.381	0.9440	0.409	0.9384
FTE doctors per 100 maternities (1 df)	0.037	0.8480		
FTE midwives per 100 maternities (1 df)	0.383	0.5362		
FTE support workers per 100 maternities (1 df)	0.614	0.4334		
FTE staff per 100 maternities (1 df)			0.135	0.7136
Doctor-to-midwife ratio (1 df)			0.005	0.9441
Support worker-to-midwife ratio (1 df)			0.802	0.3706

AUC	Model 1: FTE staffing variables					Model 2: staffing ratio variables				
	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$		AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$	
Intercept only	0.500	0.000	0.500	0.500		0.500	0.000	0.500	0.500	
Mother level	0.722	0.001	0.721	0.723		0.722	0.001	0.721	0.723	
Mother level, sociodemographics	0.724	0.001	0.723	0.725		0.724	0.001	0.723	0.725	
Mother level, sociodemographics, trust level	0.724	0.001	0.723	0.725		0.724	0.001	0.723	0.725	
Mother level, sociodemographics, trust level, staff 1	0.724	0.001	0.723	0.725						
Mother level, sociodemographics, trust level, staff 2						0.724	0.001	0.723	0.725	

AUC$_{L95\%}$, lower 95% confidence limit of the AUC; AUC$_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.
Pr > χ² refers to the probability of exceeding the χ²-statistic.
Pr > t refers to the probability of exceeding the t-statistic.

Intact perineum: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Intercept	1.079	0.263	4.103	<0.0001				1.138	0.287	3.966	0.0001			
Mother's age group (years)														
≤19	0.626	0.098	6.409	<0.0001	1.871	1.545	2.266	0.628	0.098	6.427	<0.0001	1.874	1.547	2.270
20–24	0.375	0.097	3.863	0.0001	1.455	1.203	1.760	0.377	0.097	3.883	0.0001	1.458	1.205	1.764
25–29	0.027	0.097	0.281	0.7785	1.028	0.850	1.243	0.029	0.097	0.300	0.7644	1.029	0.851	1.245
30–34	−0.190	0.097	−1.956	0.0505	0.827	0.684	1.000	−0.188	0.097	−1.938	0.0526	0.829	0.685	1.002
35–39	−0.272	0.097	−2.800	0.0051	0.762	0.630	0.922	−0.270	0.097	−2.781	0.0054	0.763	0.631	0.923
40–44	−0.267	0.099	−2.702	0.0069	0.766	0.631	0.929	−0.265	0.099	−2.683	0.0073	0.767	0.632	0.931
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−2.913	0.018	−159.009	<0.0001	0.054	0.052	0.056	−2.913	0.018	−159.011	<0.0001	0.054	0.052	0.056
1	−1.772	0.018	−99.665	<0.0001	0.170	0.164	0.176	−1.772	0.018	−99.670	<0.0001	0.170	0.164	0.176
2	−1.007	0.019	−54.071	<0.0001	0.365	0.352	0.379	−1.007	0.019	−54.075	<0.0001	0.365	0.352	0.379
3	−0.524	0.021	−24.515	<0.0001	0.592	0.568	0.618	−0.524	0.021	−24.522	<0.0001	0.592	0.568	0.617
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	−0.212	0.007	−30.747	<0.0001	0.809	0.798	0.820	−0.212	0.007	−30.745	<0.0001	0.809	0.798	0.820

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Not given/not known/not stated	0.011	0.027	0.431	0.6666	1.012	0.960	1.066	0.011	0.027	0.428	0.6683	1.011	0.960	1.066
English/Welsh/Scottish/Northern Irish/British	0.120	0.022	5.581	<0.0001	1.128	1.081	1.176	0.120	0.022	5.589	<0.0001	1.128	1.081	1.176
Irish	−0.045	0.056	−0.800	0.4236	0.956	0.856	1.068	−0.044	0.056	−0.786	0.4318	0.957	0.857	1.068
Any other white background	0.044	0.024	1.856	0.0635	1.045	0.998	1.095	0.045	0.024	1.892	0.0584	1.046	0.998	1.096
White and black Caribbean	0.602	0.052	11.639	<0.0001	1.826	1.650	2.021	0.603	0.052	11.644	<0.0001	1.827	1.651	2.022
White and black African	0.184	0.064	2.855	0.0043	1.202	1.059	1.363	0.185	0.064	2.877	0.0040	1.203	1.061	1.365
White and Asian	−0.026	0.072	−0.366	0.7146	0.974	0.845	1.122	−0.026	0.072	−0.354	0.7230	0.975	0.846	1.123
Any other mixed/multiple ethnic background	0.221	0.050	4.405	<0.0001	1.247	1.130	1.376	0.220	0.050	4.392	<0.0001	1.246	1.130	1.375
Indian	−0.572	0.030	−18.971	<0.0001	0.564	0.532	0.599	−0.573	0.030	−18.985	<0.0001	0.564	0.532	0.598
Pakistani	−0.224	0.027	−8.291	<0.0001	0.800	0.759	0.843	−0.223	0.027	−8.265	<0.0001	0.800	0.759	0.844
Bangladeshi	−0.365	0.037	−9.928	<0.0001	0.694	0.646	0.746	−0.365	0.037	−9.927	<0.0001	0.694	0.646	0.746
Chinese	−0.574	0.049	−11.807	<0.0001	0.563	0.512	0.620	−0.574	0.049	−11.803	<0.0001	0.563	0.512	0.620
Any other Asian background	−0.284	0.032	−8.785	<0.0001	0.753	0.706	0.802	−0.285	0.032	−8.801	<0.0001	0.752	0.706	0.801
African	0.173	0.028	6.136	<0.0001	1.189	1.125	1.257	0.174	0.028	6.148	<0.0001	1.189	1.125	1.257
Caribbean	0.807	0.040	20.160	<0.0001	2.242	2.072	2.425	0.809	0.040	20.191	<0.0001	2.245	2.075	2.428
Any other black/African/Caribbean background	0.373	0.041	9.016	<0.0001	1.452	1.339	1.575	0.373	0.041	9.022	<0.0001	1.452	1.339	1.575
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
IMD	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
1 = most deprived	0.436	0.013	33.937	<0.0001	1.546	1.508	1.585	0.436	0.013	33.923	<0.0001	1.546	1.507	1.585
2	0.312	0.012	25.140	<0.0001	1.366	1.333	1.400	0.312	0.012	25.133	<0.0001	1.366	1.333	1.400
3	0.194	0.012	15.645	<0.0001	1.214	1.185	1.244	0.194	0.012	15.644	<0.0001	1.214	1.185	1.244
4	0.117	0.013	9.282	<0.0001	1.124	1.097	1.152	0.117	0.013	9.279	<0.0001	1.124	1.096	1.152
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
No information/other postcode														
Urban ≥ 10,000 – sparse	0.012	0.086	0.142	0.8871	1.012	0.856	1.197	0.009	0.086	0.104	0.9174	1.009	0.853	1.193
Town and fringe – sparse	0.107	0.063	1.691	0.0908	1.113	0.983	1.260	0.105	0.063	1.659	0.0972	1.110	0.981	1.257
Village – sparse	0.154	0.067	2.313	0.0207	1.167	1.024	1.330	0.154	0.067	2.310	0.0209	1.167	1.024	1.330
Hamlet and isolated dwelling – sparse	−0.049	0.089	−0.550	0.5826	0.952	0.800	1.134	−0.050	0.089	−0.558	0.5766	0.952	0.799	1.133
Urban ≥ 10,000 – less sparse	0.047	0.026	1.765	0.0776	1.048	0.995	1.103	0.046	0.026	1.756	0.0791	1.047	0.995	1.103
Town and fringe – less sparse	0.067	0.029	2.321	0.0203	1.069	1.010	1.131	0.067	0.029	2.321	0.0203	1.069	1.010	1.131
Village – less sparse	0.058	0.030	1.936	0.0529	1.060	0.999	1.125	0.058	0.030	1.931	0.0535	1.060	0.999	1.125
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

SHA	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
---	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	0.100	0.121	0.827	0.4085	1.105	0.872	1.402	0.097	0.121	0.797	0.4254	1.102	0.868	1.397
North West	−0.111	0.094	−1.182	0.2372	0.895	0.745	1.076	−0.103	0.094	−1.102	0.2706	0.902	0.750	1.084
Yorkshire and Humber	0.089	0.101	0.874	0.3822	1.093	0.896	1.333	0.093	0.102	0.914	0.3610	1.097	0.899	1.339
East Midlands	0.074	0.122	0.611	0.5415	1.077	0.849	1.367	0.072	0.121	0.590	0.5553	1.074	0.847	1.363
West Midlands	−0.013	0.102	−0.131	0.8960	0.987	0.808	1.205	−0.012	0.102	−0.114	0.9094	0.988	0.809	1.207
East of England	0.023	0.104	0.224	0.8224	1.024	0.834	1.256	0.021	0.104	0.200	0.8415	1.021	0.832	1.253
London	−0.157	0.104	−1.511	0.1307	0.855	0.698	1.048	−0.152	0.104	−1.467	0.1425	0.859	0.701	1.053
South East Coast	−0.183	0.112	−1.643	0.1003	0.832	0.669	1.036	−0.184	0.112	−1.651	0.0988	0.832	0.669	1.035
South Central	−0.108	0.113	−0.961	0.3367	0.897	0.720	1.119	−0.116	0.112	−1.032	0.3019	0.891	0.715	1.110
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	−0.030	0.014	−2.126	0.0335	0.971	0.945	0.998	−0.031	0.014	−2.221	0.0263	0.970	0.944	0.996
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	0.063	0.055	1.130	0.2587	1.065	0.955	1.187	0.063	0.055	1.140	0.2544	1.065	0.956	1.187
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
OU	0.030	0.074	0.407	0.6838	1.030	0.892	1.190	0.033	0.074	0.449	0.6532	1.034	0.894	1.196
OU/AMU	−0.037	0.081	−0.455	0.6490	0.964	0.822	1.130	−0.037	0.082	−0.456	0.6483	0.963	0.821	1.131
OU/AMU/FMU	0.069	0.096	0.723	0.4697	1.072	0.888	1.293	0.078	0.096	0.810	0.4181	1.081	0.895	1.306
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
FTE doctors per 100 maternities	0.068	0.131	0.521	0.6022	1.070	0.829	1.383							
FTE midwives per 100 maternities	0.124	0.058	2.139	0.0324	1.132	1.010	1.268							
FTE support workers per 100 maternities	0.017	0.072	0.241	0.8095	1.017	0.884	1.171							
FTE staff per 100 maternities								0.088	0.035	2.491	0.0127	1.092	1.019	1.170
Doctor-to-midwife ratio								0.001	0.373	0.002	0.9981	1.001	0.482	2.078
Support worker-to-midwife ratio								−0.094	0.216	−0.434	0.6640	0.911	0.597	1.389

Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.283	0.006						0.283	0.006					
Mother level	0.334	0.020						0.334	0.020					
Mother level, sociodemographics	0.280	0.017						0.280	0.017					
Mother level, sociodemographics, trust level	0.267	0.016						0.267	0.016					
Mother level, sociodemographics, trust level, staff 1	0.263	0.016												
Mother level, sociodemographics, trust level, staff 2								0.263	0.016					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr > \chi^2$	χ^2	$Pr > \chi^2$
Mothers' age group (6 df)	5489.165	< 0.0001	5489.377	< 0.0001
Parity (4 df)	53,241.997	< 0.0001	53,242.208	< 0.0001
Clinical risk (1 df)	945.367	< 0.0001	945.265	< 0.0001
Ethnicity (16 df)	2521.852	< 0.0001	2524.676	< 0.0001
IMD (4 df)	1347.772	< 0.0001	1346.578	< 0.0001
Rural/urban classification (7 df)	11.543	0.1166	11.555	0.1162
SHA (9 df)	16.128	0.0643	15.755	0.0722
ONS maternities (1 df)	4.522	0.0335	4.935	0.0263
University trust (1 df)	1.276	0.2587	1.299	0.2544
Configuration (3 df)	2.187	0.5346	2.519	0.4718
FTE doctors per 100 maternities (1 df)	0.272	0.6022		
FTE midwives per 100 maternities (1 df)	4.577	0.0324		
FTE support workers per 100 maternities (1 df)	0.058	0.8095		
FTE staff per 100 maternities (1 df)			6.204	0.0127
Doctor-to-midwife ratio (1 df)			0.000	0.9981
Support worker-to-midwife ratio (1 df)			0.189	0.6640

AUC	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
	AUC	SE(AUC)	$AUC_{L95\%}$	$AUC_{U95\%}$	AUC	SE(AUC)	$AUC_{L95\%}$	$AUC_{U95\%}$
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.722	0.001	0.720	0.723	0.722	0.001	0.720	0.723
Mother level, sociodemographics	0.731	0.001	0.730	0.733	0.731	0.001	0.730	0.733
Mother level, sociodemographics, trust level	0.732	0.001	0.731	0.734	0.732	0.001	0.731	0.734
Mother level, sociodemographics, trust level, staff 1	0.732	0.001	0.731	0.734				
Mother level, sociodemographics, trust level, staff 2					0.732	0.001	0.731	0.734

$AUC_{L95\%}$, lower 95% confidence limit of the AUC; $AUC_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; $OR_{L95\%}$, lower 95% confidence limit of the odds ratio; $OR_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.
Pr > χ^2 refers to the probability of exceeding the χ^2-statistic.
Pr > t refers to the probability of exceeding the t-statistic.

Elective caesarean: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Intercept	−3.463	0.224	−15.479	<0.0001				−3.409	0.244	−13.953	<0.0001			
Mother's age group (years)														
≤19	−2.037	0.077	−26.487	<0.0001	0.130	0.112	0.152	−2.037	0.077	−26.499	<0.0001	0.130	0.112	0.152
20–24	−1.589	0.069	−23.128	<0.0001	0.204	0.178	0.233	−1.590	0.069	−23.135	<0.0001	0.204	0.178	0.233
25–29	−1.261	0.068	−18.593	<0.0001	0.283	0.248	0.324	−1.262	0.068	−18.605	<0.0001	0.283	0.248	0.323
30–34	−0.993	0.068	−14.694	<0.0001	0.370	0.324	0.423	−0.994	0.068	−14.705	<0.0001	0.370	0.324	0.423
35–39	−0.732	0.068	−10.816	<0.0001	0.481	0.421	0.549	−0.733	0.068	−10.827	<0.0001	0.480	0.421	0.549
40–44	−0.535	0.069	−7.698	<0.0001	0.586	0.511	0.671	−0.536	0.069	−7.708	<0.0001	0.585	0.511	0.671
≥45	0.000				1.000			0.000				1.000		
Parity														
0	−0.565	0.021	−26.709	<0.0001	0.568	0.545	0.592	−0.565	0.021	−26.701	<0.0001	0.568	0.545	0.592
1	0.386	0.020	19.580	<0.0001	1.472	1.416	1.530	0.387	0.020	19.587	<0.0001	1.472	1.416	1.530
2	0.446	0.021	21.459	<0.0001	1.563	1.500	1.628	0.447	0.021	21.470	<0.0001	1.563	1.501	1.628
3	0.283	0.024	11.759	<0.0001	1.328	1.266	1.392	0.284	0.024	11.766	<0.0001	1.328	1.267	1.392
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	3.121	0.022	144.425	<0.0001	22.666	21.726	23.647	3.121	0.022	144.426	<0.0001	22.665	21.725	23.645

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Not given/not known/not stated	0.004	0.039	0.095	0.9245	1.004	0.930	1.083	0.003	0.039	0.089	0.9292	1.003	0.930	1.083
English/Welsh/Scottish/Northern Irish/British	0.061	0.030	2.069	0.0386	1.063	1.003	1.126	0.061	0.030	2.073	0.0382	1.063	1.003	1.126
Irish	0.201	0.066	3.034	0.0024	1.223	1.074	1.393	0.202	0.066	3.038	0.0024	1.223	1.074	1.393
Any other white background	0.065	0.033	1.998	0.0457	1.068	1.001	1.138	0.066	0.033	2.018	0.0435	1.068	1.002	1.139
White and black Caribbean	−0.367	0.086	−4.262	<0.0001	0.692	0.585	0.820	−0.367	0.086	−4.257	<0.0001	0.693	0.585	0.820
White and black African	0.069	0.085	0.814	0.4159	1.072	0.907	1.267	0.070	0.085	0.818	0.4136	1.072	0.907	1.267
White and Asian	0.132	0.095	1.388	0.1651	1.141	0.947	1.374	0.133	0.095	1.397	0.1624	1.142	0.948	1.375
Any other mixed/multiple ethnic background	0.144	0.068	2.129	0.0333	1.155	1.011	1.318	0.144	0.068	2.130	0.0332	1.155	1.012	1.318
Indian	−0.009	0.039	−0.228	0.8200	0.991	0.918	1.070	−0.009	0.039	−0.226	0.8214	0.991	0.918	1.070
Pakistani	−0.264	0.038	−6.862	<0.0001	0.768	0.712	0.828	−0.264	0.038	−6.863	<0.0001	0.768	0.712	0.828
Bangladeshi	−0.357	0.053	−6.678	<0.0001	0.700	0.630	0.777	−0.358	0.053	−6.688	<0.0001	0.699	0.630	0.777
Chinese	−0.137	0.069	−2.001	0.0454	0.872	0.762	0.997	−0.137	0.069	−2.000	0.0455	0.872	0.762	0.997
Any other Asian background	0.002	0.043	0.049	0.9607	1.002	0.921	1.090	0.002	0.043	0.057	0.9549	1.002	0.921	1.091
African	−0.088	0.037	−2.394	0.0167	0.915	0.852	0.984	−0.088	0.037	−2.375	0.0175	0.916	0.852	0.985
Caribbean	−0.262	0.056	−4.682	<0.0001	0.770	0.690	0.859	−0.261	0.056	−4.677	<0.0001	0.770	0.690	0.859
Any other black/African/Caribbean background	−0.221	0.057	−3.871	0.0001	0.801	0.716	0.896	−0.221	0.057	−3.867	0.0001	0.802	0.717	0.897
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

IMD	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
1 = most deprived	−0.204	0.017	−11.726	<0.0001	0.816	0.789	0.844	−0.204	0.017	−11.744	<0.0001	0.816	0.788	0.844
2	−0.131	0.016	−7.974	<0.0001	0.877	0.849	0.906	−0.132	0.016	−7.981	<0.0001	0.877	0.849	0.905
3	−0.099	0.016	−6.086	<0.0001	0.906	0.877	0.935	−0.099	0.016	−6.082	<0.0001	0.906	0.877	0.935
4	−0.024	0.016	−1.475	0.1402	0.977	0.946	1.008	−0.024	0.016	−1.493	0.1355	0.976	0.946	1.008
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification														
No information/ other postcode														
Urban ≥10,000 – sparse	0.029	0.123	0.237	0.8126	1.030	0.808	1.312	0.029	0.123	0.237	0.8125	1.030	0.808	1.312
Town and fringe – sparse	−0.007	0.089	−0.081	0.9351	0.993	0.834	1.182	−0.008	0.089	−0.087	0.9305	0.992	0.833	1.182
Village – sparse	−0.145	0.095	−1.532	0.1256	0.865	0.719	1.041	−0.145	0.095	−1.536	0.1246	0.865	0.718	1.041
Hamlet and isolated dwelling – sparse	0.064	0.118	0.541	0.5885	1.066	0.846	1.343	0.064	0.118	0.540	0.5890	1.066	0.846	1.343
Urban ≥10,000 – less sparse	−0.057	0.034	−1.669	0.0952	0.945	0.884	1.010	−0.057	0.034	−1.675	0.0940	0.945	0.884	1.010
Town and fringe – less sparse	−0.063	0.038	−1.684	0.0921	0.939	0.872	1.010	−0.064	0.038	−1.694	0.0902	0.938	0.872	1.010
Village – less sparse	−0.031	0.039	−0.788	0.4304	0.970	0.898	1.047	−0.031	0.039	−0.799	0.4243	0.969	0.898	1.046
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

SHA	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	−0.051	0.105	−0.487	0.6263	0.950	0.774	1.167	−0.051	0.105	−0.487	0.6263	0.950	0.774	1.167
North West	−0.111	0.081	−1.380	0.1675	0.895	0.764	1.048	−0.113	0.081	−1.408	0.1590	0.893	0.762	1.045
Yorkshire and Humber	−0.078	0.087	−0.898	0.3692	0.925	0.779	1.097	−0.079	0.087	−0.910	0.3630	0.924	0.778	1.096
East Midlands	−0.145	0.104	−1.392	0.1638	0.865	0.705	1.061	−0.144	0.104	−1.383	0.1665	0.866	0.707	1.062
West Midlands	−0.104	0.087	−1.186	0.2356	0.901	0.759	1.070	−0.105	0.087	−1.202	0.2293	0.900	0.758	1.069
East of England	−0.035	0.090	−0.395	0.6926	0.965	0.810	1.151	−0.037	0.089	−0.415	0.6781	0.964	0.809	1.148
London	0.093	0.089	1.045	0.2960	1.097	0.922	1.306	0.087	0.089	0.983	0.3256	1.091	0.917	1.299
South East Coast	−0.014	0.095	−0.143	0.8865	0.986	0.818	1.189	−0.015	0.095	−0.162	0.8715	0.985	0.817	1.187
South Central	−0.119	0.097	−1.225	0.2205	0.888	0.735	1.074	−0.122	0.096	−1.263	0.2066	0.886	0.733	1.069
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	−0.004	0.012	−0.332	0.7401	0.996	0.973	1.020	−0.003	0.012	−0.281	0.7790	0.997	0.974	1.020
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	−0.062	0.047	−1.304	0.1922	0.940	0.857	1.031	−0.061	0.047	−1.283	0.1995	0.941	0.858	1.032
Configuration														
OU	0.039	0.063	0.608	0.5431	1.039	0.918	1.177	0.038	0.064	0.594	0.5525	1.039	0.917	1.177
OU/AMU	−0.036	0.070	−0.509	0.6111	0.965	0.842	1.107	−0.034	0.070	−0.483	0.6291	0.967	0.843	1.109
OU/AMU/FMU	−0.021	0.082	−0.258	0.7962	0.979	0.834	1.150	−0.021	0.082	−0.260	0.7947	0.979	0.833	1.150
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	SE(β)	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
FTE doctors per 100 maternities	−0.144	0.112	−1.287	0.1981	0.866	0.695	1.078							
FTE midwives per 100 maternities	0.031	0.050	0.628	0.5303	1.032	0.936	1.137							
FTE support workers per 100 maternities	0.075	0.062	1.210	0.2262	1.077	0.955	1.216							
FTE staff per 100 maternities								0.010	0.030	0.337	0.7358	1.010	0.952	1.072
Doctor-to-midwife ratio								−0.427	0.319	−1.338	0.1809	0.652	0.349	1.220
Support worker-to-midwife ratio								0.189	0.184	1.029	0.3035	1.209	0.842	1.734
Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.206	0.008						0.206	0.008					
Mother level	0.233	0.015						0.233	0.015					
Mother level, sociodemographics	0.225	0.014						0.225	0.014					
Mother level, sociodemographics, trust level	0.222	0.014						0.222	0.014					
Mother level, sociodemographics, trust level, staff 1	0.219	0.014												
Mother level, sociodemographics, trust level, staff 2								0.219	0.014					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr>\chi^2$	χ^2	$Pr>\chi^2$
Mother's age group (6 df)	4276.429	<0.0001	4275.641	<0.0001
Parity (4 df)	7138.177	<0.0001	7139.119	<0.0001
Clinical risk (1 df)	20,858.460	<0.0001	20,858.801	<0.0001
Ethnicity (16 df)	328.588	<0.0001	328.967	<0.0001
IMD (4 df)	170.993	<0.0001	171.248	<0.0001
Rural/urban classification (7 df)	7.135	0.4150	7.152	0.4132
SHA (9 df)	12.153	0.2048	11.821	0.2236
ONS maternities (1 df)	0.110	0.7401	0.079	0.7790
University trust (1 df)	1.701	0.1922	1.646	0.1995
Configuration (3 df)	2.392	0.4951	2.269	0.5185
FTE doctors per 100 maternities (1 df)	1.656	0.1981		
FTE midwives per 100 maternities (1 df)	0.394	0.5303		
FTE support workers per 100 maternities (1 df)	1.465	0.2262		
FTE staff per 100 maternities (1 df)			0.114	0.7358
Doctor-to-midwife ratio (1 df)			1.790	0.1809
Support worker-to-midwife ratio (1 df)			1.059	0.3035

AUC	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
	AUC	SE(AUC)	AUC_L95%	AUC_U95%	AUC	SE(AUC)	AUC_L95%	AUC_U95%
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.811	0.001	0.810	0.813	0.811	0.001	0.810	0.813
Mother level, sociodemographics	0.814	0.001	0.812	0.815	0.814	0.001	0.812	0.815
Mother level, sociodemographics, trust level	0.814	0.001	0.812	0.815	0.814	0.001	0.812	0.815
Mother level, sociodemographics, trust level, staff 1	0.814	0.001	0.813	0.816				
Mother level, sociodemographics, trust level, staff 2					0.814	0.001	0.813	0.816

$AUC_{L95\%}$, lower 95% confidence limit of the AUC; $AUC_{U95\%}$, upper 95% confidence limit of the AUC; $β$, beta coefficient; df, degrees of freedom; $OR_{L95\%}$, lower 95% confidence limit of the odds ratio; $OR_{U95\%}$, upper 95% confidence limit of the odds ratio; $σ$, sigma coefficient; $SE(AUC)$, SE of area under the curve; $SE(β)$, SE of beta coefficient; $SE(σ)$, SE of sigma coefficient.
$Pr > χ^2$ refers to the probability of exceeding the $χ^2$-statistic.
$Pr > t$ refers to the probability of exceeding the t-statistic.

Emergency caesarean: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Intercept	−2.696	0.201	−13.441	<0.0001				−2.655	0.218	−12.159	<0.0001			
Mother's age group (years)														
≤19	−1.054	0.073	−14.479	<0.0001	0.348	0.302	0.402	−1.055	0.073	−14.495	<0.0001	0.348	0.302	0.401
20–24	−0.735	0.071	−10.351	<0.0001	0.479	0.417	0.551	−0.735	0.071	−10.349	<0.0001	0.479	0.417	0.551
25–29	−0.535	0.071	−7.571	<0.0001	0.585	0.510	0.672	−0.535	0.071	−7.570	<0.0001	0.585	0.510	0.673
30–34	−0.374	0.071	−5.302	<0.0001	0.688	0.599	0.790	−0.374	0.071	−5.301	<0.0001	0.688	0.599	0.790
35–39	−0.212	0.071	−2.991	0.0028	0.809	0.704	0.930	−0.212	0.071	−2.990	0.0028	0.809	0.704	0.930
40–44	−0.049	0.072	−0.680	0.4963	0.952	0.826	1.097	−0.051	0.072	−0.699	0.4849	0.951	0.825	1.096
≥45	0.000				1.000			0.000				1.000		
Parity														
0	1.356	0.020	67.343	<0.0001	3.882	3.732	4.038	1.356	0.020	67.343	<0.0001	3.882	3.732	4.038
1	0.428	0.020	20.955	<0.0001	1.534	1.473	1.596	0.427	0.020	20.950	<0.0001	1.533	1.473	1.596
2	0.122	0.022	5.481	<0.0001	1.130	1.082	1.180	0.122	0.022	5.480	<0.0001	1.130	1.082	1.180
3	0.070	0.026	2.720	0.0065	1.073	1.020	1.129	0.070	0.026	2.721	0.0065	1.073	1.020	1.129
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	1.183	0.009	135.601	<0.0001	3.263	3.208	3.319	1.183	0.009	135.619	<0.0001	3.264	3.208	3.320

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	$SE(\beta)$	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$	β	$SE(\beta)$	t-statistic	Pr>t	OR	OR$_{L95\%}$	OR$_{U95\%}$
Not given/not known/not stated	−0.141	0.029	−4.894	<0.0001	0.869	0.821	0.919	−0.141	0.029	−4.889	<0.0001	0.869	0.821	0.919
English/Welsh/Scottish/Northern Irish/British	−0.129	0.023	−5.557	<0.0001	0.879	0.840	0.920	−0.128	0.023	−5.541	<0.0001	0.879	0.840	0.920
Irish	−0.054	0.058	−0.927	0.3541	0.947	0.845	1.062	−0.053	0.058	−0.913	0.3612	0.948	0.846	1.063
Any other white background	−0.174	0.026	−6.755	<0.0001	0.841	0.799	0.884	−0.173	0.026	−6.735	<0.0001	0.841	0.800	0.885
White and black Caribbean	0.018	0.060	0.298	0.7659	1.018	0.904	1.146	0.018	0.060	0.300	0.7640	1.018	0.905	1.146
White and black African	0.015	0.072	0.214	0.8307	1.015	0.882	1.168	0.015	0.072	0.203	0.8391	1.015	0.882	1.168
White and Asian	−0.041	0.079	−0.524	0.6002	0.959	0.822	1.120	−0.042	0.079	−0.529	0.5968	0.959	0.821	1.120
Any other mixed/multiple ethnic background	−0.102	0.056	−1.797	0.0723	0.903	0.809	1.009	−0.101	0.056	−1.786	0.0741	0.904	0.809	1.010
Indian	0.118	0.030	3.925	<0.0001	1.125	1.061	1.193	0.118	0.030	3.938	<0.0001	1.125	1.061	1.193
Pakistani	−0.018	0.030	−0.602	0.5474	0.982	0.926	1.042	−0.018	0.030	−0.608	0.5432	0.982	0.925	1.042
Bangladeshi	0.103	0.040	2.561	0.0104	1.109	1.025	1.200	0.103	0.040	2.554	0.0106	1.109	1.024	1.200
Chinese	−0.170	0.052	−3.299	0.0010	0.843	0.762	0.933	−0.170	0.052	−3.297	0.0010	0.843	0.762	0.933
Any other Asian background	0.093	0.033	2.785	0.0054	1.098	1.028	1.172	0.093	0.033	2.794	0.0052	1.098	1.028	1.172
African	0.374	0.029	13.017	<0.0001	1.453	1.374	1.537	0.374	0.029	13.015	<0.0001	1.453	1.374	1.537
Caribbean	0.196	0.042	4.710	<0.0001	1.216	1.121	1.319	0.196	0.042	4.705	<0.0001	1.216	1.121	1.319
Any other black/African/Caribbean background	0.276	0.042	6.560	<0.0001	1.318	1.214	1.432	0.276	0.042	6.545	<0.0001	1.317	1.213	1.431
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
IMD	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
1 = most deprived	0.107	0.015	7.331	<0.0001	1.113	1.081	1.145	0.107	0.015	7.335	<0.0001	1.113	1.081	1.145
2	0.086	0.014	6.191	<0.0001	1.090	1.061	1.120	0.086	0.014	6.188	<0.0001	1.090	1.060	1.120
3	0.054	0.014	3.890	0.0001	1.055	1.027	1.084	0.054	0.014	3.881	0.0001	1.055	1.027	1.084
4	0.020	0.014	1.418	0.1561	1.020	0.992	1.048	0.020	0.014	1.408	0.1592	1.020	0.992	1.048
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
No information/other postcode														
Urban ≥ 10,000 – sparse	-0.050	0.108	-0.460	0.6457	0.952	0.770	1.175	-0.049	0.108	-0.457	0.6475	0.952	0.771	1.176
Town and fringe – sparse	-0.029	0.078	-0.366	0.7142	0.972	0.834	1.132	-0.028	0.078	-0.359	0.7198	0.972	0.834	1.133
Village – sparse	-0.074	0.083	-0.896	0.3703	0.928	0.789	1.092	-0.074	0.083	-0.892	0.3724	0.929	0.789	1.093
Hamlet and isolated dwelling – sparse	0.038	0.105	0.361	0.7182	1.039	0.845	1.276	0.038	0.105	0.361	0.7178	1.039	0.845	1.276
Urban ≥ 10,000 – less sparse	0.037	0.030	1.239	0.2153	1.038	0.978	1.101	0.037	0.030	1.240	0.2150	1.038	0.978	1.101
Town and fringe – less sparse	0.006	0.033	0.168	0.8667	1.006	0.942	1.073	0.005	0.033	0.158	0.8745	1.005	0.942	1.073
Village – less sparse	-0.018	0.035	-0.507	0.6120	0.982	0.918	1.052	-0.017	0.035	-0.497	0.6192	0.983	0.918	1.052
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
SHA	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
North East	−0.007	0.093	−0.080	0.9359	0.993	0.828	1.190	−0.008	0.093	−0.083	0.9338	0.992	0.828	1.190
North West	−0.090	0.071	−1.258	0.2083	0.914	0.795	1.051	−0.090	0.071	−1.260	0.2075	0.914	0.795	1.051
Yorkshire and Humber	0.018	0.077	0.229	0.8191	1.018	0.875	1.184	0.017	0.077	0.220	0.8259	1.017	0.874	1.183
East Midlands	−0.080	0.092	−0.870	0.3843	0.923	0.770	1.106	−0.078	0.092	−0.853	0.3938	0.925	0.772	1.107
West Midlands	0.021	0.077	0.267	0.7891	1.021	0.877	1.188	0.019	0.077	0.243	0.8082	1.019	0.876	1.186
East of England	0.119	0.079	1.505	0.1322	1.127	0.965	1.316	0.121	0.079	1.529	0.1263	1.128	0.967	1.317
London	0.094	0.079	1.197	0.2313	1.099	0.942	1.282	0.097	0.078	1.241	0.2147	1.102	0.945	1.286
South East Coast	0.011	0.084	0.134	0.8935	1.011	0.857	1.193	0.013	0.084	0.149	0.8814	1.013	0.859	1.194
South Central	0.084	0.086	0.982	0.3263	1.088	0.920	1.286	0.085	0.085	0.999	0.3179	1.089	0.922	1.286
South West	0.000				1.000			0.000				1.000		
Trust size	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
ONS maternities (thousands)	0.005	0.011	0.466	0.6414	1.005	0.984	1.026	0.005	0.010	0.501	0.6166	1.005	0.985	1.026
University trust	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Yes	0.000				1.000			0.000				1.000		
No	−0.037	0.042	−0.876	0.3813	0.964	0.888	1.046	−0.036	0.042	−0.864	0.3877	0.965	0.889	1.047
Configuration	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
OU	−0.031	0.056	−0.559	0.5759	0.969	0.869	1.081	−0.032	0.056	−0.578	0.5634	0.968	0.867	1.081
OU/AMU	−0.061	0.062	−0.982	0.3260	0.941	0.834	1.062	−0.060	0.062	−0.974	0.3302	0.942	0.834	1.063
OU/AMU/FMU	−0.028	0.072	−0.380	0.7036	0.973	0.844	1.121	−0.029	0.073	−0.393	0.6943	0.972	0.843	1.121
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
FTE doctors per 100 maternities	−0.105	0.099	−1.057	0.2903	0.900	0.741	1.094							
FTE midwives per 100 maternities	−0.023	0.044	−0.512	0.6085	0.978	0.897	1.066							
FTE support workers per 100 maternities	−0.007	0.055	−0.133	0.8941	0.993	0.892	1.105							
FTE staff per 100 maternities								−0.032	0.027	−1.204	0.2287	0.968	0.919	1.020
Doctor-to-midwife ratio								−0.274	0.282	−0.971	0.3314	0.760	0.437	1.322
Support worker-to-midwife ratio								0.079	0.163	0.482	0.6296	1.082	0.786	1.489

Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.203	0.007						0.203	0.007					
Mother level	0.226	0.014						0.226	0.014					
Mother level, sociodemographics	0.198	0.013						0.198	0.013					
Mother level, sociodemographics, trust level	0.197	0.012						0.197	0.012					
Mother level, sociodemographics, trust level, staff 1	0.195	0.012												
Mother level, sociodemographics, trust level, staff 2								0.195	0.012					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr>\chi^2$	χ^2	$Pr>\chi^2$
Mother's age group (6 df)	2921.832	<0.0001	2921.593	<0.0001
Parity (4 df)	17,160.218	<0.0001	17,162.241	<0.0001
Clinical risk (1 df)	18,387.718	<0.0001	18,392.505	<0.0001
Ethnicity (16 df)	954.095	<0.0001	952.142	<0.0001
IMD (4 df)	70.290	<0.0001	70.430	<0.0001
Rural/urban classification (7 df)	13.060	0.0706	13.024	0.0715
SHA (9 df)	12.906	0.1669	13.280	0.1503
ONS maternities (1 df)	0.217	0.6414	0.251	0.6166
University trust (1 df)	0.767	0.3813	0.746	0.3877
Configuration (3 df)	1.042	0.7911	1.011	0.7986
FTE doctors per 100 maternities (1 df)	1.118	0.2903		
FTE midwives per 100 maternities (1 df)	0.262	0.6085		
FTE support workers per 100 maternities (1 df)	0.018	0.8941		
FTE staff per 100 maternities (1 df)			1.449	0.2287
Doctor-to-midwife ratio (1 df)			0.943	0.3314
Support worker-to-midwife ratio (1 df)			0.233	0.6296

AUC	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
	AUC	SE(AUC)	$AUC_{L95\%}$	$AUC_{U95\%}$	AUC	SE(AUC)	$AUC_{L95\%}$	$AUC_{U95\%}$
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.692	0.001	0.690	0.694	0.692	0.001	0.690	0.694
Mother level, sociodemographics	0.698	0.001	0.697	0.700	0.698	0.001	0.697	0.700
Mother level, sociodemographics, trust level	0.698	0.001	0.696	0.700	0.698	0.001	0.696	0.700
Mother level, sociodemographics, trust level, staff 1	0.698	0.001	0.696	0.700				
Mother level, sociodemographics, trust level, staff 2					0.698	0.001	0.697	0.700

$AUC_{L95\%}$, lower 95% confidence limit of the AUC; $AUC_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; $OR_{L95\%}$, lower 95% confidence limit of the odds ratio; $OR_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(σ), SE of sigma coefficient; SE(β), SE of beta coefficient; SE(AUC), SE of area under the curve; SE(AUC), SE of area under the curve; SE(σ), SE of sigma coefficient.
$Pr > \chi^2$ refers to the probability of exceeding the χ^2-statistic.
$Pr > t$ refers to the probability of exceeding the t-statistic.

Caesarean: multilevel model

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Intercept	−1.595	0.192	−8.311	<0.0001				−1.524	0.209	−7.284	<0.0001			
Mother's age group (years)														
≤19	−1.841	0.064	−28.771	<0.0001	0.159	0.140	0.180	−1.842	0.064	−28.781	<0.0001	0.159	0.140	0.180
20–24	−1.486	0.062	−23.904	<0.0001	0.226	0.200	0.256	−1.486	0.062	−23.910	<0.0001	0.226	0.200	0.256
25–29	−1.214	0.062	−19.635	<0.0001	0.297	0.263	0.335	−1.215	0.062	−19.646	<0.0001	0.297	0.263	0.335
30–34	−0.973	0.062	−15.752	<0.0001	0.378	0.335	0.427	−0.974	0.062	−15.766	<0.0001	0.378	0.335	0.426
35–39	−0.696	0.062	−11.247	<0.0001	0.499	0.442	0.563	−0.697	0.062	−11.266	<0.0001	0.498	0.441	0.562
40–44	−0.445	0.063	−7.043	<0.0001	0.641	0.566	0.725	−0.447	0.063	−7.063	<0.0001	0.640	0.565	0.724
≥45	0.000				1.000			0.000				1.000		
Parity														
0	0.804	0.016	51.134	<0.0001	2.235	2.167	2.305	0.804	0.016	51.117	<0.0001	2.234	2.166	2.304
1	0.508	0.016	32.552	<0.0001	1.661	1.611	1.713	0.507	0.016	32.540	<0.0001	1.661	1.611	1.712
2	0.371	0.017	22.263	<0.0001	1.449	1.403	1.497	0.371	0.017	22.262	<0.0001	1.449	1.403	1.497
3	0.225	0.019	11.643	<0.0001	1.253	1.206	1.301	0.225	0.019	11.627	<0.0001	1.252	1.206	1.301
≥4	0.000				1.000			0.000				1.000		
Clinical risk														
Lower	0.000				1.000			0.000				1.000		
Higher	1.867	0.008	234.270	<0.0001	6.467	6.366	6.568	1.867	0.008	234.270	<0.0001	6.466	6.366	6.568

Ethnicity	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
Not given/not known/not stated	−0.108	0.026	−4.213	<0.0001	0.898	0.854	0.944	−0.108	0.026	−4.217	<0.0001	0.898	0.854	0.944
English/Welsh/Scottish/Northern Irish/British	−0.069	0.020	−3.385	0.0007	0.934	0.897	0.971	−0.069	0.020	−3.400	0.0007	0.933	0.897	0.971
Irish	0.066	0.049	1.329	0.1838	1.068	0.969	1.176	0.063	0.049	1.285	0.1989	1.065	0.967	1.174
Any other white background	−0.101	0.022	−4.505	<0.0001	0.904	0.865	0.944	−0.102	0.022	−4.514	<0.0001	0.903	0.865	0.944
White and black Caribbean	−0.145	0.054	−2.701	0.0069	0.865	0.778	0.961	−0.150	0.054	−2.777	0.0055	0.861	0.775	0.957
White and black African	0.048	0.061	0.787	0.4313	1.049	0.931	1.183	0.048	0.061	0.779	0.4357	1.049	0.930	1.182
White and Asian	0.036	0.068	0.537	0.5915	1.037	0.908	1.185	0.032	0.068	0.476	0.6342	1.033	0.904	1.180
Any other mixed/multiple ethnic background	−0.004	0.048	−0.076	0.9391	0.996	0.907	1.095	−0.004	0.048	−0.090	0.9280	0.996	0.906	1.094
Indian	0.091	0.027	3.443	0.0006	1.096	1.040	1.154	0.091	0.027	3.432	0.0006	1.095	1.040	1.154
Pakistani	−0.153	0.026	−5.831	<0.0001	0.858	0.815	0.903	−0.152	0.026	−5.797	<0.0001	0.859	0.816	0.904
Bangladeshi	−0.107	0.036	−3.006	0.0027	0.899	0.838	0.964	−0.107	0.036	−3.021	0.0025	0.898	0.838	0.963
Chinese	−0.191	0.045	−4.194	<0.0001	0.826	0.756	0.903	−0.191	0.045	−4.198	<0.0001	0.826	0.756	0.903
Any other Asian background	0.077	0.030	2.599	0.0093	1.080	1.019	1.144	0.075	0.030	2.545	0.0109	1.078	1.017	1.142
African	0.235	0.026	9.215	<0.0001	1.265	1.203	1.330	0.235	0.026	9.223	<0.0001	1.265	1.204	1.330
Caribbean	0.023	0.037	0.606	0.5443	1.023	0.951	1.100	0.021	0.037	0.574	0.5661	1.022	0.950	1.099
Any other black/African/Caribbean background	0.108	0.038	2.847	0.0044	1.114	1.034	1.200	0.108	0.038	2.833	0.0046	1.114	1.034	1.200
Any other ethnic group, please describe	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
IMD	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
1 = most deprived	−0.029	0.012	−2.349	0.019	0.971	0.948	0.995	−0.030	0.012	−2.395	0.017	0.971	0.948	0.995
2	−0.008	0.012	−0.692	0.489	0.992	0.969	1.015	−0.009	0.012	−0.735	0.462	0.991	0.969	1.015
3	−0.015	0.012	−1.317	0.188	0.985	0.962	1.008	−0.016	0.012	−1.333	0.183	0.985	0.962	1.007
4	−0.001	0.012	−0.073	0.942	0.999	0.976	1.022	−0.001	0.012	−0.088	0.930	0.999	0.976	1.022
5 = least deprived	0.000				1.000			0.000				1.000		
Rural/urban classification	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
No information/other postcode														
Urban ≥10,000 – sparse	−0.019	0.089	−0.218	0.827	0.981	0.824	1.168	−0.021	0.089	−0.231	0.818	0.980	0.823	1.166
Town and fringe – sparse	−0.027	0.064	−0.424	0.672	0.973	0.858	1.104	−0.026	0.064	−0.400	0.689	0.975	0.859	1.106
Village – sparse	−0.129	0.068	−1.878	0.060	0.879	0.769	1.006	−0.130	0.068	−1.895	0.058	0.878	0.768	1.004
Hamlet and isolated dwelling – sparse	0.059	0.087	0.682	0.495	1.061	0.895	1.258	0.058	0.087	0.666	0.506	1.060	0.894	1.256
Urban ≥10,000 – less sparse	−0.005	0.025	−0.203	0.839	0.995	0.947	1.045	−0.006	0.025	−0.250	0.803	0.994	0.946	1.044
Town and fringe – less sparse	−0.031	0.028	−1.121	0.262	0.970	0.919	1.023	−0.032	0.028	−1.159	0.246	0.969	0.918	1.022
Village – less sparse	−0.025	0.029	−0.865	0.387	0.975	0.922	1.032	−0.026	0.029	−0.911	0.362	0.974	0.921	1.031
Hamlet and isolated dwelling – less sparse	0.000				1.000			0.000				1.000		

	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
SHA	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**
North East	−0.030	0.090	−0.337	0.736	0.970	0.813	1.157	−0.040	0.090	−0.446	0.656	0.961	0.806	1.146
North West	−0.125	0.069	−1.804	0.071	0.882	0.770	1.011	−0.137	0.069	−1.975	0.048	0.872	0.762	0.999
Yorkshire and Humber	−0.028	0.075	−0.371	0.711	0.973	0.839	1.127	−0.034	0.075	−0.460	0.646	0.966	0.834	1.119
East Midlands	−0.134	0.090	−1.494	0.135	0.875	0.733	1.043	−0.141	0.089	−1.580	0.114	0.868	0.729	1.035
West Midlands	−0.042	0.075	−0.553	0.581	0.959	0.828	1.112	−0.051	0.075	−0.684	0.494	0.950	0.820	1.101
East of England	0.065	0.077	0.839	0.402	1.067	0.917	1.241	0.056	0.077	0.731	0.465	1.058	0.910	1.230
London	0.112	0.077	1.464	0.143	1.119	0.963	1.300	0.106	0.076	1.391	0.164	1.112	0.957	1.292
South East Coast	−0.004	0.082	−0.046	0.963	0.996	0.848	1.171	−0.014	0.082	−0.167	0.868	0.986	0.840	1.158
South Central	0.003	0.083	0.038	0.970	1.003	0.852	1.181	−0.008	0.083	−0.092	0.927	0.992	0.844	1.167
South West	0.000				1.000			0.000				1.000		
Trust size	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**
ONS maternities (thousands)	0.001	0.010	0.100	0.920	1.001	0.981	1.021	0.001	0.010	0.123	0.902	1.001	0.981	1.021
University trust	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**
Yes	0.000				1.000			0.000				1.000		
No	−0.059	0.041	−1.442	0.149	0.943	0.870	1.021	−0.059	0.041	−1.444	0.149	0.943	0.871	1.021
Configuration	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**	**β**	**SE(β)**	**t-statistic**	**Pr>t**	**OR**	**OR_L95%**	**OR_U95%**
OU	−0.006	0.054	−0.111	0.912	0.994	0.893	1.106	−0.003	0.055	−0.056	0.956	0.997	0.896	1.110
OU/AMU	−0.062	0.060	−1.035	0.301	0.940	0.835	1.057	−0.059	0.060	−0.980	0.327	0.943	0.838	1.061
OU/AMU/FMU	−0.033	0.071	−0.471	0.638	0.967	0.842	1.111	−0.031	0.071	−0.432	0.666	0.970	0.844	1.114
OU/FMU	0.000				1.000			0.000				1.000		

Staffing variables	Model 1: FTE staffing variables							Model 2: staffing ratio variables						
	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%	β	SE(β)	t-statistic	Pr>t	OR	OR_L95%	OR_U95%
FTE doctors per 100 maternities	−0.154	0.097	−1.594	0.1110	0.857	0.709	1.036							
FTE midwives per 100 maternities	0.000	0.043	−0.005	0.9962	1.000	0.919	1.087							
FTE support workers per 100 maternities	0.031	0.053	0.583	0.5600	1.031	0.930	1.144							
FTE staff per 100 maternities								−0.020	0.026	−0.774	0.4391	0.980	0.931	1.031
Doctor-to-midwife ratio								−0.430	0.275	−1.566	0.1172	0.650	0.380	1.114
Support worker-to-midwife ratio								0.167	0.159	1.055	0.2914	1.182	0.866	1.613
Random variation (trust level)	σ	SE(σ)						σ	SE(σ)					
Intercept only	0.208	0.006						0.208	0.006					
Mother level	0.215	0.013						0.215	0.013					
Mother level, sociodemographics	0.196	0.012						0.196	0.012					
Mother level, sociodemographics, trust level	0.194	0.012						0.194	0.012					
Mother level, sociodemographics, trust level, staff 1	0.192	0.012												
Mother level, sociodemographics, trust level, staff 2								0.191	0.012					

Global tests (df)	Model 1: FTE staffing variables		Model 2: staffing ratio variables	
	χ^2	$Pr>\chi^2$	χ^2	$Pr>\chi^2$
Mother's age group (6 df)	8272.444	< 0.0001	8261.615	< 0.0001
Parity (4 df)	4244.311	< 0.0001	4241.702	< 0.0001
Clinical risk (1 df)	54,882.322	< 0.0001	54,882.242	< 0.0001
Ethnicity (16 df)	504.113	< 0.0001	503.361	< 0.0001
IMD (4 df)	8.436	0.0769	8.629	0.0711
Rural/urban classification (7 df)	8.628	0.2805	8.637	0.2798
SHA (9 df)	17.276	0.0446	18.216	0.0328
ONS maternities (1 df)	0.010	0.9202	0.015	0.9018
University trust (1 df)	2.078	0.1494	2.086	0.1486
Configuration (3 df)	1.916	0.5900	1.858	0.6024
FTE doctors per 100 maternities (1 df)	2.540	0.1110		
FTE midwives per 100 maternities (1 df)	0.000	0.9962		
FTE support workers per 100 maternities (1 df)	0.340	0.5600		
FTE staff per 100 maternities (1 df)			0.599	0.4391
Doctor-to-midwife ratio (1 df)			2.454	0.1172
Support worker-to-midwife ratio (1 df)			1.113	0.2914

AUC	Model 1: FTE staffing variables				Model 2: staffing ratio variables			
	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$	AUC	SE(AUC)	AUC$_{L95\%}$	AUC$_{U95\%}$
Intercept only	0.500	0.000	0.500	0.500	0.500	0.000	0.500	0.500
Mother level	0.737	0.001	0.736	0.738	0.737	0.001	0.736	0.738
Mother level, sociodemographics	0.740	0.001	0.738	0.741	0.740	0.001	0.738	0.741
Mother level, sociodemographics, trust level	0.740	0.001	0.738	0.741	0.740	0.001	0.738	0.741
Mother level, sociodemographics, trust level, staff 1	0.740	0.001	0.739	0.741				
Mother level, sociodemographics, trust level, staff 2					0.740	0.001	0.739	0.741

AUC$_{L95\%}$, lower 95% confidence limit of the AUC; AUC$_{U95\%}$, upper 95% confidence limit of the AUC; β, beta coefficient; df, degrees of freedom; OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio; σ, sigma coefficient; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient; SE(AUC), SE of area under the curve; SE(AUC), SE of area under the curve.
Pr > χ² refers to the probability of exceeding the χ²-statistic.
Pr > t refers to the probability of exceeding the t-statistic.

Appendix 5 Multilevel model sensitivity analyses

Healthy mother: sensitivity analysis

Variable	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Intercept	−0.003	0.275	0.004	0.276	0.950	0.410
Mother's age group (years)	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
≤ 19	1.320	0.092	1.301	0.092	1.224	0.179
20–24	0.999	0.091	0.982	0.091	0.878	0.178
25–29	0.629	0.090	0.613	0.091	0.513	0.177
30–34	0.363	0.090	0.349	0.091	0.283	0.177
35–39	0.189	0.090	0.176	0.091	0.132	0.178
40–44	0.035	0.092	0.020	0.093	0.048	0.181
≥ 45	0.000		0.000		0.000	
Parity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
0	−2.631	0.017	−2.624	0.017	−2.529	0.033
1	−1.359	0.016	−1.350	0.016	−1.274	0.031
2	−0.726	0.017	−0.714	0.017	−0.634	0.033
3	−0.362	0.019	−0.347	0.019	−0.290	0.038
≥ 4	0.000		0.000		0.000	
Clinical risk	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Lower	0.000		0.000		0.000	
Individual assessment	N/A		−0.400	0.018		
Higher	−0.984	0.008	−1.041	0.008	−0.955	0.014
Ethnicity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Not given/not known/not stated	0.128	0.031	0.129	0.031	0.104	0.058
English/Welsh/Scottish/Northern Irish/British	0.145	0.025	0.147	0.025	0.105	0.047
Irish	−0.123	0.066	−0.118	0.066	−0.015	0.133
Any other white background	0.107	0.028	0.108	0.028	0.067	0.052
White and black Caribbean	0.575	0.056	0.577	0.056	0.521	0.120
White and black African	0.146	0.070	0.148	0.070	0.062	0.150
White and Asian	−0.093	0.083	−0.099	0.084	−0.212	0.175
Any other mixed/multiple ethnic background	0.162	0.057	0.164	0.057	0.225	0.122
Indian	−0.503	0.035	−0.504	0.035	−0.547	0.068
Pakistani	−0.056	0.031	−0.059	0.031	−0.104	0.068

Variable	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Bangladeshi	−0.194	0.040	−0.190	0.040	−0.251	0.093
Chinese	−0.372	0.058	−0.374	0.058	−0.327	0.111
Any other Asian background	−0.200	0.038	−0.200	0.038	−0.222	0.075
African	0.027	0.032	0.031	0.032	0.031	0.062
Caribbean	0.507	0.042	0.510	0.042	0.442	0.083
Any other black/African/Caribbean background	0.192	0.045	0.196	0.045	0.294	0.082
Any other ethnic group, please describe	0.000		0.000		0.000	
IMD	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
1 = most deprived	0.324	0.014	0.327	0.015	0.356	0.026
2	0.236	0.014	0.238	0.014	0.280	0.026
3	0.149	0.014	0.150	0.014	0.169	0.026
4	0.096	0.015	0.097	0.015	0.117	0.026
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Urban ≥ 10,000 – sparse	−0.146	0.096	−0.148	0.096	−0.295	0.229
Town and fringe – sparse	0.006	0.070	−0.001	0.070	−0.026	0.118
Village – sparse	0.073	0.073	0.073	0.073	−0.049	0.133
Hamlet and isolated dwelling – sparse	−0.100	0.098	−0.108	0.098	−0.502	0.205
Urban ≥ 10,000 – less sparse	0.059	0.030	0.059	0.030	0.038	0.051
Town and fringe – less sparse	0.109	0.032	0.109	0.032	0.105	0.055
Village – less sparse	0.092	0.034	0.092	0.034	0.068	0.057
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
North East	0.215	0.119	0.219	0.120	0.503	0.114
North West	0.032	0.092	0.037	0.093	0.149	0.100
Yorkshire and Humber	0.205	0.098	0.210	0.098	0.137	0.106
East Midlands	0.178	0.115	0.179	0.115	0.267	0.121
West Midlands	0.063	0.109	0.065	0.109	−0.027	0.110
East of England	0.093	0.109	0.094	0.109	0.093	0.111
London	−0.116	0.101	−0.115	0.102	0.137	0.113
South East Coast	−0.030	0.121	−0.029	0.122	−0.115	0.155
South Central	−0.084	0.115	−0.084	0.115	0.215	0.145
South West	0.000		0.000		0.000	

Variable	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Trust size	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
ONS maternities (thousands)	−0.028	0.015	−0.028	0.015	−0.122	0.032
University trust	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Yes	0.000		0.000		0.000	
No	0.057	0.057	0.058	0.057	−0.128	0.098
Configuration	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
OU	0.006	0.071	0.005	0.072	0.000	
OU/AMU	−0.052	0.081	−0.054	0.082		
OU/AMU/FMU	0.003	0.098	0.001	0.098		
OU/FMU	0.000		0.000			
Staffing variables	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
FTE doctors per 100 maternities	0.069	0.134	0.060	0.135	−0.222	0.146
FTE midwives per 100 maternities	0.084	0.062	0.086	0.063	−0.021	0.096
FTE support workers per 100 maternities	−0.114	0.071	−0.116	0.072	−0.073	0.104
Random variation (trust level)	σ	$SE(\sigma)$	σ	$SE(\sigma)$	σ	$SE(\sigma)$
Intercept only	0.283	0.011	0.283	0.011	0.214	0.019
Mother level	0.308	0.021	0.310	0.021	0.260	0.029
Mother level, sociodemographics	0.256	0.018	0.257	0.018	0.192	0.022
Mother level, sociodemographics, trust level	0.242	0.017	0.242	0.017	0.162	0.019
Mother level, sociodemographics, trust level, staff 1	0.238	0.016	0.239	0.016	0.154	0.018
Global tests (df)	χ^2	p-value	χ^2	p-value	χ^2	p-value
Mother's age group (6 df)	6495.901	< 0.0001	6418.4640	< 0.0001	1736.611	< 0.0001
Parity (4 df)	40,071.673	< 0.0001	40,051.9730	< 0.0001	11,364.017	< 0.0001
Clinical risk (1/2 df)[a]	15,840.786	< 0.0001	16,779.6590	< 0.0001	4382.500	< 0.0001
Ethnicity (16 df)	1190.077	< 0.0001	1198.2640	< 0.0001	276.915	< 0.0001
IMD (4 df)	582.919	< 0.0001	593.3470	< 0.0001	214.923	< 0.0001
Rural/urban classification (7 df)	27.018	0.0003	27.7000	0.0002	18.399	0.0103
SHA (9 df)	19.979	0.0180	20.1330	0.0171	30.470	0.0004
ONS maternities (1 df)	3.533	0.0602	3.5600	0.0592	14.206	0.0002
University trust (1 df)	1.004	0.3164	1.0520	0.3050	1.702	0.1921
Configuration (3 df)	0.966	0.8095	1.0140	0.7978		
FTE doctors per 100 maternities (1 df)	0.260	0.6099	0.2000	0.6549	2.302	0.1292
FTE midwives per 100 maternities (1 df)	1.832	0.1759	1.8800	0.1703	0.048	0.8263
FTE support workers per 100 maternities (1 df)	2.570	0.1089	2.6260	0.1051	0.497	0.4810

Variable	All mothers		Clinical risk (three categories)		Trusts with single OUs	
AUC	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.735	0.001	0.737	0.001	0.729	0.002
Mother level, sociodemographics	0.741	0.001	0.743	0.001	0.737	0.002
Mother level, sociodemographics, trust level	0.742	0.001	0.744	0.001	0.738	0.002
Mother level, sociodemographics, trust level, staff 1	0.742	0.001	0.744	0.001	0.739	0.002

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Healthy baby: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Intercept	2.817	0.272	2.847	0.278	2.807	0.433
Mother's age group (years)	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
≤ 19	0.092	0.086	0.042	0.087	0.112	0.172
20–24	0.116	0.085	0.074	0.085	0.135	0.170
25–29	0.193	0.084	0.155	0.085	0.212	0.169
30–34	0.208	0.084	0.175	0.085	0.207	0.169
35–39	0.212	0.085	0.185	0.085	0.209	0.170
40–44	0.267	0.087	0.240	0.087	0.238	0.174
≥ 45	0.000		0.000		0.000	
Parity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
0	−0.317	0.020	−0.293	0.020	−0.249	0.040
1	0.185	0.020	0.217	0.020	0.222	0.040
2	0.154	0.022	0.186	0.022	0.185	0.043
3	0.080	0.025	0.115	0.026	0.086	0.051
≥ 4	0.000		0.000		0.000	
Clinical risk	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Lower	0.000		0.000		0.000	
Individual assessment			−0.499	0.032		
Higher	−1.986	0.012	−2.069	0.012	−2.002	0.024
Ethnicity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Not given/not known/not stated	−0.058	0.038	−0.056	0.038	−0.100	0.073
English/Welsh/Scottish/Northern Irish/British	−0.110	0.030	−0.109	0.030	−0.149	0.059
Irish	−0.166	0.072	−0.157	0.072	−0.372	0.147
Any other white background	−0.016	0.033	−0.017	0.033	−0.002	0.066
White and black Caribbean	−0.165	0.071	−0.166	0.072	−0.195	0.162
White and black African	0.080	0.089	0.063	0.090	0.286	0.208
White and Asian	−0.151	0.095	−0.163	0.095	−0.257	0.200
Any other mixed/multiple ethnic background	−0.046	0.069	−0.046	0.070	−0.253	0.149
Indian	−0.136	0.039	−0.141	0.039	−0.149	0.077
Pakistani	−0.118	0.038	−0.122	0.038	−0.144	0.084
Bangladeshi	−0.161	0.048	−0.155	0.048	−0.237	0.112
Chinese	0.173	0.069	0.171	0.069	0.264	0.146
Any other Asian background	−0.028	0.044	−0.031	0.044	−0.049	0.090
African	−0.017	0.038	−0.006	0.038	−0.075	0.076
Caribbean	−0.138	0.052	−0.140	0.053	−0.361	0.102

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Any other black/African/Caribbean background	−0.180	0.053	−0.167	0.053	−0.178	0.104
Any other ethnic group, please describe	0.000		0.000		0.000	
IMD	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
1 = most deprived	−0.158	0.017	−0.154	0.017	−0.160	0.032
2	−0.096	0.017	−0.093	0.017	−0.064	0.031
3	−0.059	0.017	−0.060	0.017	−0.072	0.031
4	−0.032	0.017	−0.033	0.017	−0.032	0.031
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Urban ≥ 10,000 – sparse	−0.227	0.115	−0.233	0.115	−0.361	0.220
Town and fringe – sparse	0.049	0.085	0.033	0.086	−0.005	0.137
Village – sparse	0.078	0.090	0.077	0.091	−0.039	0.150
Hamlet and isolated dwelling – sparse	−0.121	0.112	−0.138	0.113	−0.132	0.211
Urban ≥ 10,000 – less sparse	0.060	0.035	0.061	0.035	0.072	0.060
Town and fringe – less sparse	0.048	0.038	0.049	0.038	0.085	0.065
Village – less sparse	0.099	0.040	0.102	0.041	0.081	0.068
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
North East	−0.089	0.118	−0.084	0.121	−0.120	0.123
North West	0.234	0.092	0.248	0.094	0.452	0.108
Yorkshire and Humber	0.103	0.097	0.113	0.099	0.153	0.114
East Midlands	0.209	0.114	0.209	0.116	0.401	0.130
West Midlands	0.144	0.108	0.144	0.110	0.107	0.118
East of England	0.169	0.108	0.168	0.110	0.099	0.119
London	0.263	0.100	0.266	0.103	0.172	0.121
South East Coast	0.173	0.120	0.173	0.123	0.243	0.166
South Central	0.098	0.114	0.098	0.116	0.168	0.156
South West	0.000		0.000		0.000	
Trust size	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
ONS maternities (thousands)	−0.002	0.015	−0.003	0.015	0.032	0.035
University trust	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Yes	0.000		0.000		0.000	
No	0.125	0.056	0.130	0.057	0.163	0.105
Configuration	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
OU	−0.040	0.071	−0.043	0.072	0.000	
OU/AMU	−0.126	0.081	−0.136	0.082		
OU/AMU/FMU	−0.167	0.097	−0.176	0.099		
OU/FMU	0.000		0.000			

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Staffing variables	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
FTE doctors per 100 maternities	0.149	0.133	0.127	0.136	0.114	0.158
FTE midwives per 100 maternities	0.028	0.062	0.030	0.063	0.018	0.103
FTE support workers per 100 maternities	−0.034	0.071	−0.037	0.072	−0.221	0.112
Random variation (trust level)	σ	$SE(\sigma)$	σ	$SE(\sigma)$	σ	$SE(\sigma)$
Intercept only	0.227	0.011	0.227	0.011	0.217	0.023
Mother level	0.268	0.019	0.274	0.019	0.252	0.029
Mother level, sociodemographics	0.250	0.017	0.257	0.018	0.177	0.021
Mother level, sociodemographics, trust level	0.235	0.016	0.241	0.017	0.173	0.021
Mother level, sociodemographics, trust level, staff 1	0.233	0.016	0.239	0.017	0.163	0.020
Global tests (df)	χ^2	**p-value**	χ^2	**p-value**	χ^2	**p-value**
Mother's age group (6 df)	85.881	< 0.0001	104.197	< 0.0001	17.326	0.0082
Parity (4 df)	2461.647	< 0.0001	2505.192	< 0.0001	604.059	< 0.0001
Clinical risk (1/2 df)[a]	26,717.623	< 0.0001	30,271.981	< 0.0001	7221.888	< 0.0001
Ethnicity (16 df)	89.968	< 0.0001	89.280	< 0.0001	52.427	< 0.0001
IMD (4 df)	103.297	< 0.0001	96.079	< 0.0001	32.140	< 0.0001
Rural/urban classification (7 df)	16.186	0.0235	17.409	0.0149	7.488	0.3799
SHA (9 df)	15.450	0.0793	15.409	0.0803	36.889	< 0.0001
ONS maternities (1 df)	0.019	0.8890	0.051	0.8212	0.874	0.3499
University trust (1 df)	5.005	0.0253	5.155	0.0232	2.430	0.1190
Configuration (3 df)	4.507	0.2116	4.865	0.1820		
FTE doctors per 100 maternities (1 df)	1.246	0.2642	0.869	0.3512	0.523	0.4696
FTE midwives per 100 maternities (1 df)	0.211	0.6456	0.229	0.6320	0.030	0.8618
FTE support workers per 100 maternities (1 df)	0.231	0.6308	0.264	0.6075	3.911	0.0480
AUC	**AUC**	**SE(AUC)**	**AUC**	**SE(AUC)**	**AUC**	**SE(AUC)**
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.720	0.001	0.732	0.001	0.717	0.002
Mother level, sociodemographics	0.724	0.001	0.736	0.001	0.728	0.002
Mother level, sociodemographics, trust level	0.726	0.001	0.738	0.001	0.728	0.002
Mother level, sociodemographics, trust level, staff 1	0.726	0.001	0.738	0.001	0.729	0.002

a　All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Healthy mother/healthy baby dyad: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	SE(β)	β	SE(β)	β	SE(β)
Intercept	–0.107	0.276	–0.092	0.277	0.768	0.419
Mother's age group (years)	β	SE(β)	β	SE(β)	β	SE(β)
≤ 19	1.216	0.097	1.191	0.097	1.155	0.190
20–24	0.905	0.096	0.882	0.096	0.811	0.188
25–29	0.574	0.095	0.552	0.096	0.483	0.188
30–34	0.319	0.095	0.299	0.096	0.248	0.188
35–39	0.158	0.096	0.140	0.096	0.124	0.188
40–44	0.042	0.097	0.022	0.098	0.073	0.191
≥ 45	0.000		0.000		0.000	
Parity	β	SE(β)	β	SE(β)	β	SE(β)
0	–2.596	0.018	–2.588	0.018	–2.500	0.034
1	–1.277	0.016	–1.266	0.016	–1.199	0.032
2	–0.669	0.017	–0.653	0.017	–0.576	0.034
3	–0.319	0.019	–0.297	0.020	–0.242	0.039
≥ 4	0.000		0.000		0.000	
Clinical risk	β	SE(β)	β	SE(β)	β	SE(β)
Lower	0.000		0.000		0.000	
Individual assessment	N/A		–0.429	0.019	N/A	
Higher	–1.267	0.008	–1.357	0.009	–1.223	0.015
Ethnicity	β	SE(β)	β	SE(β)	β	SE(β)
Not given/not known/not stated	0.104	0.033	0.108	0.033	0.098	0.061
English/Welsh/Scottish/Northern Irish/British	0.118	0.026	0.121	0.026	0.085	0.049
Irish	–0.134	0.069	–0.128	0.069	–0.087	0.141
Any other white background	0.086	0.029	0.087	0.029	0.059	0.054
White and black Caribbean	0.486	0.058	0.489	0.059	0.425	0.125
White and black African	0.190	0.072	0.191	0.072	0.118	0.155
White and Asian	–0.091	0.087	–0.100	0.088	–0.121	0.180
Any other mixed/multiple ethnic background	0.170	0.059	0.177	0.059	0.188	0.128
Indian	–0.516	0.037	–0.516	0.037	–0.563	0.072
Pakistani	–0.054	0.033	–0.058	0.033	–0.101	0.071
Bangladeshi	–0.197	0.042	–0.191	0.042	–0.249	0.097
Chinese	–0.350	0.060	–0.351	0.060	–0.284	0.115
Any other Asian background	–0.213	0.039	–0.217	0.040	–0.207	0.079
African	0.067	0.033	0.072	0.033	0.078	0.064
Caribbean	0.505	0.044	0.508	0.044	0.420	0.086
Any other black/African/Caribbean background	0.162	0.047	0.170	0.047	0.271	0.085
Any other ethnic group, please describe	0.000		0.000		0.000	

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
IMD	β	SE(β)	β	SE(β)	β	SE(β)
1 = most deprived	0.280	0.015	0.286	0.015	0.315	0.027
2	0.213	0.015	0.217	0.015	0.252	0.027
3	0.132	0.015	0.134	0.015	0.147	0.027
4	0.077	0.015	0.078	0.015	0.089	0.027
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	β	SE(β)	β	SE(β)	β	SE(β)
Urban ≥ 10,000 – sparse	−0.177	0.101	−0.177	0.101	−0.419	0.252
Town and fringe – sparse	0.037	0.073	0.030	0.073	−0.001	0.123
Village – sparse	0.075	0.076	0.074	0.076	−0.129	0.142
Hamlet and isolated dwelling – sparse	−0.113	0.103	−0.128	0.103	−0.551	0.219
Urban ≥ 10,000 – less sparse	0.073	0.031	0.072	0.031	0.034	0.053
Town and fringe – less sparse	0.114	0.034	0.113	0.034	0.091	0.057
Village – less sparse	0.109	0.036	0.109	0.036	0.054	0.060
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	β	SE(β)	β	SE(β)	β	SE(β)
North East	0.191	0.119	0.199	0.119	0.464	0.116
North West	0.058	0.092	0.068	0.092	0.203	0.101
Yorkshire and Humber	0.205	0.097	0.213	0.098	0.143	0.107
East Midlands	0.225	0.114	0.227	0.115	0.340	0.123
West Midlands	0.066	0.108	0.068	0.109	−0.027	0.112
East of England	0.109	0.108	0.111	0.109	0.088	0.113
London	−0.098	0.101	−0.097	0.101	0.129	0.115
South East Coast	−0.019	0.121	−0.017	0.122	−0.111	0.157
South Central	−0.051	0.114	−0.051	0.115	0.266	0.147
South West	0.000		0.000		0.000	
Trust size	β	SE(β)	β	SE(β)	β	SE(β)
ONS maternities (thousands)	−0.025	0.015	−0.026	0.015	−0.107	0.033
University trust	β	SE(β)	β	SE(β)	β	SE(β)
Yes	0.000		0.000		0.000	
No	0.076	0.056	0.076	0.057	−0.080	0.099
Configuration	β	SE(β)	β	SE(β)	β	SE(β)
OU	−0.001	0.071	−0.001	0.071	0.000	
OU/AMU	−0.077	0.081	−0.080	0.081		
OU/AMU/FMU	−0.044	0.097	−0.046	0.098		
OU/FMU	0.000		0.000			

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Staffing variables	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
FTE doctors per 100 maternities	0.106	0.134	0.093	0.135	–0.197	0.148
FTE midwives per 100 maternities	0.089	0.062	0.091	0.062	–0.002	0.097
FTE support workers per 100 maternities	–0.108	0.071	–0.110	0.071	–0.101	0.105
Random variation (trust level)	*σ*	*SE(σ)*	*σ*	*SE(σ)*	*σ*	*SE(σ)*
Intercept only	0.275	0.011	0.275	0.011	0.210	0.019
Mother level	0.306	0.021	0.310	0.021	0.256	0.029
Mother level, sociodemographics	0.257	0.018	0.259	0.018	0.190	0.022
Mother level, sociodemographics, trust level	0.241	0.017	0.242	0.017	0.162	0.019
Mother level, sociodemographics, trust level, staff 1	0.236	0.016	0.237	0.016	0.155	0.019
Global tests (df)	χ^2	*p-value*	χ^2	*p-value*	χ^2	*p-value*
Mother's age group (6 df)	5099.826	< 0.0001	5009.142	< 0.0001	1408.035	< 0.0001
Parity (4 df)	36,597.737	< 0.0001	36,615.453	< 0.0001	10,520.835	< 0.0001
Clinical risk (1/2 df)[a]	23,436.464	< 0.0001	25,064.757	< 0.0001	6465.242	< 0.0001
Ethnicity (16 df)	999.337	< 0.0001	1007.598	< 0.0001	230.396	< 0.0001
IMD (4 df)	413.126	< 0.0001	427.101	< 0.0001	161.531	< 0.0001
Rural/urban classification (7 df)	26.818	0.0004	27.258	0.0003	17.690	0.0135
SHA (9 df)	18.926	0.0258	19.350	0.0224	30.028	0.0004
ONS maternities (1 df)	2.761	0.0966	2.956	0.0856	10.724	0.0011
University trust (1 df)	1.801	0.1796	1.799	0.1798	0.654	0.4186
Configuration (3 df)	1.653	0.6474	1.798	0.6153		
FTE doctors per 100 maternities (1 df)	0.628	0.4281	0.479	0.4889	1.760	0.1847
FTE midwives per 100 maternities (1 df)	2.037	0.1536	2.117	0.1457	0.000	0.9828
FTE support workers per 100 maternities (1 df)	2.339	0.1262	2.374	0.1234	0.915	0.3387
AUC	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.747	0.001	0.751	0.001	0.741	0.002
Mother level, sociodemographics	0.752	0.001	0.756	0.001	0.748	0.002
Mother level, sociodemographics, trust level	0.753	0.001	0.757	0.001	0.750	0.002
Mother level, sociodemographics, trust level, staff 1	0.753	0.001	0.757	0.001	0.750	0.002

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Delivery with bodily integrity: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Intercept	0.178	0.226	0.188	0.226	0.626	0.408
Mother's age group (years)	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
≤ 19	1.292	0.073	1.278	0.073	1.248	0.155
20–24	0.971	0.072	0.960	0.073	0.913	0.153
25–29	0.598	0.072	0.587	0.072	0.527	0.153
30–34	0.343	0.072	0.333	0.072	0.294	0.153
35–39	0.173	0.072	0.165	0.073	0.136	0.153
40–44	0.054	0.074	0.047	0.074	0.099	0.156
≤ 45	0.000		0.000		0.000	
Parity	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
0	−2.616	0.014	−2.609	0.014	−2.546	0.030
1	−1.483	0.014	−1.474	0.014	−1.401	0.029
2	−0.823	0.014	−0.811	0.014	−0.720	0.030
3	−0.417	0.016	−0.403	0.016	−0.343	0.035
≤ 4	0.000		0.000		0.000	
Clinical risk	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Lower	0.000		0.000		0.000	
Individual assessment	N/A		−0.342	0.015	N/A	
Higher	−0.845	0.006	−0.892	0.007	−0.799	0.012
Ethnicity	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Not given/not known/not stated	0.029	0.025	0.029	0.025	0.045	0.051
English/Welsh/Scottish/Northern Irish/British	0.111	0.020	0.112	0.020	0.128	0.041
Irish	−0.083	0.051	−0.080	0.051	−0.097	0.116
Any other white background	0.059	0.022	0.060	0.022	0.069	0.046
White and black Caribbean	0.532	0.046	0.533	0.046	0.530	0.108
White and black African	0.132	0.057	0.130	0.057	0.083	0.133
White and Asian	−0.042	0.066	−0.047	0.066	0.007	0.140
Any other mixed/multiple ethnic background	0.176	0.045	0.181	0.045	0.185	0.109
Indian	−0.521	0.028	−0.522	0.028	−0.526	0.059
Pakistani	−0.110	0.025	−0.113	0.025	−0.083	0.054
Bangladeshi	−0.233	0.033	−0.230	0.033	−0.351	0.082
Chinese	−0.467	0.046	−0.470	0.046	−0.356	0.097
Any other Asian background	−0.251	0.030	−0.253	0.030	−0.284	0.067
African	0.028	0.025	0.032	0.025	0.062	0.055

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Caribbean	0.587	0.034	0.593	0.034	0.610	0.076
Any other black/African/Caribbean background	0.244	0.036	0.246	0.036	0.256	0.075
Any other ethnic group, please describe	0.000		0.000		0.000	
IMD	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
1 = most deprived	0.376	0.012	0.380	0.012	0.402	0.022
2	0.267	0.011	0.270	0.011	0.296	0.022
3	0.175	0.011	0.177	0.011	0.183	0.022
4	0.098	0.012	0.099	0.012	0.111	0.022
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Urban ≥ 10,000 – sparse	0.011	0.077	0.011	0.077	−0.009	0.154
Town and fringe – sparse	0.107	0.057	0.101	0.057	0.000	0.096
Village – sparse	0.170	0.060	0.168	0.061	0.159	0.106
Hamlet and isolated dwelling – sparse	−0.050	0.081	−0.058	0.081	−0.189	0.145
Urban ≥ 10,000 – less sparse	0.042	0.024	0.042	0.024	0.034	0.044
Town and fringe – less sparse	0.068	0.026	0.068	0.026	0.058	0.048
Village – less sparse	0.057	0.028	0.058	0.028	0.040	0.050
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
North East	0.078	0.106	0.082	0.106	0.361	0.120
North West	−0.056	0.082	−0.051	0.082	−0.019	0.096
Yorkshire and Humber	0.092	0.089	0.096	0.089	0.069	0.104
East Midlands	0.102	0.107	0.104	0.106	0.252	0.128
West Midlands	−0.003	0.089	−0.003	0.089	−0.014	0.112
East of England	−0.005	0.092	−0.004	0.091	0.016	0.118
London	−0.180	0.091	−0.178	0.091	0.121	0.120
South East Coast	−0.169	0.098	−0.170	0.097	0.022	0.148
South Central	−0.095	0.099	−0.095	0.098	0.218	0.135
South West	0.000		0.000		0.000	
Trust size	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
ONS maternities (thousands)	−0.025	0.012	−0.025	0.012	−0.105	0.034
University trust	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Yes	0.000		0.000		0.000	
No	0.074	0.049	0.074	0.048	−0.108	0.098
Configuration	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
OU	0.009	0.065	0.008	0.064	0.000	
OU/AMU	−0.022	0.071	−0.024	0.071		
OU/AMU/FMU	0.048	0.084	0.048	0.084		
OU/FMU	0.000		0.000			

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Staffing variables	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
FTE doctors per 100 maternities	0.096	0.115	0.088	0.114	−0.131	0.149
FTE midwives per 100 maternities	0.105	0.051	0.103	0.051	0.113	0.093
FTE support workers per 100 maternities	0.000	0.063	−0.001	0.063	−0.026	0.109
Random variation (trust level)	*σ*	*SE(σ)*	*σ*	*SE(σ)*	*σ*	*SE(σ)*
Intercept only	0.283	0.006	0.283	0.006	0.253	0.016
Mother level	0.297	0.018	0.299	0.018	0.260	0.027
Mother level, sociodemographics	0.244	0.015	0.245	0.015	0.203	0.021
Mother level, sociodemographics, trust level	0.233	0.014	0.234	0.014	0.173	0.018
Mother level, sociodemographics, trust level, staff 1	0.230	0.014	0.229	0.014	0.169	0.018
Global tests (df)	χ^2	**p-value**	χ^2	**p-value**	χ^2	**p-value**
Mother's age group (6 df)	9408.646	< 0.0001	9312.553	< 0.0001	2422.172	< 0.0001
Parity (4 df)	56,739.596	< 0.0001	56,694.660	< 0.0001	14,598.140	< 0.0001
Clinical risk (1/2 df)[a]	17,470.422	< 0.0001	18,585.809	< 0.0001	4107.283	< 0.0001
Ethnicity (16 df)	1976.990	< 0.0001	1996.954	< 0.0001	441.925	< 0.0001
IMD (4 df)	1199.651	< 0.0001	1218.406	< 0.0001	370.907	< 0.0001
Rural/urban classification (7 df)	16.701	0.0194	16.623	0.0200	6.460	0.4872
SHA (9 df)	18.824	0.0267	19.137	0.0240	19.976	0.0181
ONS maternities (1 df)	4.172	0.0411	4.278	0.0386	9.702	0.0018
University trust (1 df)	2.321	0.1276	2.303	0.1291	1.227	0.2680
Configuration (3 df)	1.030	0.7940	1.079	0.7821		
FTE doctors per 100 maternities (1 df)	0.698	0.4035	0.593	0.4412	0.775	0.3786
FTE midwives per 100 maternities (1 df)	4.223	0.0399	4.151	0.0416	1.477	0.2243
FTE support workers per 100 maternities (1 df)	0.000	0.9951	0.000	0.9852	0.056	0.8126
AUC	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.724	0.001	0.726	0.001	0.721	0.001
Mother level, sociodemographics	0.732	0.001	0.733	0.001	0.729	0.001
Mother level, sociodemographics, trust level	0.732	0.001	0.734	0.001	0.730	0.001
Mother level, sociodemographics, trust level, staff 1	0.733	0.001	0.734	0.001	0.731	0.001

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Normal birth: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	SE(β)	β	SE(β)	β	SE(β)
Intercept	0.509	0.222	0.526	0.225	1.209	0.529
Mother's age group (years)	β	SE(β)	β	SE(β)	β	SE(β)
≤ 19	1.415	0.090	1.394	0.091	1.431	0.189
20–24	1.179	0.090	1.162	0.090	1.173	0.187
25–29	0.986	0.089	0.969	0.090	0.956	0.187
30–34	0.830	0.089	0.816	0.090	0.855	0.187
35–39	0.685	0.090	0.673	0.090	0.713	0.187
40–44	0.419	0.091	0.404	0.092	0.493	0.190
≤ 45	0.000		0.000		0.000	
Parity	β	SE(β)	β	SE(β)	β	SE(β)
0	−1.751	0.016	−1.739	0.016	−1.722	0.034
1	−0.607	0.015	−0.586	0.016	−0.543	0.032
2	−0.351	0.016	−0.324	0.017	−0.269	0.035
3	−0.199	0.019	−0.167	0.019	−0.189	0.040
≤ 4	0.000		0.000		0.000	
Clinical risk	β	SE(β)	β	SE(β)	β	SE(β)
Lower	0.000		0.000		0.000	
Individual assessment	N/A		−0.852	0.016	N/A	
Higher	−1.822	0.007	−1.926	0.008	−1.818	0.014
Ethnicity	β	SE(β)	β	SE(β)	β	SE(β)
Not given/not known/not stated	0.010	0.027	0.014	0.027	−0.014	0.054
English/Welsh/Scottish/Northern Irish/British	−0.108	0.022	−0.104	0.022	−0.065	0.045
Irish	−0.314	0.056	−0.307	0.056	−0.172	0.123
Any other white background	−0.025	0.024	−0.022	0.024	0.028	0.050
White and black Caribbean	0.159	0.053	0.162	0.054	0.396	0.125
White and black African	0.129	0.064	0.122	0.064	0.018	0.147
White and Asian	−0.106	0.073	−0.115	0.073	−0.106	0.154
Any other mixed/multiple ethnic background	−0.018	0.051	−0.012	0.051	−0.066	0.120
Indian	−0.205	0.029	−0.207	0.029	−0.248	0.062
Pakistani	0.081	0.028	0.081	0.028	0.063	0.066
Bangladeshi	0.107	0.036	0.119	0.036	−0.012	0.092
Chinese	0.182	0.045	0.177	0.045	0.353	0.097
Any other Asian background	−0.042	0.032	−0.040	0.032	−0.119	0.073
African	0.036	0.028	0.046	0.028	0.024	0.064
Caribbean	0.223	0.040	0.231	0.040	0.274	0.099
Any other black/African/Caribbean background	0.099	0.044	0.108	0.044	0.141	0.115
Any other ethnic group, please describe	0.000		0.000		0.000	

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
IMD	β	SE(β)	β	SE(β)	β	SE(β)
1 = most deprived	0.118	0.013	0.124	0.013	0.143	0.024
2	0.068	0.012	0.073	0.012	0.094	0.023
3	0.022	0.012	0.025	0.012	0.046	0.023
4	0.010	0.012	0.010	0.012	0.029	0.023
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	β	SE(β)	β	SE(β)	β	SE(β)
Urban ≥ 10,000 – sparse	−0.083	0.084	−0.085	0.085	−0.169	0.161
Town and fringe – sparse	0.108	0.063	0.102	0.063	0.032	0.098
Village – sparse	0.199	0.066	0.202	0.066	0.041	0.110
Hamlet and isolated dwelling – sparse	−0.049	0.089	−0.061	0.089	−0.185	0.146
Urban ≥ 10,000 – less sparse	−0.024	0.026	−0.024	0.026	0.010	0.046
Town and fringe – less sparse	−0.001	0.029	−0.002	0.029	0.018	0.049
Village – less sparse	0.006	0.030	0.007	0.030	0.020	0.052
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	β	SE(β)	β	SE(β)	β	SE(β)
North East	0.072	0.099	0.078	0.100	0.140	0.127
North West	0.112	0.076	0.120	0.077	0.074	0.104
Yorkshire and Humber	0.130	0.082	0.136	0.083	0.174	0.116
East Midlands	0.236	0.095	0.236	0.097	0.417	0.136
West Midlands	0.179	0.087	0.182	0.088	0.146	0.122
East of England	0.128	0.086	0.129	0.087	0.078	0.126
London	−0.033	0.088	−0.030	0.089	0.163	0.168
South East Coast	0.067	0.099	0.062	0.100	0.269	0.180
South Central	0.184	0.100	0.185	0.101	0.456	0.164
South West	0.000		0.000		0.000	
Trust size	β	SE(β)	β	SE(β)	β	SE(β)
ONS maternities (thousands)	−0.017	0.011	−0.018	0.012	−0.134	0.040
University trust	β	SE(β)	β	SE(β)	β	SE(β)
Yes	0.000		0.000		0.000	
No	0.042	0.046	0.043	0.046	−0.207	0.106
Configuration	β	SE(β)	β	SE(β)	β	SE(β)
OU	−0.122	0.058	−0.128	0.059	0.000	
OU/AMU	−0.076	0.067	−0.087	0.068		
OU/AMU/FMU	−0.120	0.081	−0.125	0.082		
OU/FMU	0.000		0.000			

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Staffing variables	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
FTE doctors per 100 maternities	−0.075	0.113	−0.094	0.115	−0.047	0.168
FTE midwives per 100 maternities	0.060	0.048	0.062	0.048	−0.101	0.117
FTE support workers per 100 maternities	0.010	0.060	0.009	0.061	0.071	0.121
Random variation (trust level)	*σ*	*SE(σ)*	*σ*	*SE(σ)*	*σ*	*SE(σ)*
Intercept only	0.270	0.012	0.270	0.012	0.214	0.021
Mother level	0.225	0.015	0.229	0.015	0.236	0.026
Mother level, sociodemographics	0.209	0.014	0.212	0.014	0.205	0.023
Mother level, sociodemographics, trust level	0.199	0.013	0.203	0.014	0.180	0.021
Mother level, sociodemographics, trust level, staff 1	0.198	0.013	0.201	0.014	0.179	0.020
Global tests (df)	*χ²*	*p-value*	*χ²*	*p-value*	*χ²*	*p-value*
Mother's age group (6 df)	3455.872	< 0.0001	3359.947	< 0.0001	829.814	< 0.0001
Parity (4 df)	28,952.972	< 0.0001	29,090.456	< 0.0001	7942.722	< 0.0001
Clinical risk (1/2 df)[a]	63,029.595	< 0.0001	65,791.942	< 0.0001	16,606.664	< 0.0001
Ethnicity (16 df)	452.266	< 0.0001	454.095	< 0.0001	97.427	< 0.0001
IMD (4 df)	126.536	< 0.0001	137.401	< 0.0001	43.605	< 0.0001
Rural/urban classification (7 df)	23.057	0.0017	22.878	0.0018	3.690	0.8147
SHA (9 df)	15.832	0.0705	15.460	0.079	15.829	0.0705
ONS maternities (1 df)	2.112	0.1461	2.383	0.1227	11.485	0.0007
University trust (1 df)	0.856	0.3549	0.872	0.3504	3.827	0.0504
Configuration (3 df)	4.877	0.1810	5.028	0.1698		
FTE doctors per 100 maternities (1 df)	0.433	0.5107	0.669	0.4134	0.078	0.7807
FTE midwives per 100 maternities (1 df)	1.608	0.2048	1.623	0.2027	0.750	0.3864
FTE support workers per 100 maternities (1 df)	0.030	0.8614	0.023	0.8785	0.346	0.5562
AUC	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*	*AUC*	*SE(AUC)*
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.754	0.001	0.761	0.001	0.752	0.001
Mother level, sociodemographics	0.756	0.001	0.763	0.001	0.756	0.001
Mother level, sociodemographics, trust level	0.756	0.001	0.763	0.001	0.758	0.001
Mother level, sociodemographics, trust level, staff 1	0.756	0.001	0.763	0.001	0.758	0.001

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Spontaneous vaginal delivery: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	SE(β)	β	SE(β)	β	SE(β)
Intercept	1.250	0.181	1.273	0.183	1.846	0.337
Mother's age group (years)	β	SE(β)	β	SE(β)	β	SE(β)
≤ 19	1.771	0.065	1.756	0.065	1.668	0.136
20–24	1.406	0.064	1.393	0.064	1.258	0.134
25–29	1.088	0.064	1.074	0.064	0.911	0.134
30–34	0.842	0.064	0.829	0.064	0.693	0.134
35–39	0.602	0.064	0.591	0.064	0.453	0.134
40–44	0.388	0.065	0.377	0.065	0.305	0.137
≤ 45	0.000		0.000		0.000	
Parity	β	SE(β)	β	SE(β)	β	SE(β)
0	−1.610	0.015	−1.601	0.015	−1.557	0.031
1	−0.670	0.015	−0.651	0.015	−0.605	0.031
2	−0.401	0.016	−0.378	0.016	−0.330	0.033
3	−0.225	0.018	−0.200	0.018	−0.209	0.038
≤ 4	0.000		0.000		0.000	
Clinical risk	β	SE(β)	β	SE(β)	β	SE(β)
Lower	0.000		0.000		0.000	
Individual assessment	N/A		−0.526	0.015	N/A	
Higher	−1.417	0.006	−1.491	0.006	−1.396	0.013
Ethnicity	β	SE(β)	β	SE(β)	β	SE(β)
Not given/not known/not stated	0.044	0.023	0.047	0.023	−0.024	0.048
English/Welsh/Scottish/Northern Irish/British	−0.001	0.018	0.001	0.018	−0.018	0.040
Irish	−0.161	0.045	−0.156	0.046	−0.170	0.104
Any other white background	0.044	0.020	0.045	0.020	0.027	0.044
White and black Caribbean	0.265	0.048	0.268	0.049	0.481	0.119
White and black African	0.034	0.056	0.022	0.056	−0.126	0.129
White and Asian	−0.135	0.061	−0.143	0.061	−0.263	0.131
Any other mixed/multiple ethnic background	0.047	0.044	0.046	0.044	−0.110	0.105
Indian	−0.187	0.024	−0.189	0.024	−0.230	0.052
Pakistani	0.069	0.024	0.067	0.024	0.100	0.053
Bangladeshi	0.059	0.032	0.065	0.032	−0.031	0.080
Chinese	0.111	0.040	0.103	0.040	0.300	0.088
Any other Asian background	−0.087	0.027	−0.089	0.027	−0.236	0.060
African	−0.039	0.024	−0.033	0.024	−0.105	0.053

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Caribbean	0.188	0.034	0.194	0.035	0.199	0.077
Any other black/African/Caribbean background	0.125	0.035	0.132	0.036	0.155	0.075
Any other ethnic group, please describe	0.000		0.000		0.000	
IMD	***β***	***SE(β)***	***β***	***SE(β)***	***β***	***SE(β)***
1 = most deprived	0.096	0.011	0.101	0.011	0.126	0.021
2	0.059	0.011	0.063	0.011	0.080	0.021
3	0.029	0.010	0.032	0.011	0.047	0.020
4	0.018	0.011	0.019	0.011	0.046	0.020
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	***β***	***SE(β)***	***β***	***SE(β)***	***β***	***SE(β)***
Urban ≥ 10,000 – sparse	−0.031	0.079	−0.032	0.079	−0.221	0.149
Town and fringe – sparse	0.089	0.057	0.082	0.058	0.025	0.094
Village – sparse	0.218	0.061	0.219	0.061	0.159	0.106
Hamlet and isolated dwelling – sparse	−0.073	0.078	−0.080	0.078	−0.179	0.135
Urban ≥ 10,000 – less sparse	0.010	0.023	0.011	0.023	0.017	0.041
Town and fringe – less sparse	0.052	0.025	0.054	0.025	0.055	0.045
Village – less sparse	0.053	0.026	0.056	0.026	0.041	0.047
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	***β***	***SE(β)***	***β***	***SE(β)***	***β***	***SE(β)***
North East	−0.039	0.084	−0.035	0.085	0.045	0.098
North West	0.092	0.065	0.098	0.066	0.191	0.079
Yorkshire and Humber	0.047	0.070	0.051	0.071	0.024	0.085
East Midlands	0.062	0.084	0.059	0.085	0.150	0.104
West Midlands	0.028	0.070	0.025	0.071	−0.019	0.092
East of England	−0.027	0.072	−0.029	0.073	−0.047	0.096
London	−0.099	0.071	−0.100	0.072	0.036	0.097
South East Coast	−0.045	0.077	−0.052	0.078	0.182	0.120
South Central	−0.040	0.078	−0.040	0.079	0.110	0.111
South West	0.000		0.000		0.000	
Trust size	***β***	***SE(β)***	***β***	***SE(β)***	***β***	***SE(β)***
ONS maternities (thousands)	−0.011	0.010	−0.013	0.010	−0.072	0.028
University trust	***β***	***SE(β)***	***β***	***SE(β)***	***β***	***SE(β)***
Yes	0.000		0.000		0.000	
No	0.086	0.038	0.089	0.039	−0.096	0.079
Configuration	***β***	***SE(β)***	***β***	***SE(β)***	***β***	***SE(β)***
OU	−0.030	0.051	−0.034	0.051	0.000	
OU/AMU	−0.020	0.056	−0.027	0.057		
OU/AMU/FMU	−0.012	0.066	−0.018	0.067		
OU/FMU	0.000		0.000			

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Staffing variables	β	SE(β)	β	SE(β)	β	SE(β)
FTE doctors per 100 maternities	0.017	0.090	0.001	0.091	0.154	0.122
FTE midwives per 100 maternities	0.025	0.040	0.025	0.041	–0.104	0.075
FTE support workers per 100 maternities	–0.039	0.049	–0.042	0.050	–0.013	0.089
Random variation (trust level)	σ	SE(σ)	σ	SE(σ)	σ	SE(σ)
Intercept only	0.226	0.006	0.226	0.006	0.177	0.014
Mother level	0.200	0.012	0.204	0.013	0.167	0.018
Mother level, sociodemographics	0.186	0.011	0.189	0.012	0.152	0.016
Mother level, sociodemographics, trust level	0.180	0.011	0.182	0.011	0.141	0.015
Mother level, sociodemographics, trust level, staff 1	0.179	0.011	0.182	0.011	0.136	0.015
Global tests (df)	χ^2	p-value	χ^2	p-value	χ^2	p-value
Mother's age group (6 df)	10,476.241	< 0.0001	10,328.463	< 0.0001	2752.466	< 0.0001
Parity (4 df)	30,418.259	< 0.0001	30,538.921	< 0.0001	7757.249	< 0.0001
Clinical risk (1/2 df)[a]	49,583.344	< 0.0001	53,674.530	< 0.0001	12,437.977	< 0.0001
Ethnicity (16 df)	323.687	< 0.0001	327.126	< 0.0001	140.068	< 0.0001
IMD (4 df)	91.541	< 0.0001	101.717	< 0.0001	37.967	< 0.0001
Rural/urban classification (7 df)	33.710	< 0.0001	33.966	< 0.0001	10.891	0.1435
SHA (9 df)	12.780	0.1728	13.215	0.1531	13.802	0.1296
ONS maternities (1 df)	1.438	0.2304	1.773	0.1830	6.896	0.0086
University trust (1 df)	5.109	0.0238	5.255	0.0219	1.451	0.2284
Configuration (3 df)	0.381	0.9440	0.453	0.9291		
FTE doctors per 100 maternities (1 df)	0.037	0.8480	0.000	0.9941	1.595	0.2067
FTE midwives per 100 maternities (1 df)	0.383	0.5362	0.367	0.5446	1.884	0.1699
FTE support workers per 100 maternities (1 df)	0.614	0.4334	0.698	0.4035	0.022	0.8815
AUC	AUC	SE(AUC)	AUC	SE(AUC)	AUC	SE(AUC)
Intercept only	0.500	0.000	0.500	0.0000	0.500	0.000
Mother level	0.722	0.001	0.729	0.0010	0.720	0.001
Mother level, sociodemographics	0.724	0.001	0.731	0.0010	0.722	0.001
Mother level, sociodemographics, trust level	0.724	0.001	0.731	0.0010	0.723	0.001
Mother level, sociodemographics, trust level, staff 1	0.724	0.001	0.731	0.0010	0.723	0.001

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Intact perineum: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Intercept	1.079	0.263	1.076	0.263	1.567	0.524
Mother's age group (years)	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
≤ 19	0.626	0.098	0.626	0.098	0.667	0.198
20–24	0.375	0.097	0.374	0.097	0.402	0.197
25–29	0.027	0.097	0.026	0.097	0.046	0.197
30–34	−0.190	0.097	−0.191	0.097	−0.149	0.197
35–39	−0.272	0.097	−0.273	0.097	−0.224	0.197
40–44	−0.267	0.099	−0.267	0.099	−0.134	0.201
≥ 45	0.000		0.000		0.000	
Parity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
0	−2.913	0.018	−2.915	0.018	−2.912	0.039
1	−1.772	0.018	−1.773	0.018	−1.747	0.039
2	−1.007	0.019	−1.008	0.019	−0.928	0.041
3	−0.524	0.021	−0.525	0.021	−0.451	0.047
≥ 4	0.000		0.000		0.000	
Clinical risk	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Lower	0.000		0.000		0.000	
Individual assessment	N/A		−0.246	0.016	N/A	
Higher	−0.212	0.007	−0.209	0.007	−0.147	0.014
Ethnicity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Not given/not known/not stated	0.011	0.027	0.011	0.027	0.029	0.055
English/Welsh/Scottish/Northern Irish/British	0.120	0.022	0.120	0.022	0.145	0.045
Irish	−0.045	0.056	−0.045	0.056	−0.058	0.130
Any other white background	0.044	0.024	0.044	0.024	0.066	0.050
White and black Caribbean	0.602	0.052	0.602	0.052	0.533	0.120
White and black African	0.184	0.064	0.183	0.064	0.279	0.154
White and Asian	−0.026	0.072	−0.025	0.072	0.044	0.154
Any other mixed/multiple ethnic background	0.221	0.050	0.219	0.050	0.251	0.122
Indian	−0.572	0.030	−0.572	0.030	−0.567	0.064
Pakistani	−0.224	0.027	−0.222	0.027	−0.186	0.060
Bangladeshi	−0.365	0.037	−0.365	0.037	−0.441	0.091
Chinese	−0.574	0.049	−0.574	0.049	−0.488	0.102
Any other Asian background	−0.284	0.032	−0.284	0.032	−0.252	0.073
African	0.173	0.028	0.174	0.028	0.299	0.063
Caribbean	0.807	0.040	0.807	0.040	0.846	0.089
Any other black/African/Caribbean background	0.373	0.041	0.374	0.041	0.434	0.087
Any other ethnic group, please describe	0.000		0.000		0.000	

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
IMD	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
1 = most deprived	0.436	0.013	0.435	0.013	0.460	0.025
2	0.312	0.012	0.312	0.012	0.343	0.024
3	0.194	0.012	0.194	0.012	0.192	0.024
4	0.117	0.013	0.117	0.013	0.117	0.024
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Urban ≥ 10,000 – sparse	0.012	0.086	0.011	0.086	0.096	0.178
Town and fringe – sparse	0.107	0.063	0.109	0.063	−0.020	0.106
Village – sparse	0.154	0.067	0.155	0.067	0.117	0.119
Hamlet and isolated dwelling – sparse	−0.049	0.089	−0.046	0.089	−0.152	0.162
Urban ≥ 10,000 – less sparse	0.047	0.026	0.047	0.026	0.012	0.048
Town and fringe – less sparse	0.067	0.029	0.067	0.029	0.024	0.052
Village – less sparse	0.058	0.030	0.058	0.030	0.016	0.055
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
North East	0.100	0.121	0.099	0.121	0.372	0.153
North West	−0.111	0.094	−0.112	0.094	−0.125	0.123
Yorkshire and Humber	0.089	0.101	0.089	0.101	0.013	0.133
East Midlands	0.074	0.122	0.074	0.121	0.183	0.164
West Midlands	−0.013	0.102	−0.012	0.102	−0.057	0.144
East of England	0.023	0.104	0.025	0.104	0.016	0.151
London	−0.157	0.104	−0.156	0.103	0.081	0.153
South East Coast	−0.183	0.112	−0.183	0.111	−0.046	0.190
South Central	−0.108	0.113	−0.107	0.112	0.179	0.173
South West	0.000		0.000		0.000	
Trust size	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
ONS maternities (thousands)	−0.030	0.014	−0.029	0.014	−0.107	0.043
University trust	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Yes	0.000		0.000		0.000	
No	0.063	0.055	0.063	0.055	−0.097	0.126
Configuration	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
OU	0.030	0.074	0.029	0.073	0.000	
OU/AMU	−0.037	0.081	−0.038	0.081		
OU/AMU/FMU	0.069	0.096	0.068	0.096		
OU/FMU	0.000		0.000			

Staffing variables	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	SE(β)	β	SE(β)	β	SE(β)
FTE doctors per 100 maternities	0.068	0.131	0.069	0.130	−0.222	0.191
FTE midwives per 100 maternities	0.124	0.058	0.125	0.058	0.147	0.119
FTE support workers per 100 maternities	0.017	0.072	0.017	0.072	−0.035	0.140
Random variation (trust level)	σ	SE(σ)	σ	SE(σ)	σ	SE(σ)
Intercept only	0.283	0.006	0.283	0.006	0.266	0.017
Mother level	0.334	0.020	0.335	0.020	0.310	0.032
Mother level, sociodemographics	0.280	0.017	0.280	0.017	0.249	0.026
Mother level, sociodemographics, trust level	0.267	0.016	0.267	0.016	0.224	0.023
Mother level, sociodemographics, trust level, staff 1	0.263	0.016	0.262	0.016	0.219	0.023
Global tests (df)	χ^2	p-value	χ^2	p-value	χ^2	p-value
Mothers' age group (6 df)	5489.165	< 0.0001	5491.286	< 0.0001	1415.394	< 0.0001
Parity (4 df)	53,241.997	< 0.0001	53,228.888	< 0.0001	13,921.949	< 0.0001
Clinical risk (1/2 df)[a]	945.367	< 0.0001	950.803	< 0.0001	118.292	< 0.0001
Ethnicity (16 df)	2521.852	< 0.0001	2519.513	< 0.0001	545.690	< 0.0001
IMD (4 df)	1347.772	< 0.0001	1343.963	< 0.0001	416.223	< 0.0001
Rural/urban classification (7 df)	11.543	0.1166	11.635	0.1132	2.818	0.9013
SHA (9 df)	16.128	0.0643	16.232	0.0622	16.006	0.0668
ONS maternities (1 df)	4.522	0.0335	4.414	0.0356	6.060	0.0138
University trust (1 df)	1.276	0.2587	1.313	0.2518	0.594	0.4407
Configuration (3 df)	2.187	0.5346	2.187	0.5345		
FTE doctors per 100 maternities (1 df)	0.272	0.6022	0.278	0.5982	1.342	0.2467
FTE midwives per 100 maternities (1 df)	4.577	0.0324	4.678	0.0305	1.533	0.2156
FTE support workers per 100 maternities (1 df)	0.058	0.8095	0.056	0.8126	0.063	0.8019
AUC	AUC	SE(AUC)	AUC	SE(AUC)	AUC	SE(AUC)
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.722	0.001	0.722	0.001	0.721	0.002
Mother level, sociodemographics	0.731	0.001	0.731	0.001	0.733	0.001
Mother level, sociodemographics, trust level	0.732	0.001	0.732	0.001	0.734	0.001
Mother level, sociodemographics, trust level, staff 1	0.732	0.001	0.732	0.001	0.735	0.001

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Elective caesarean: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	SE(β)	β	SE(β)	β	SE(β)
Intercept	–3.463	0.224	–3.463	0.228	–4.115	0.462
Mother's age group (years)	β	SE(β)	β	SE(β)	β	SE(β)
≤ 19	–2.037	0.077	–2.005	0.078	–1.900	0.163
20–24	–1.589	0.069	–1.566	0.069	–1.389	0.150
25–29	–1.261	0.068	–1.238	0.069	–1.089	0.148
30–34	–0.993	0.068	–0.973	0.068	–0.845	0.148
35–39	–0.732	0.068	–0.717	0.068	–0.601	0.148
40–44	–0.535	0.069	–0.520	0.070	–0.439	0.152
≥ 45	0.000		0.000		0.000	
Parity	β	SE(β)	β	SE(β)	β	SE(β)
0	–0.565	0.021	–0.611	0.021	–0.686	0.043
1	0.386	0.020	0.347	0.020	0.380	0.040
2	0.446	0.021	0.407	0.021	0.425	0.043
3	0.283	0.024	0.240	0.024	0.232	0.050
≤ 4	0.000		0.000		0.000	
Clinical risk	β	SE(β)	β	SE(β)	β	SE(β)
Lower	0.000		0.000		0.000	
Individual assessment	N/A		1.264	0.041	N/A	
Higher	3.121	0.022	3.206	0.022	3.242	0.045
Ethnicity	β	SE(β)	β	SE(β)	β	SE(β)
Not given/not known/not stated	0.004	0.039	–0.005	0.039	0.042	0.082
English/Welsh/Scottish/Northern Irish/British	0.061	0.030	0.056	0.030	0.053	0.066
Irish	0.201	0.066	0.192	0.067	0.191	0.155
Any other white background	0.065	0.033	0.064	0.033	0.105	0.073
White and black Caribbean	–0.367	0.086	–0.367	0.087	–0.513	0.216
White and black African	0.069	0.085	0.080	0.086	0.371	0.184
White and Asian	0.132	0.095	0.147	0.096	0.464	0.202
Any other mixed/multiple ethnic background	0.144	0.068	0.144	0.068	0.334	0.161
Indian	–0.009	0.039	–0.011	0.039	0.013	0.084
Pakistani	–0.264	0.038	–0.262	0.039	–0.314	0.088
Bangladeshi	–0.357	0.053	–0.376	0.054	–0.245	0.131
Chinese	–0.137	0.069	–0.128	0.069	–0.158	0.152
Any other Asian background	0.002	0.043	0.001	0.043	0.162	0.095
African	–0.088	0.037	–0.103	0.037	0.026	0.083

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Caribbean	−0.262	0.056	−0.273	0.056	−0.141	0.122
Any other black/African/Caribbean background	−0.221	0.057	−0.233	0.058	−0.112	0.121
Any other ethnic group, please describe	0.000		0.000		0.000	
IMD	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
1 = most deprived	−0.204	0.017	−0.215	0.017	−0.191	0.033
2	−0.131	0.016	−0.142	0.017	−0.099	0.032
3	−0.099	0.016	−0.106	0.016	−0.092	0.032
4	−0.024	0.016	−0.026	0.016	−0.055	0.031
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Urban ≥ 10,000 – sparse	0.029	0.123	0.034	0.124	0.261	0.222
Town and fringe – sparse	−0.007	0.089	0.004	0.090	−0.022	0.149
Village – sparse	−0.145	0.095	−0.144	0.095	0.009	0.159
Hamlet and isolated dwelling – sparse	0.064	0.118	0.081	0.119	0.143	0.201
Urban ≥ 10,000 – less sparse	−0.057	0.034	−0.056	0.034	−0.052	0.063
Town and fringe – less sparse	−0.063	0.038	−0.064	0.038	−0.064	0.069
Village – less sparse	−0.031	0.039	−0.031	0.039	−0.042	0.072
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
North East	−0.051	0.105	−0.066	0.107	−0.043	0.140
North West	−0.111	0.081	−0.128	0.082	−0.077	0.111
Yorkshire and Humber	−0.078	0.087	−0.094	0.089	−0.030	0.120
East Midlands	−0.145	0.104	−0.151	0.106	−0.095	0.146
West Midlands	−0.104	0.087	−0.113	0.089	0.063	0.129
East of England	−0.035	0.090	−0.042	0.091	−0.060	0.135
London	0.093	0.089	0.085	0.091	0.083	0.137
South East Coast	−0.014	0.095	−0.013	0.097	0.039	0.168
South Central	−0.119	0.097	−0.125	0.099	−0.085	0.157
South West	0.000		0.000		0.000	
Trust size	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
ONS maternities (thousands)	−0.004	0.012	−0.003	0.012	0.032	0.039
University trust	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Yes	0.000		0.000		0.000	
No	−0.062	0.047	−0.064	0.048	0.023	0.112
Configuration	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
OU	0.039	0.063	0.043	0.065	0.000	
OU/AMU	−0.036	0.070	−0.027	0.071		
OU/AMU/FMU	−0.021	0.082	−0.017	0.084		
OU/FMU	0.000		0.000			

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Staffing variables	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
FTE doctors per 100 maternities	−0.144	0.112	−0.123	0.114	−0.243	0.171
FTE midwives per 100 maternities	0.031	0.050	0.033	0.051	0.164	0.106
FTE support workers per 100 maternities	0.075	0.062	0.078	0.063	−0.047	0.125
Random variation (trust level)	σ	$SE(\sigma)$	σ	$SE(\sigma)$	σ	$SE(\sigma)$
Intercept only	0.206	0.008	0.206	0.008	0.172	0.017
Mother level	0.233	0.015	0.239	0.015	0.208	0.023
Mother level, sociodemographics	0.225	0.014	0.230	0.015	0.197	0.022
Mother level, sociodemographics, trust level	0.222	0.014	0.227	0.014	0.196	0.022
Mother level, sociodemographics, trust level, staff 1	0.219	0.014	0.224	0.014	0.188	0.021
Global tests (df)	χ^2	**p-value**	χ^2	**p-value**	χ^2	**p-value**
Mother's age group (6 df)	4276.429	< 0.0001	4131.216	< 0.0001	968.450	< 0.0001
Parity (4 df)	7138.177	< 0.0001	7113.285	< 0.0001	2250.082	< 0.0001
Clinical risk (1/2 df)[a]	20,858.460	< 0.0001	24,380.611	< 0.0001	5232.044	< 0.0001
Ethnicity (16 df)	328.588	< 0.0001	330.503	< 0.0001	75.705	< 0.0001
IMD (4 df)	170.993	< 0.0001	189.110	< 0.0001	34.369	< 0.0001
Rural/urban classification (7 df)	7.135	0.4150	7.521	0.3767	3.897	0.7916
SHA (9 df)	12.153	0.2048	12.599	0.1816	3.774	0.9257
ONS maternities (1 df)	0.110	0.7401	0.044	0.8330	0.688	0.4069
University trust (1 df)	1.701	0.1922	1.739	0.1873	0.041	0.8400
Configuration (3 df)	2.392	0.4951	2.158	0.5404		
FTE doctors per 100 maternities (1 df)	1.656	0.1981	1.156	0.2824	2.013	0.1559
FTE midwives per 100 maternities (1 df)	0.394	0.5303	0.416	0.5191	2.393	0.1219
FTE support workers per 100 maternities (1 df)	1.465	0.2262	1.530	0.2161	0.139	0.7095
AUC	**AUC**	**SE(AUC)**	**AUC**	**SE(AUC)**	**AUC**	**SE(AUC)**
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.811	0.001	0.822	0.001	0.818	0.001
Mother level, sociodemographics	0.814	0.001	0.825	0.001	0.820	0.001
Mother level, sociodemographics, trust level	0.814	0.001	0.825	0.001	0.820	0.001
Mother level, sociodemographics, trust level, staff 1	0.814	0.001	0.825	0.001	0.821	0.001

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Emergency caesarean: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Intercept	–2.696	0.201	–2.709	0.201	–2.928	0.389
Mother's age group (years)	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
≤ 19	–1.054	0.073	–1.037	0.073	–0.949	0.159
20–24	–0.735	0.071	–0.721	0.071	–0.618	0.157
25–29	–0.535	0.071	–0.521	0.071	–0.361	0.156
30–34	–0.374	0.071	–0.361	0.071	–0.207	0.156
35–39	–0.212	0.071	–0.201	0.071	–0.068	0.156
40–44	–0.049	0.072	–0.039	0.073	0.089	0.160
≥ 45	0.000		0.000		0.000	
Parity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
0	1.356	0.020	1.345	0.020	1.364	0.043
1	0.428	0.020	0.414	0.020	0.354	0.044
2	0.122	0.022	0.107	0.022	0.041	0.048
3	0.070	0.026	0.055	0.026	0.068	0.055
≥ 4	0.000		0.000		0.000	
Clinical risk	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Lower	0.000		0.000		0.000	
Individual assessment	N/A		0.687	0.020	N/A	
Higher	1.183	0.009	1.219	0.009	1.174	0.017
Ethnicity	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Not given/not known/not stated	–0.141	0.029	–0.142	0.029	–0.126	0.061
English/Welsh/Scottish/Northern Irish/British	–0.129	0.023	–0.130	0.023	–0.116	0.051
Irish	–0.054	0.058	–0.058	0.058	–0.006	0.135
Any other white background	–0.174	0.026	–0.174	0.026	–0.193	0.057
White and black Caribbean	0.018	0.060	0.018	0.060	–0.120	0.152
White and black African	0.015	0.072	0.025	0.072	0.069	0.166
White and Asian	–0.041	0.079	–0.037	0.079	0.032	0.168
Any other mixed/multiple ethnic background	–0.102	0.056	–0.100	0.057	–0.069	0.139
Indian	0.118	0.030	0.118	0.030	0.156	0.065
Pakistani	–0.018	0.030	–0.018	0.030	–0.011	0.069
Bangladeshi	0.103	0.040	0.100	0.040	0.210	0.101
Chinese	–0.170	0.052	–0.165	0.052	–0.256	0.116
Any other Asian background	0.093	0.033	0.093	0.033	0.141	0.075
African	0.374	0.029	0.371	0.029	0.411	0.065
Caribbean	0.196	0.042	0.194	0.042	0.085	0.097
Any other black/African/Caribbean background	0.276	0.042	0.273	0.042	0.201	0.092
Any other ethnic group, please describe	0.000		0.000		0.000	

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
IMD	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
1 = most deprived	0.107	0.015	0.104	0.015	0.080	0.028
2	0.086	0.014	0.084	0.014	0.056	0.027
3	0.054	0.014	0.053	0.014	0.049	0.027
4	0.020	0.014	0.019	0.014	0.028	0.027
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
Urban ≥ 10,000 – sparse	−0.050	0.108	−0.049	0.108	0.148	0.186
Town and fringe – sparse	−0.029	0.078	−0.024	0.078	−0.048	0.125
Village – sparse	−0.074	0.083	−0.074	0.083	−0.257	0.146
Hamlet and isolated dwelling – sparse	0.038	0.105	0.041	0.105	0.086	0.173
Urban ≥ 10,000 – less sparse	0.037	0.030	0.038	0.030	−0.018	0.054
Town and fringe – less sparse	0.006	0.033	0.006	0.033	−0.070	0.059
Village – less sparse	−0.018	0.035	−0.019	0.035	−0.067	0.061
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
North East	−0.007	0.093	−0.010	0.093	−0.298	0.114
North West	−0.090	0.071	−0.095	0.072	−0.283	0.090
Yorkshire and Humber	0.018	0.077	0.015	0.077	−0.154	0.098
East Midlands	−0.080	0.092	−0.079	0.092	−0.404	0.120
West Midlands	0.021	0.077	0.021	0.078	−0.145	0.105
East of England	0.119	0.079	0.120	0.079	0.014	0.110
London	0.094	0.079	0.093	0.079	−0.286	0.111
South East Coast	0.011	0.084	0.014	0.085	−0.364	0.137
South Central	0.084	0.086	0.084	0.086	−0.254	0.128
South West	0.000		0.000		0.000	
Trust size	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
ONS maternities (thousands)	0.005	0.011	0.006	0.011	0.074	0.032
University trust						
Yes	0.000		0.000		0.000	
No	−0.037	0.042	−0.037	0.042	0.139	0.091
Configuration	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
OU	−0.031	0.056	−0.029	0.056	0.000	
OU/AMU	−0.061	0.062	−0.056	0.062		
OU/AMU/FMU	−0.028	0.072	−0.026	0.073		
OU/FMU	0.000		0.000			

Staffing variables	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	$SE(\beta)$	β	$SE(\beta)$	β	$SE(\beta)$
FTE doctors per 100 maternities	−0.105	0.099	−0.096	0.100	−0.103	0.140
FTE midwives per 100 maternities	−0.023	0.044	−0.022	0.044	−0.055	0.086
FTE support workers per 100 maternities	−0.007	0.055	−0.007	0.055	0.071	0.102
Random variation (trust level)	σ	$SE(\sigma)$	σ	$SE(\sigma)$	σ	$SE(\sigma)$
Intercept only	0.203	0.007	0.203	0.007	0.174	0.016
Mother level	0.226	0.014	0.227	0.014	0.199	0.022
Mother level, sociodemographics	0.198	0.013	0.199	0.013	0.171	0.019
Mother level, sociodemographics, trust level	0.197	0.012	0.197	0.013	0.156	0.018
Mother level, sociodemographics, trust level, staff 1	0.195	0.012	0.196	0.012	0.153	0.017
Global tests (df)	χ^2	**p-value**	χ^2	**p-value**	χ^2	**p-value**
Mother's age group (6 df)	2921.832	< 0.0001	2875.943	< 0.0001	834.262	< 0.0001
Parity (4 df)	17,160.218	< 0.0001	17,175.321	< 0.0001	4795.734	< 0.0001
Clinical risk (1/2 df)[a]	18,387.718	< 0.0001	19,267.674	< 0.0001	4606.877	< 0.0001
Ethnicity (16 df)	954.095	< 0.0001	947.416	< 0.0001	208.822	< 0.0001
IMD (4 df)	70.290	< 0.0001	66.743	< 0.0001	8.677	0.0697
Rural/urban classification (7 df)	13.060	0.0706	13.274	0.0657	8.785	0.2685
SHA (9 df)	12.906	0.1669	13.157	0.1557	24.034	0.0042
ONS maternities (1 df)	0.217	0.6414	0.281	0.5958	5.420	0.0199
University trust (1 df)	0.767	0.3813	0.793	0.3731	2.329	0.1270
Configuration (3 df)	1.042	0.7911	0.897	0.8261		
FTE Doctors per 100 maternities (1 df)	1.118	0.2903	0.937	0.3331	0.536	0.4639
FTE Midwives per 100 maternities (1 df)	0.262	0.6085	0.248	0.6184	0.400	0.5273
FTE Support Workers per 100 maternities (1 df)	0.018	0.8941	0.017	0.8975	0.488	0.4849
AUC	**AUC**	**SE(AUC)**	**AUC**	**SE(AUC)**	**AUC**	**SE(AUC)**
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.692	0.001	0.696	0.001	0.697	0.002
Mother level, sociodemographics	0.698	0.001	0.701	0.001	0.703	0.002
Mother level, sociodemographics, trust level	0.698	0.001	0.701	0.001	0.703	0.002
Mother level, sociodemographics, trust level, staff 1	0.698	0.001	0.701	0.001	0.703	0.002

a   All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Caesarean: sensitivity analysis

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
	β	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Intercept	−1.595	0.192	−1.592	0.195	−2.046	0.364
Mother's age group (years)	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
≤ 19	−1.841	0.064	−1.823	0.065	−1.644	0.135
20–24	−1.486	0.062	−1.471	0.063	−1.260	0.132
25–29	−1.214	0.062	−1.198	0.062	−0.952	0.131
30–34	−0.973	0.062	−0.958	0.062	−0.725	0.131
35–39	−0.696	0.062	−0.683	0.062	−0.474	0.131
40–44	−0.445	0.063	−0.432	0.064	−0.257	0.134
≥ 45	0.000		0.000		0.000	
Parity	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
0	0.804	0.016	0.783	0.016	0.735	0.033
1	0.508	0.016	0.481	0.016	0.446	0.032
2	0.371	0.017	0.341	0.017	0.313	0.035
3	0.225	0.019	0.193	0.019	0.189	0.040
≥ 4	0.000		0.000		0.000	
Clinical risk	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Lower	0.000		0.000		0.000	
Individual assessment	N/A		0.753	0.019	N/A	
Higher	1.867	0.008	1.944	0.008	1.873	0.016
Ethnicity	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Not given/not known/not stated	−0.108	0.026	−0.114	0.026	−0.089	0.054
English/Welsh/Scottish/Northern Irish/British	−0.069	0.020	−0.074	0.020	−0.066	0.045
Irish	0.066	0.049	0.055	0.050	0.111	0.114
Any other white background	−0.101	0.022	−0.104	0.023	−0.093	0.049
White and black Caribbean	−0.145	0.054	−0.151	0.054	−0.309	0.133
White and black African	0.048	0.061	0.063	0.061	0.239	0.138
White and Asian	0.036	0.068	0.042	0.068	0.234	0.145
Any other mixed/multiple ethnic background	−0.004	0.048	−0.003	0.048	0.123	0.116
Indian	0.091	0.027	0.093	0.027	0.125	0.057
Pakistani	−0.153	0.026	−0.152	0.026	−0.175	0.060
Bangladeshi	−0.107	0.036	−0.118	0.036	0.015	0.088
Chinese	−0.191	0.045	−0.183	0.046	−0.266	0.101
Any other Asian background	0.077	0.030	0.076	0.030	0.192	0.066
African	0.235	0.026	0.228	0.026	0.322	0.057
Caribbean	0.023	0.037	0.016	0.037	−0.023	0.084

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Any other black/African/Caribbean background	0.108	0.038	0.101	0.038	0.096	0.082
Any other ethnic group, please describe	0.000		0.000		0.000	
IMD	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
1 = most deprived	−0.029	0.012	−0.036	0.012	−0.042	0.024
2	−0.008	0.012	−0.015	0.012	−0.013	0.023
3	−0.015	0.012	−0.019	0.012	−0.014	0.023
4	−0.001	0.012	−0.002	0.012	−0.011	0.023
5 = least deprived	0.000		0.000		0.000	
Rural/urban classification	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Urban ≥ 10,000 – sparse	−0.019	0.089	−0.018	0.089	0.227	0.160
Town and fringe – sparse	−0.027	0.064	−0.016	0.065	−0.033	0.106
Village – sparse	−0.129	0.068	−0.129	0.069	−0.161	0.118
Hamlet and isolated dwelling – sparse	0.059	0.087	0.069	0.088	0.149	0.147
Urban ≥ 10,000 – less sparse	−0.005	0.025	−0.006	0.025	−0.032	0.045
Town and fringe – less sparse	−0.031	0.028	−0.032	0.028	−0.077	0.050
Village – less sparse	−0.025	0.029	−0.028	0.029	−0.062	0.052
Hamlet and isolated dwelling – less sparse	0.000		0.000		0.000	
SHA	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
North East	−0.030	0.090	−0.047	0.092	−0.238	0.108
North West	−0.125	0.069	−0.145	0.071	−0.250	0.086
Yorkshire and Humber	−0.028	0.075	−0.041	0.076	−0.136	0.093
East Midlands	−0.134	0.090	−0.142	0.091	−0.339	0.114
West Midlands	−0.042	0.075	−0.049	0.077	−0.085	0.100
East of England	0.065	0.077	0.057	0.078	−0.032	0.105
London	0.112	0.077	0.104	0.078	−0.172	0.107
South East Coast	−0.004	0.082	−0.007	0.084	−0.249	0.131
South Central	0.003	0.083	−0.006	0.085	−0.239	0.122
South West	0.000		0.000		0.000	
Trust size	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
ONS maternities (thousands)	0.001	0.010	0.002	0.010	0.064	0.030
University trust	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
Yes	0.000		0.000		0.000	
No	−0.059	0.041	−0.062	0.042	0.101	0.087
Configuration	*β*	*SE(β)*	*β*	*SE(β)*	*β*	*SE(β)*
OU	−0.006	0.054	−0.001	0.055	0.000	
OU/AMU	−0.062	0.060	−0.054	0.061		
OU/AMU/FMU	−0.033	0.071	−0.029	0.072		
OU/FMU	0.000		0.000			

	All mothers		Clinical risk (three categories)		Trusts with single OUs	
Staffing variables	β	SE(β)	β	SE(β)	β	SE(β)
FTE doctors per 100 maternities	–0.154	0.097	–0.138	0.098	–0.209	0.134
FTE midwives per 100 maternities	0.000	0.043	–0.001	0.044	0.042	0.083
FTE support workers per 100 maternities	0.031	0.053	0.033	0.054	0.030	0.098
Random variation (trust level)	σ	SE(σ)	σ	SE(σ)	σ	SE(σ)
Intercept only	0.208	0.006	0.208	0.006	0.175	0.014
Mother level	0.215	0.013	0.219	0.013	0.188	0.020
Mother level, sociodemographics	0.196	0.012	0.199	0.012	0.163	0.018
Mother level, sociodemographics, trust level	0.194	0.012	0.197	0.012	0.153	0.017
Mother level, sociodemographics, trust level, staff 1	0.192	0.012	0.195	0.012	0.149	0.016
Global tests (df)	χ^2	p-value	χ^2	p-value	χ^2	p-value
Mother's age group (6 df)	8272.444	< 0.0001	8114.330	< 0.0001	2093.793	< 0.0001
Parity (4 df)	4244.311	< 0.0001	4187.323	< 0.0001	942.449	< 0.0001
Clinical risk (1/2 df)[a]	54,882.322	< 0.0001	60,219.451	< 0.0001	14,112.580	< 0.0001
Ethnicity (16 df)	504.113	< 0.0001	494.845	< 0.0001	168.344	< 0.0001
IMD (4 df)	8.436	0.0769	11.702	0.0197	3.829	0.4296
Rural/urban classification (7 df)	8.628	0.2805	9.039	0.2499	10.555	0.1592
SHA (9 df)	17.276	0.0446	18.248	0.0324	16.598	0.0554
ONS maternities (1 df)	0.010	0.9202	0.040	0.8413	4.530	0.0333
University trust (1 df)	2.078	0.1494	2.239	0.1346	1.337	0.2476
Configuration (3 df)	1.916	0.5900	1.588	0.6620		
FTE doctors per 100 maternities (1 df)	2.540	0.1110	1.973	0.1601	2.444	0.1180
FTE midwives per 100 maternities (1 df)	0.000	0.9962	0.000	0.9849	0.263	0.6081
FTE support workers per 100 maternities (1 df)	0.340	0.5600	0.378	0.5385	0.094	0.7597
AUC	AUC	SE(AUC)	AUC	SE(AUC)	AUC	SE(AUC)
Intercept only	0.500	0.000	0.500	0.000	0.500	0.000
Mother level	0.737	0.001	0.746	0.001	0.738	0.001
Mother level, sociodemographics	0.740	0.001	0.748	0.001	0.741	0.001
Mother level, sociodemographics, trust level	0.740	0.001	0.748	0.001	0.741	0.001
Mother level, sociodemographics, trust level, staff 1	0.740	0.001	0.748	0.001	0.741	0.001

a All mothers and trusts with single OUs 1 df; clinical risk (three categories) 2 df.
β, beta coefficient; df, degrees of freedom; σ, sigma coefficient; N/A, not applicable; SE(AUC), SE of area under the curve; SE(β), SE of beta coefficient; SE(σ), SE of sigma coefficient.

Appendix 6 Parity and clinical risk: tests of interaction with staffing variables

Parity by full-time equivalent doctors per 100 maternities: test of interaction

	Healthy mother				Healthy baby			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Parity								
0	0.062	0.054	0.071	< 0.0001	0.640	0.542	0.756	< 0.0001
1	0.237	0.209	0.270	< 0.0001	1.186	0.999	1.407	0.0509
2	0.470	0.409	0.540	< 0.0001	1.086	0.902	1.308	0.3828
3	0.688	0.586	0.807	< 0.0001	1.256	1.013	1.559	0.0379
≥ 4	1.000				1.000			
Staffing variables								
FTE doctors per 100 maternities	0.950	0.707	1.275	0.7307	1.071	0.781	1.468	0.6710
Parity × FTE doctors per 100 maternities								
0	1.195	1.021	1.399	0.0267	1.170	0.962	1.424	0.1151
1	1.099	0.944	1.281	0.2249	1.019	0.833	1.246	0.8579
2	1.034	0.879	1.217	0.6847	1.091	0.876	1.358	0.4366
3	1.013	0.840	1.223	0.8904	0.836	0.649	1.077	0.1668
≥ 4	1.000				1.000			
Chi-squared test	χ^2			*p-value*	χ^2			*p-value*
Parity × FTE doctors per 100 maternities (4 df)	9.957			0.0412	15.115			0.0045

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicator, the model was unable to be fitted because of the inability to find initial covariance matrix estimates.

Healthy mother/healthy baby dyad				Delivery with bodily integrity				Normal birth			
OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
0.063	0.055	0.073	< 0.0001	0.070	0.062	0.078	< 0.0001	0.239	0.207	0.276	< 0.0001
0.259	0.226	0.296	< 0.0001	0.231	0.206	0.259	< 0.0001	0.739	0.641	0.852	< 0.0001
0.491	0.426	0.566	< 0.0001	0.441	0.391	0.499	< 0.0001	0.917	0.786	1.069	0.2659
0.742	0.630	0.874	0.0004	0.648	0.563	0.747	< 0.0001	0.986	0.825	1.179	0.8812
1.000				1.000				1.000			
0.985	0.734	1.322	0.9192	1.069	0.828	1.380	0.6095	1.301	0.993	1.705	0.0562
1.220	1.038	1.435	0.0162	1.058	0.920	1.218	0.4288	0.679	0.574	0.803	< 0.0001
1.095	0.937	1.281	0.2536	0.977	0.850	1.122	0.7376	0.692	0.585	0.819	< 0.0001
1.052	0.891	1.243	0.5494	0.992	0.856	1.150	0.9166	0.727	0.607	0.872	0.0006
0.975	0.804	1.182	0.7969	1.019	0.859	1.209	0.8251	0.800	0.648	0.987	0.0371
1.000				1.000				1.000			
χ^2			**p-value**	χ^2			**p-value**	χ^2			**p-value**
12.502			0.0140	4.532			0.0140	23.572			< 0.0001

	Spontaneous vaginal delivery				Intact perineum			
	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
Parity								
0	0.247	0.218	0.279	< 0.0001	0.050	0.043	0.059	< 0.0001
1	0.613	0.540	0.695	< 0.0001	0.167	0.143	0.194	< 0.0001
2	0.760	0.663	0.870	< 0.0001	0.351	0.300	0.412	< 0.0001
3	0.872	0.744	1.022	0.0913	0.554	0.461	0.665	< 0.0001
≥ 4	1.000				1.000			
Staffing variables								
FTE doctors per 100 maternities	1.270	1.014	1.590	0.0371	0.992	0.731	1.347	0.9594
Parity × FTE doctors per 100 maternities								
0	0.771	0.664	0.895	0.0006	1.098	0.916	1.317	0.3112
1	0.801	0.688	0.932	0.0041	1.025	0.856	1.228	0.7899
2	0.856	0.727	1.009	0.0639	1.049	0.867	1.270	0.6204
3	0.897	0.741	1.086	0.2638	1.087	0.871	1.356	0.4598
≥ 4	1.000				1.000			
Chi-squared test	χ^2			*p-value*	χ^2			*p-value*
Parity × FTE doctors per 100 maternities (4 df)	16.508			0.0024	3.410			0.4916

$OR_{L95\%}$, lower 95% confidence limit of the odds ratio; $OR_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicators, the model was unable to be fitted because of the mobility to find initial covariance matrix estimates.

Elective caesarean				Emergency caesarean				Caesarean			
OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
0.464	0.389	0.553	< 0.0001	3.780	3.190	4.479	< 0.0001	1.952	1.712	2.227	< 0.0001
1.309	1.108	1.546	0.0015	1.403	1.177	1.671	0.0002	1.405	1.230	1.605	< 0.0001
1.608	1.346	1.920	< 0.0001	0.960	0.792	1.163	0.6769	1.324	1.147	1.528	0.0001
1.354	1.101	1.664	0.0040	0.978	0.781	1.224	0.8442	1.176	0.995	1.390	0.0573
1.000				1.000				1.000			
0.777	0.585	1.032	0.0813	0.842	0.641	1.107	0.2187	0.732	0.576	0.929	0.0104
1.285	1.039	1.590	0.0206	1.034	0.843	1.267	0.7505	1.181	1.008	1.384	0.0389
1.156	0.946	1.413	0.1566	1.117	0.905	1.378	0.3028	1.229	1.047	1.443	0.0115
0.964	0.778	1.195	0.7401	1.224	0.972	1.542	0.0863	1.118	0.940	1.329	0.2071
0.976	0.760	1.252	0.8463	1.123	0.857	1.470	0.4002	1.081	0.883	1.322	0.4507
1.000				1.000				1.000			
χ^2			**p-*value*	χ^2			**p-*value*	χ^2			**p-*value*
18.256			0.0011	7.684			0.1039	9.029			0.0604

Parity by full-time equivalent midwives per 100 maternities: test of interaction

	Delivery with bodily integrity				Spontaneous vaginal delivery				Intact perineum			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Parity												
0	0.066	0.055	0.079	< 0.0001	0.195	0.161	0.237	< 0.0001	0.075	0.059	0.096	< 0.0001
1	0.220	0.184	0.263	< 0.0001	0.502	0.413	0.611	< 0.0001	0.239	0.188	0.303	< 0.0001
2	0.394	0.325	0.477	< 0.0001	0.634	0.513	0.783	< 0.0001	0.431	0.335	0.554	< 0.0001
3	0.656	0.526	0.819	0.0002	0.765	0.598	0.979	0.0333	0.753	0.563	1.008	0.0562
≥ 4	1.000				1.000				1.000			
Staffing variables												
FTE midwives per 100 maternities	1.082	0.967	1.212	0.1687	1.015	0.920	1.120	0.7645	1.248	1.090	1.428	0.0014
Parity × FTE midwives per 100 maternities												
0	1.036	0.976	1.100	0.2433	1.007	0.945	1.074	0.8180	0.896	0.828	0.970	0.0068
1	1.011	0.952	1.072	0.7268	1.006	0.944	1.073	0.8457	0.893	0.825	0.967	0.0051
2	1.037	0.973	1.104	0.2637	1.018	0.950	1.092	0.6074	0.946	0.871	1.029	0.1971
3	1.002	0.931	1.078	0.9639	1.014	0.935	1.100	0.7344	0.923	0.838	1.016	0.1030
≥ 4	1.000				1.000				1.000			
Chi-squared test	χ^2			**p-value**	χ^2			**p-value**	χ^2			**p-value**
Parity × FTE midwives per 100 maternities (4 df)	4.595			0.3314	0.466			0.9767	14.150			0.0068

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For healthy mother, healthy baby, healthy mother/healthy baby dyad and normal birth indicators, this model was unable to be fitted because of the inability to find initial covariance matrix estimates.

Elective caesarean				Emergency caesarean				Caesarean			
OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
0.418	0.318	0.549	< 0.0001	5.035	3.844	6.597	< 0.0001	2.537	2.066	3.117	< 0.0001
1.324	1.022	1.714	0.0335	1.977	1.497	2.610	< 0.0001	1.800	1.462	2.216	< 0.0001
1.603	1.218	2.111	0.0008	1.306	0.963	1.770	0.0858	1.623	1.297	2.030	< 0.0001
1.373	0.996	1.892	0.0526	1.367	0.958	1.950	0.0847	1.480	1.140	1.922	0.0033
1.000				1.000				1.000			
0.992	0.876	1.123	0.8951	1.056	0.936	1.192	0.3726	1.042	0.938	1.157	0.4475
1.107	1.012	1.211	0.0265	0.917	0.839	1.003	0.0569	0.959	0.896	1.026	0.2238
1.036	0.951	1.128	0.4197	0.919	0.839	1.007	0.0711	0.974	0.909	1.043	0.4480
0.991	0.905	1.085	0.8508	0.953	0.862	1.054	0.3492	0.963	0.895	1.037	0.3209
0.989	0.889	1.099	0.8323	0.923	0.821	1.038	0.1804	0.946	0.868	1.031	0.2089
1.000				1.000				1.000			
χ^2			p-*value*	χ^2			p-*value*	χ^2			p-*value*
16.219			0.0027	5.067			0.2804	2.550			0.6356

Parity by full-time equivalent support workers per 100 maternities: test of interaction

	Healthy mother				Healthy baby			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Parity								
0	0.071	0.065	0.077	< 0.0001	0.744	0.666	0.833	< 0.0001
1	0.251	0.230	0.274	< 0.0001	1.193	1.064	1.338	0.0026
2	0.454	0.414	0.498	< 0.0001	1.265	1.117	1.434	0.0002
3	0.672	0.604	0.749	< 0.0001	1.188	1.028	1.374	0.0200
≤ 4	1.000				1.000			
Staffing variables								
FTE support workers per 100 maternities	0.866	0.738	1.016	0.0780	0.991	0.834	1.177	0.9145
Parity × FTE support workers per 100 maternities								
0	1.019	0.930	1.116	0.6912	0.977	0.872	1.095	0.6876
1	1.025	0.938	1.119	0.5906	1.011	0.899	1.136	0.8605
2	1.071	0.975	1.178	0.1539	0.916	0.806	1.040	0.1742
3	1.039	0.931	1.159	0.4958	0.906	0.782	1.050	0.1899
≤ 4	1.000				1.000			
Chi-squared	χ^2			*p-value*	χ^2			*p-value*
Parity × FTE support workers per 100 maternities (4 df)	3.357			0.4999	7.226			0.1244

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicator, the model was unable to be fitted because of the inability to find initial covariance matrix estimates.

Healthy mother/healthy baby dyad				Delivery with bodily integrity				Normal birth			
OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
0.072	0.066	0.079	< 0.0001	0.073	0.067	0.079	< 0.0001	0.179	0.164	0.196	< 0.0001
0.269	0.246	0.294	< 0.0001	0.232	0.215	0.251	< 0.0001	0.560	0.514	0.610	< 0.0001
0.494	0.450	0.544	< 0.0001	0.431	0.397	0.468	< 0.0001	0.750	0.683	0.823	< 0.0001
0.704	0.631	0.785	< 0.0001	0.640	0.581	0.705	< 0.0001	0.854	0.766	0.951	0.0042
1.000				1.000				1.000			
0.867	0.739	1.017	0.0795	1.002	0.868	1.156	0.9820	1.049	0.910	1.211	0.5083
1.041	0.948	1.143	0.3991	1.007	0.927	1.093	0.8756	0.964	0.881	1.054	0.4181
1.040	0.950	1.138	0.3972	0.976	0.900	1.059	0.5599	0.969	0.886	1.061	0.4995
1.039	0.944	1.145	0.4342	1.020	0.935	1.113	0.6573	0.932	0.845	1.027	0.1564
1.037	0.927	1.159	0.5277	1.033	0.934	1.143	0.5280	0.954	0.852	1.069	0.4204
1.000				1.000				1.000			
χ^2			**p-*value*	χ^2			**p-*value*	χ^2			**p-*value*
0.800			0.9384	4.334			0.3627	2.607			0.6256

	Spontaneous vaginal delivery				Intact perineum			
	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
Parity								
0	0.204	0.187	0.222	< 0.0001	0.061	0.055	0.068	< 0.0001
1	0.513	0.471	0.559	< 0.0001	0.194	0.174	0.215	< 0.0001
2	0.657	0.600	0.720	< 0.0001	0.389	0.348	0.434	< 0.0001
3	0.817	0.734	0.909	0.0002	0.587	0.517	0.667	< 0.0001
≤ 4	1.000				1.000			
Staffing variables								
FTE support workers per 100 maternities	0.965	0.850	1.095	0.5789	1.147	0.964	1.366	0.1226
Parity × FTE support workers per 100 maternities								
0	0.978	0.896	1.067	0.6157	0.872	0.780	0.975	0.0162
1	0.997	0.912	1.089	0.9391	0.865	0.774	0.967	0.0105
2	1.020	0.927	1.123	0.6828	0.932	0.829	1.048	0.2415
3	0.974	0.871	1.090	0.6455	1.009	0.881	1.156	0.8941
≤ 4	1.000				1.000			
Chi-squared	χ^2			**p-value**	χ^2			**p-value**
Parity × FTE support workers per 100 maternities (4 df)	2.690			0.6110	19.483			0.0006

$OR_{L95\%}$, lower 95% confidence limit of the odds ratio; $OR_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicators, the model was unable to be fitted because of the mobility to find initial covariance matrix estimates.

Elective caesarean				Emergency caesarean				Caesarean			
OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
0.505	0.448	0.570	< 0.0001	4.386	3.906	4.924	< 0.0001	2.301	2.104	2.516	< 0.0001
1.275	1.138	1.428	< 0.0001	1.776	1.576	2.001	< 0.0001	1.663	1.520	1.820	< 0.0001
1.430	1.267	1.614	< 0.0001	1.257	1.102	1.433	0.0007	1.458	1.323	1.607	< 0.0001
1.208	1.049	1.391	0.0085	1.158	0.993	1.350	0.0613	1.236	1.104	1.385	0.0003
1.000				1.000				1.000			
0.950	0.809	1.117	0.5344	1.132	0.969	1.323	0.1185	1.049	0.917	1.200	0.4839
1.139	1.004	1.291	0.0426	0.873	0.774	0.985	0.0270	0.968	0.882	1.062	0.4882
1.172	1.041	1.320	0.0088	0.849	0.750	0.961	0.0099	0.998	0.909	1.097	0.9714
1.103	0.971	1.252	0.1311	0.889	0.775	1.019	0.0914	0.993	0.897	1.099	0.8889
1.110	0.957	1.287	0.1665	0.919	0.783	1.079	0.3042	1.014	0.901	1.142	0.8142
1.000				1.000				1.000			
χ^2			*p-value*	χ^2			*p-value*	χ^2			*p-value*
8.113			0.0875	7.650			0.1053	2.569			0.6323

Clinical risk by full-time equivalent doctors per 100 maternities: test of interaction

	Healthy mother				Healthy baby			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Clinical risk								
Low	1.000				1.000			
High	0.416	0.390	0.444	< 0.0001	0.137	0.124	0.152	< 0.0001
Staffing variables								
FTE doctors per 100 maternities	1.136	0.871	1.483	0.3472	1.157	0.874	1.531	0.3075
Clinical risk × FTE doctors per 100 maternities								
Low	1.000				1.000			
High	0.877	0.812	0.947	0.0008	0.998	0.883	1.128	0.9772
Chi-squared test	χ^2			**p-value**	χ^2			**p-value**
Clinical risk × FTE doctors per 100 maternities (1 df)	11.346			0.0008	0.001			0.9772

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicator, the model was unable to be fitted because of the inability to find initial covariance matrix estimates.

Healthy mother/healthy baby dyad				Delivery with bodily integrity				Normal birth			
OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
1.000				1.000				1.000			
0.293	0.274	0.314	< 0.0001	0.428	0.406	0.451	< 0.0001	0.147	0.138	0.156	< 0.0001
1.129	0.867	1.471	0.3675	1.097	0.875	1.376	0.4225	0.879	0.702	1.100	0.2600
1.000				1.000				1.000			
0.954	0.880	1.033	0.2479	1.005	0.943	1.071	0.8766	1.127	1.047	1.213	0.0014
χ^2			**p-*value***	χ^2			**p-*value***	χ^2			**p-*value***
1.335			0.2479	0.024			0.8766	10.164			0.0014

	Spontaneous vaginal delivery				Intact perineum			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Clinical risk								
Low	1.000				1.000			
High	0.219	0.208	0.231	< 0.0001	0.819	0.772	0.868	< 0.0001
Staffing variables								
FTE doctors per 100 maternities	0.935	0.780	1.120	0.4643	1.077	0.832	1.393	0.5744
Clinical risk × FTE doctors per 100 maternities								
Low	1.000				1.000			
High	1.134	1.064	1.209	0.0001	0.985	0.918	1.057	0.6749
Chi-squared test	χ^2			**p-value**	χ^2			**p-value**
Clinical risk × FTE doctors per 100 maternities (1 df)	14.926			0.0001	0.176			0.6749

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicators, the model was unable to be fitted because of the mobility to find initial covariance matrix estimates.

Elective caesarean				Emergency caesarean				Caesarean			
OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
1.000				1.000				1.000			
35.116	29.294	42.095	< 0.0001	3.372	3.131	3.632	< 0.0001	7.322	6.843	7.834	< 0.0001
1.432	1.062	1.932	0.0185	0.926	0.755	1.137	0.4642	0.964	0.790	1.177	0.7189
1.000				1.000				1.000			
0.584	0.471	0.723	< 0.0001	0.960	0.878	1.050	0.3753	0.858	0.791	0.930	0.0002
χ^2			p-*value*	χ^2			p-*value*	χ^2			p-*value*
24.234			< 0.0001	0.786			0.3753	13.716			0.0002

Clinical risk by full-time equivalent midwives per 100 maternities: test of interaction

	Healthy mother				Healthy baby			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Clinical risk								
Low	1.000				1.000			
High	0.441	0.399	0.487	< 0.0001	0.169	0.144	0.199	< 0.0001
Staffing variables								
FTE midwives per 100 maternities	1.117	0.987	1.263	0.0794	1.089	0.957	1.238	0.1950
Clinical risk × FTE midwives per 100 maternities								
Low	1.000				1.000			
High	0.947	0.917	0.978	0.0009	0.933	0.886	0.983	0.0089
Chi-squared test	χ^2			**p-value**	χ^2			**p-value**
Clinical risk × FTE midwives per 100 maternities (1 df)	11.002			0.0009	6.836			0.0089

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicator, the model was unable to be fitted because of the inability to find initial covariance matrix estimates.

Healthy mother/healthy baby dyad				Delivery with bodily integrity				Spontaneous vaginal delivery			
OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
1.000				1.000				1.000			
0.325	0.293	0.361	< 0.0001	0.455	0.421	0.493	< 0.0001	0.242	0.224	0.262	< 0.0001
1.115	0.986	1.259	0.0819	1.119	1.012	1.237	0.0282	1.026	0.947	1.112	0.5255
1.000				1.000				1.000			
0.954	0.923	0.987	0.0065	0.981	0.956	1.007	0.1461	1.000	0.975	1.026	0.9818
χ^2			p-*value*	χ^2			p-*value*	χ^2			p-*value*
7.413			0.0065	2.112			0.1461	0.001			0.9818

	Intact perineum				Elective caesarean			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Clinical risk								
Low	1.000				1.000			
High	0.819	0.751	0.894	< 0.0001	20.816	15.758	27.496	< 0.0001
Staffing variables								
FTE midwives per 100 maternities	1.136	1.013	1.273	0.0292	1.008	0.885	1.149	0.9005
Clinical risk × FTE midwives per 100 maternities								
Low	1.000				1.000			
High	0.996	0.968	1.025	0.7709	1.029	0.939	1.127	0.5456
Chi-squared test	χ^2			**p-value**	χ^2			**p-value**
Clinical risk × FTE midwives per 100 maternities (1 df)	0.085			0.7709	0.365			0.5456

Note the intact perineum model could not be estimated.
OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.

| Emergency caesarean | | | | Caesarean | | | |
OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
1.000				1.000			
2.951	2.641	3.297	< 0.0001	5.775	5.217	6.393	< 0.0001
0.948	0.866	1.037	0.2429	0.961	0.880	1.049	0.3746
1.000				1.000			
1.034	0.997	1.072	0.0724	1.038	1.004	1.074	0.0273
χ^2			**p-*value***	χ^2			**p-*value***
3.229			0.0724	4.872			0.0273

Clinical risk by full-time equivalent support workers per 100 maternities: test of interaction

	Healthy mother				Healthy baby			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Clinical risk								
Low	1.000				1.000			
High	0.401	0.385	0.418	< 0.0001	0.148	0.139	0.159	< 0.0001
Staffing variables								
FTE support workers per 100 maternities	0.924	0.802	1.065	0.2748	1.035	0.890	1.204	0.6511
Clinical risk × FTE support workers per 100 maternities								
Low	1.000				1.000			
High	0.925	0.886	0.966	0.0004	0.917	0.854	0.986	0.0184
Chi-squared test	χ^2			***p-value***	χ^2			***p-value***
Clinical risk × FTE support workers per 100 maternities (1 df)	12.605			0.0004	5.555			0.0184

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicator, the model was unable to be fitted because of the inability to find initial covariance matrix estimates.

Healthy mother/healthy baby dyad				Delivery with bodily integrity				Normal birth			
OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
1.000				1.000				1.000			
0.304	0.290	0.317	< 0.0001	0.462	0.446	0.478	< 0.0001	0.176	0.170	0.183	< 0.0001
0.931	0.809	1.072	0.3201	1.039	0.918	1.177	0.5449	1.059	0.940	1.192	0.3463
1.000				1.000				1.000			
0.921	0.880	0.964	0.0004	0.924	0.891	0.958	< 0.0001	0.908	0.873	0.945	< 0.0001
χ^2			**p-*value***	χ^2			**p-*value***	χ^2			**p-*value***
12.681			0.0004	18.242			< 0.0001	22.991			< 0.0001

	Spontaneous vaginal delivery				Intact perineum			
	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value	OR	OR$_{L95\%}$	OR$_{U95\%}$	*p*-value
Clinical risk								
Low	1.000				1.000			
High	0.251	0.242	0.260	< 0.0001	0.845	0.813	0.879	< 0.0001
Staffing variables								
FTE support workers per 100 maternities	0.983	0.890	1.086	0.7348	1.038	0.901	1.197	0.6034
Clinical risk × FTE support workers per 100 maternities								
Low	1.000				1.000			
High	0.963	0.928	0.999	0.0440	0.952	0.915	0.991	0.0165
Chi-squared test	χ^2			*p-value*	χ^2			*p-value*
Clinical risk × FTE support workers per 100 maternities (1 df)	4.058			0.0440	5.747			0.0165

OR$_{L95\%}$, lower 95% confidence limit of the odds ratio; OR$_{U95\%}$, upper 95% confidence limit of the odds ratio.
For the normal birth indicators, the model was unable to be fitted because of the mobility to find initial covariance matrix estimates.

Elective caesarean				Emergency caesarean				Caesarean			
OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value	OR	$OR_{L95\%}$	$OR_{U95\%}$	*p*-value
1.000				1.000				1.000			
20.735	18.346	23.435	< 0.0001	3.101	2.949	3.260	< 0.0001	6.018	5.750	6.299	< 0.0001
0.980	0.825	1.163	0.8140	0.954	0.851	1.068	0.4121	0.975	0.874	1.089	0.6588
1.000				1.000				1.000			
1.104	0.971	1.256	0.1323	1.059	1.005	1.116	0.0332	1.084	1.033	1.137	0.0010
χ^2			**p-value**	χ^2			**p-value**	χ^2			**p-value**
2.265			0.1323	4.535			0.0332	10.788			0.0010

Appendix 7 Completeness of data by trust

Percentage of women for whom the indicator was known by trust

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
5QT Isle of Wight NHS PCT	10	97.6	97.6	97.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RA4 Yeovil District Hospital NHS Foundation Trust	10	87.7	87.7	87.7	100.0	98.6	100.0	100.0	100.0	100.0	100.0
RA7 University Hospitals Bristol NHS Foundation Trust/Weston Area Health	10	90.3	90.3	90.3	100.0	98.7	100.0	100.0	100.0	100.0	100.0
RA9 South Devon Healthcare NHS Foundation Trust	10	98.4	98.4	98.4	100.0	99.8	100.0	100.0	100.0	100.0	100.0
RAJ Southend University Hospital NHS Foundation Trust	10	99.4	99.4	99.4	100.0	96.5	100.0	100.0	100.0	100.0	100.0
RAL Royal Free Hampstead NHS Trust	10	97.5	97.5	97.5	100.0	99.7	100.0	100.0	100.0	100.0	100.0
RAP North Middlesex University Hospital NHS Trust	10	93.1	93.1	93.1	100.0	91.1	100.0	100.0	100.0	100.0	100.0
RAS The Hillingdon Hospital NHS Trust	10	92.7	92.7	92.7	100.0	95.1	100.0	100.0	100.0	100.0	100.0
RAX Kingston Hospital NHS Trust	10	98.9	98.9	98.9	100.0	99.3	100.0	100.0	100.0	100.0	100.0
RBD Dorset County Hospital NHS Foundation Trust	10	98.7	98.7	98.7	100.0	98.9	100.0	100.0	100.0	100.0	100.0
RBK Walsall Hospitals NHS Trust	10	98.5	98.5	98.5	100.0	92.9	100.0	100.0	100.0	100.0	100.0
RBL Wirral University Teaching Hospital NHS Foundation Trust	10	96.7	96.7	96.7	100.0	97.3	100.0	100.0	100.0	100.0	100.0
RBN St Helens and Knowsley Hospitals NHS Trust	10	98.8	98.8	98.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RC1 Bedford Hospital NHS Trust	10	99.8	99.8	99.8	100.0	99.9	100.0	100.0	100.0	100.0	100.0
RCB York Hospitals NHS Foundation Trust	10	94.9	94.9	94.9	100.0	85.9	100.0	100.0	100.0	100.0	100.0
RCC Scarborough and North East Yorkshire Health Care NHS Trust	10	98.8	98.8	98.8	100.0	99.0	100.0	100.0	100.0	100.0	100.0
RCD Harrogate and District NHS Foundation Trust	10	99.2	99.2	99.2	100.0	99.8	100.0	100.0	100.0	100.0	100.0
RCF Airedale NHS Trust	10	95.3	95.3	95.3	100.0	95.6	100.0	100.0	100.0	100.0	100.0
RCX The Queen Elizabeth Hospital King's Lynn NHS Trust	10	98.7	98.7	98.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RD3 Poole Hospital NHS Foundation Trust/Bournemouth NHS Trust	10	94.1	94.1	94.1	100.0	97.9	100.0	100.0	100.0	100.0	100.0
RD8 Milton Keynes Hospital NHS Foundation Trust	10	98.4	98.4	98.4	99.9	96.0	99.9	99.9	99.9	99.9	99.9
RDE Colchester Hospital University NHS Foundation Trust	10	99.0	99.0	99.0	100.0	98.1	100.0	99.9	100.0	100.0	100.0
RDU Frimley Park Hospital NHS Foundation Trust	10	97.4	97.4	97.4	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RE9 South Tyneside NHS Foundation Trust	10	99.3	99.3	99.3	100.0	99.9	100.0	100.0	100.0	100.0	100.0
REF Royal Cornwall Hospitals NHS Trust	10	92.9	92.9	92.9	99.6	98.1	99.6	99.5	99.6	99.6	99.6
RF4 Barking, Havering and Redbridge University Hospitals NHS Trust	10	97.2	97.2	97.2	100.0	94.1	100.0	100.0	100.0	100.0	100.0

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RFF Barnsley Hospital NHS Foundation Trust	10	93.1	93.1	93.1	100.0	95.9	100.0	100.0	100.0	100.0	100.0
RFR The Rotherham NHS Foundation Trust	10	98.9	98.9	98.9	100.0	98.2	100.0	100.0	100.0	100.0	100.0
RFS Chesterfield Royal Hospital NHS Foundation Trust	10	98.0	98.0	98.0	100.0	97.4	100.0	100.0	100.0	100.0	100.0
RFW West Middlesex University Hospital NHS Trust	10	88.6	88.6	88.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RGC Whipps Cross University Hospital NHS Trust	10	98.6	98.6	98.6	100.0	95.2	100.0	100.0	100.0	100.0	100.0
RGN Peterborough and Stamford Hospitals NHS Foundation Trust	10	98.1	98.1	98.1	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RGP James Paget University Hospitals NHS Foundation Trust	10	98.1	98.1	98.1	100.0	98.6	100.0	100.0	100.0	100.0	100.0
RGR West Suffolk Hospitals NHS Trust	10	98.3	98.3	98.3	100.0	99.6	100.0	100.0	100.0	100.0	100.0
RGT Cambridge University Hospitals NHS Foundation Trust	10	99.7	99.7	99.7	100.0	99.9	100.0	100.0	100.0	100.0	100.0
RH8 Royal Devon and Exeter NHS Foundation Trust/Devon PCT	10	88.8	88.8	88.8	100.0	99.0	100.0	100.0	100.0	100.0	100.0
RHQ Sheffield Teaching Hospitals NHS Foundation Trust	10	99.2	99.2	99.2	99.9	95.9	99.9	99.9	99.9	99.9	99.9
RHU Portsmouth Hospitals NHS Trust	10	83.8	83.8	83.8	100.0	99.7	100.0	100.0	100.0	100.0	100.0
RHW Royal Berkshire NHS Foundation Trust	10	97.1	97.1	97.1	100.0	97.5	100.0	100.0	100.0	100.0	100.0

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RJ1 Guy's and St Thomas' NHS Foundation Trust	10	99.3	99.3	99.3	100.0	99.2	100.0	100.0	100.0	100.0	100.0
RJE University Hospital of North Staffordshire NHS Trust	10	99.5	99.5	99.5	100.0	99.7	100.0	100.0	100.0	100.0	100.0
RJL Northern Lincolnshire and Goole Hospitals NHS Foundation Trust	10	87.2	87.2	87.2	100.0	98.7	100.0	100.0	100.0	100.0	100.0
RJN East Cheshire NHS Trust	10	84.3	84.3	84.3	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RJR Countess of Chester Hospital NHS Foundation Trust	10	97.4	97.4	97.4	100.0	97.8	100.0	100.0	100.0	100.0	100.0
RK5 Sherwood Forest Hospitals NHS Foundation Trust	10	95.7	95.7	95.7	100.0	96.1	100.0	100.0	100.0	100.0	100.0
RK9 Plymouth Hospitals NHS Trust	10	99.5	99.5	99.5	100.0	98.0	100.0	100.0	100.0	100.0	100.0
RKB University Hospitals Coventry and Warwickshire NHS Trust	10	99.5	99.5	99.5	100.0	99.8	100.0	100.0	100.0	100.0	100.0
RKE The Whittington Hospital NHS Trust	10	98.4	98.4	98.4	100.0	96.4	100.0	100.0	100.0	100.0	100.0
RL4 The Royal Wolverhampton Hospitals NHS Trust	10	95.2	95.2	95.2	100.0	91.6	100.0	99.9	100.0	100.0	100.0
RLN City Hospitals Sunderland NHS Foundation Trust	10	95.8	95.8	95.8	100.0	97.5	100.0	100.0	100.0	100.0	100.0
RLQ Hereford Hospitals NHS Trust	10	95.9	95.9	95.9	100.0	97.8	100.0	100.0	100.0	100.0	100.0
RLT George Eliot Hospital NHS Trust	10	99.5	99.5	99.5	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RM1 Norfolk and Norwich University Hospitals NHS Foundation Trust	10	98.7	98.7	98.7	100.0	99.1	100.0	100.0	100.0	100.0	100.0
RM3 Salford Royal NHS Foundation Trust	10	93.1	93.1	93.1	100.0	97.1	100.0	100.0	100.0	100.0	100.0
RMC Royal Bolton Hospital NHS Foundation Trust	10	90.1	90.1	90.1	100.0	93.5	100.0	100.0	100.0	100.0	100.0
RMP Tameside Hospital NHS Foundation Trust	10	96.7	96.7	96.7	99.3	98.3	99.3	99.1	99.3	99.3	99.3
RN1 Winchester and Eastleigh Healthcare Trust	10	83.2	83.2	83.2	99.3	83.7	99.3	99.0	99.3	99.3	99.3
RN3 Great Western Hospitals NHS Foundation Trust	10	97.7	97.7	97.7	100.0	99.9	100.0	100.0	100.0	100.0	100.0
RNA The Dudley Group of Hospitals NHS Foundation Trust	10	99.3	99.3	99.3	100.0	99.7	100.0	100.0	100.0	100.0	100.0
RNH Newham University Hospital NHS Trust	10	96.7	96.7	96.7	100.0	96.0	100.0	100.0	100.0	100.0	100.0
RNJ Barts and The London NHS Trust	10	96.2	96.2	96.2	100.0	97.0	100.0	100.0	100.0	100.0	100.0
RNL North Cumbria University Hospitals NHS Trust	10	99.0	99.0	99.0	100.0	98.9	100.0	100.0	100.0	100.0	100.0
RNQ Kettering General Hospital NHS Foundation Trust	10	99.3	99.3	99.3	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RNS Northampton General Hospital NHS Trust	10	97.8	97.8	97.8	100.0	96.2	100.0	100.0	100.0	100.0	100.0
RNZ Salisbury NHS Foundation Trust	10	99.0	99.0	99.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RP5 Doncaster and Bassetlaw Hospitals NHS Foundation Trust	10	97.4	97.4	97.4	100.0	95.7	100.0	100.0	100.0	100.0	100.0
RQW The Princess Alexandra Hospital NHS Trust	10	97.5	97.5	97.5	100.0	96.0	100.0	100.0	100.0	100.0	100.0
RQX Homerton University Hospital NHS Foundation Trust	10	98.6	98.6	98.6	100.0	97.6	100.0	100.0	100.0	100.0	100.0
RR7 Gateshead Health NHS Foundation Trust	10	97.5	97.5	97.5	100.0	98.2	100.0	100.0	100.0	100.0	100.0
RR8 Leeds Teaching Hospitals NHS Trust	10	98.2	98.2	98.2	99.9	85.6	99.9	99.9	99.9	99.9	99.9
RRF Wrightington, Wigan and Leigh NHS Foundation Trust	10	96.4	96.4	96.4	100.0	99.5	100.0	100.0	100.0	100.0	100.0
RTD The Newcastle upon Tyne Hospitals NHS Foundation Trust	10	95.0	95.0	95.0	100.0	99.6	100.0	100.0	100.0	100.0	100.0
RTF Northumbria Healthcare NHS Foundation Trust	10	98.0	98.0	98.0	100.0	97.7	100.0	100.0	100.0	100.0	100.0
RTG Derby Hospitals NHS Foundation Trust	10	96.9	96.9	96.9	99.9	97.4	99.9	99.9	99.9	99.9	99.9
RTH Oxford Radcliffe Hospitals NHS Trust	10	98.3	98.3	98.3	100.0	99.9	100.0	100.0	100.0	100.0	100.0
RTK Ashford and St Peter's Hospitals NHS Trust	10	99.7	99.7	99.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RTX University Hospitals of Morecambe Bay NHS Trust	10	95.2	95.2	95.2	99.9	87.6	99.9	99.9	99.9	99.9	99.9
RVJ North Bristol NHS Trust	10	92.3	92.3	92.3	99.9	92.9	99.9	99.9	99.9	99.9	99.9

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RVL Barnet and Chase Farm Hospitals NHS Trust	10	95.8	95.8	95.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RVR Epsom and St Helier University Hospitals NHS Trust	10	99.3	99.3	99.3	100.0	99.9	100.0	100.0	100.0	100.0	100.0
RVW North Tees and Hartlepool NHS Foundation Trust	10	99.5	99.5	99.5	100.0	99.7	100.0	100.0	100.0	100.0	100.0
RVY Southport and Ormskirk Hospital NHS Trust	10	98.7	98.7	98.7	100.0	96.4	100.0	100.0	100.0	100.0	100.0
RW3 Central Manchester University Hospitals NHS Foundation Trust	10	96.8	96.8	96.8	100.0	99.0	100.0	100.0	100.0	100.0	100.0
RW6 Pennine Acute Hospitals NHS Trust	10	95.7	95.7	95.7	100.0	99.9	100.0	99.9	100.0	100.0	100.0
RWA Hull and East Yorkshire Hospitals NHS Trust	10	98.8	98.8	98.8	100.0	96.0	100.0	99.9	100.0	100.0	100.0
RWD United Lincolnshire Hospitals NHS Trust	10	96.0	96.0	96.0	100.0	91.2	100.0	100.0	100.0	100.0	100.0
RWE University Hospitals of Leicester NHS Trust	10	96.5	96.5	96.5	99.5	94.9	99.5	99.4	99.5	99.5	99.5
RWG West Hertfordshire Hospitals NHS Trust	10	95.9	95.9	95.9	99.8	99.2	99.8	99.8	99.8	99.8	99.8
RWJ Stockport NHS Foundation Trust	10	97.7	97.7	97.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RWW Warrington and Halton Hospitals NHS Foundation Trust	10	80.9	80.9	80.9	99.9	94.1	99.9	99.9	99.9	99.9	99.9

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RWY Calderdale and Huddersfield NHS Foundation Trust	10	99.3	99.3	99.3	100.0	99.7	100.0	100.0	100.0	100.0	100.0
RXC East Sussex Hospitals NHS Trust	10	96.5	96.5	96.5	100.0	98.7	100.0	99.9	100.0	100.0	100.0
RXF Mid Yorkshire Hospitals NHS Trust	10	85.9	85.9	85.9	97.4	90.9	97.4	96.7	97.4	97.4	97.4
RXH Brighton and Sussex University Hospitals NHS Trust	10	93.9	93.9	93.9	100.0	93.8	100.0	100.0	100.0	100.0	100.0
RXK Sandwell and West Birmingham Hospitals NHS Trust	10	97.9	97.9	97.9	99.9	98.1	99.9	99.9	99.9	99.9	99.9
RXL Blackpool, Fylde and Wyre Hospitals NHS Foundation Trust	10	96.3	96.3	96.3	100.0	98.6	100.0	100.0	100.0	100.0	100.0
RXP County Durham and Darlington NHS Foundation Trust	10	96.7	96.7	96.7	100.0	99.6	100.0	100.0	100.0	100.0	100.0
RXQ Buckinghamshire Hospitals NHS Trust	10	98.2	98.2	98.2	99.9	99.0	99.9	99.8	99.9	99.9	99.9
RXR East Lancashire Hospitals NHS Trust	10	96.2	96.2	96.2	100.0	98.3	100.0	100.0	100.0	100.0	100.0
RYJ Imperial College Healthcare NHS Trust	10	92.6	92.6	92.6	99.9	97.1	99.9	99.9	99.9	99.9	99.9
RYR Western Sussex Hospitals NHS Trust	10	99.1	99.1	99.1	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RC3 Ealing Hospital NHS Trust	9	87.4	87.4	87.4	99.4	0.0	99.4	99.2	99.4	99.4	99.4
RDD Basildon and Thurrock University Hospitals NHS Foundation Trust	9	80.1	80.1	80.1	100.0	72.5	100.0	100.0	100.0	100.0	100.0

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RHM Southampton University Hospitals NHS Trust	9	90.8	90.8	90.8	99.9	0.0	99.9	99.9	99.9	99.9	99.9
RJF Burton Hospitals NHS Foundation Trust/South Staffs PCT	9	98.9	98.9	98.9	100.0	0.0	100.0	100.0	100.0	100.0	100.0
RJZ King's College Hospital NHS Foundation Trust	9	98.0	98.0	98.0	100.0	0.0	100.0	100.0	100.0	100.0	100.0
RN7 Dartford and Gravesham NHS Trust	9	99.1	99.1	99.1	100.0	0.0	100.0	100.0	100.0	100.0	100.0
RQM Chelsea and Westminster Hospital NHS Foundation Trust	9	88.4	88.4	88.4	99.9	0.0	99.9	99.8	99.9	99.9	99.9
RTE+ +Gloucestershire Hospitals NHS Foundation Trust/Gloucestershire PCT	9	94.0	94.0	94.0	100.0	68.3	100.0	100.0	100.0	100.0	100.0
RTP Surrey and Sussex Healthcare NHS Trust	9	83.5	83.5	83.5	99.9	0.1	99.9	99.8	99.9	99.9	99.9
RXW Shrewsbury and Telford Hospital NHS Trust	9	96.5	96.5	96.5	100.0	42.1	100.0	100.0	100.0	100.0	100.0
RBZ Northern Devon Healthcare NHS Trust	7	0.0	0.0	0.0	100.0	96.9	100.0	100.0	100.0	100.0	100.0
RC9 Luton and Dunstable Hospital NHS Foundation Trust	7	0.0	0.0	0.0	100.0	98.1	100.0	100.0	100.0	100.0	100.0
REP Liverpool Women's NHS Foundation Trust	7	66.0	66.0	66.0	100.0	99.5	100.0	100.0	100.0	100.0	100.0
RGQ Ipswich Hospital NHS Trust	7	0.0	0.0	0.0	100.0	93.9	100.0	100.0	100.0	100.0	100.0
RJC South Warwickshire NHS Foundation Trust	7	0.0	0.0	0.0	99.9	99.6	99.9	99.9	99.9	99.9	99.9

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
RJD Mid Staffordshire NHS Foundation Trust	7	0.0	0.0	0.0	95.4	93.1	95.4	94.6	95.4	95.4	95.4
RLU Birmingham Women's NHS Foundation Trust	7	29.8	29.8	29.8	100.0	97.7	100.0	100.0	100.0	100.0	100.0
RM2 University Hospital of South Manchester NHS Foundation Trust	7	0.0	0.0	0.0	100.0	96.5	100.0	100.0	100.0	100.0	100.0
RQ8 Mid Essex Hospital Services NHS Trust	7	0.0	0.0	0.0	100.0	98.4	100.0	99.9	100.0	100.0	100.0
RQQ Hinchingbrooke Health Care NHS Trust	7	3.2	3.2	3.2	99.9	91.6	99.9	99.9	99.9	99.9	99.9
RV8 North West London Hospitals NHS Trust	7	0.0	0.0	0.0	99.8	85.2	99.8	99.7	99.8	99.8	99.8
RVV East Kent Hospitals University NHS Foundation Trust	7	3.0	3.0	3.0	99.8	83.8	99.8	99.8	99.8	99.8	99.8
RWF Maidstone and Tunbridge Wells NHS Trust	7	0.6	0.6	0.6	100.0	99.9	100.0	100.0	100.0	100.0	100.0
RWH East and North Hertfordshire NHS Trust	7	0.0	0.0	0.0	99.1	80.0	99.1	98.7	99.1	99.1	99.1
RWP Worcestershire Acute Hospitals NHS Trust	7	1.6	1.6	1.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0
RX1 Nottingham University Hospitals NHS Trust	7	55.6	55.6	55.6	100.0	97.3	100.0	100.0	100.0	100.0	100.0
RXN Lancashire Teaching Hospitals NHS Foundation Trust	7	71.6	71.6	71.6	99.3	82.0	99.3	99.0	99.3	99.3	99.3
RYQ South London Healthcare NHS Trust	7	42.7	42.7	42.7	99.2	94.7	99.2	99.0	99.2	99.2	99.2

Trust	Number of times 80% criteria met	Healthy mother	Healthy baby	Healthy mother/healthy baby dyad	Delivery with bodily integrity	Normal birth	Spontaneous vaginal delivery	Intact perineum	Elective caesarean	Emergency caesarean	Caesarean
5QK Wiltshire PCT	6	0.0	0.0	0.0	98.0	0.0	98.0	97.5	98.0	98.0	98.0
RA2 Royal Surrey County NHS Foundation Trust	6	0.4	0.4	0.4	99.9	0.0	99.9	99.9	99.9	99.9	99.9
RAE Bradford Teaching Hospitals NHS Foundation Trust	6	72.8	72.8	72.8	99.9	73.1	99.9	99.9	99.9	99.9	99.9
RBA Taunton and Somerset NHS Foundation Trust	6	23.4	23.4	23.4	99.1	0.0	99.1	98.8	99.1	99.1	99.1
RBT Mid Cheshire Hospitals NHS Foundation Trust	6	4.3	4.3	4.3	100.0	0.0	100.0	99.8	100.0	100.0	100.0
RD7 Heatherwood and Wexham Park Hospitals NHS Foundation Trust	6	0.0	0.0	0.0	100.0	0.0	100.0	100.0	100.0	100.0	100.0
RJ2 The Lewisham Hospital NHS Trust	6	62.6	62.6	62.6	99.7	73.0	99.7	99.5	99.7	99.7	99.7
RJ6 Mayday Healthcare NHS Trust	6	44.0	44.0	44.0	100.0	77.8	100.0	100.0	100.0	100.0	100.0
RJ7 St George's Healthcare NHS Trust	6	0.0	0.0	0.0	100.0	0.0	100.0	100.0	100.0	100.0	100.0
RN5 Hampshire Hospitals NHS Foundation Trust	6	41.0	41.0	41.0	99.5	79.4	99.5	99.4	99.5	99.5	99.5
RPA Medway NHS Foundation Trust	6	49.1	49.1	49.1	100.0	48.3	100.0	100.0	100.0	100.0	100.0
RR1 Heart of England NHS Foundation Trust	6	75.8	75.8	75.8	99.8	0.0	99.8	99.8	99.8	99.8	99.8
RRV University College London Hospitals NHS Foundation Trust	6	0.0	0.0	0.0	100.0	5.5	100.0	100.0	100.0	100.0	100.0
RTR South Tees Hospitals NHS Foundation Trust	6	47.5	47.5	47.5	100.0	0.0	100.0	100.0	100.0	100.0	100.0

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library