THE UNIVERSITY *of* EDINBURGH

COMPREHENSIVE DATA ANALYSIS TO STUDY PARTURITION

GEMMA C SHARP

THE UNIVERSITY
of EDINBURGH

THESIS SUBMITTED IN THE FULFILMENT OF THE DEGREE OF DOCTOR OF PHILOSOPHY.

SEPTEMBER 2013

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

DECLARATION

Except where due acknowledgement is made by reference below, the studies undertaken in this thesis were the unaided work of the author. I am very grateful for the following assistance:

Chapter 3

Nanette Hibbert synthesised cDNA, conducted quantitative real time PCR and gave me guidance in conducting PCR myself.

Doctor Lawrence Hutchinson assisted and guided me in extracting and amplifying RNA.

Doctor Donald Dunbar and Doctor Jon Manning of CVS Bioinformatics preprocessed and built a database of the microarray data.

Chapter 4

Dr Hongwu Ma gave me guidance in computer model development.

Chapter 6

Professor Graham Horgan of Biomathematics and Biostatistics Scotland, and Daniel Ayoubkhani of the UK Office for National Statistics provided advice during model design.

Chapter 7

In accordance with Cochrane guidelines, Doctor Sarah Stock and I independently extracted data from trial papers. Sarah Stock and Professor Jane Norman provided a clinical perspective and revised the protocol and review before submission.

No part of the work described in this thesis has been previously accepted for, or is currently being submitted in candidature for another degree.

Signed ……..
Gemma Sharp

ACKNOWLEDGEMENTS

I'm grateful to the many people who have helped me during this PhD. Firstly, I would like to thank my supervisor, Prof Jane Norman, who has always been supportive, encouraging and given me confidence that I would get to the point where I'd be writing these acknowledgements. I am also very grateful to Prof Philippa Saunders for coming on board as my second supervisor and for always giving positive, useful comments on my work.

I would like to say a big thank you to Dr Sarah Stock, Dr Donald Dunbar, Dr Jon Manning, Prof Tom Freeman, Mike Millar, Prof Graham Horgan and Nanette Hibbert for the work they have helped me with. Particular thanks go to Dr Hongwu Ma and Dr Lawrence Hutchinson for the many long chats about the direction of my work, which I really valued. I would also like to extend my thanks to everyone at the Centre for Reproductive Health for being lovely, helpful, friendly people.

Offering personal and sometimes financial (thanks Mum!) support, my family and friends have been invaluable sources of kind words and encouragement. To Mum, Dad, Carn, Jack, Fiona, Gary, Owen, Guy, Shalini, Lawrence, Stuart, Paul, Nick, Ellie and anyone else who has ever helped me remember that there is life outside of the PhD, I'd like to say a huge thank you. Particular thanks go to Lorraine, Sara and Sofia for keeping me sane/joining me when I'm feeling less than sane at work. I will never forget R&B Fridays (even though it only happened once), and if I ever earn enough when I get a proper job I'll be buying a cold room to store my Prosecco in. Finally, a special thanks goes to James for reminding me to "do science" when I've been tempted by daytime TV over the past couple of months, and for proof-reading all these words.

PUBLICATIONS AND PRESENTATIONS

Publications that have arisen from the work in this PhD:

- Sharp GC, Saunders PTK, Greene SA, Morris AD, Norman JE (2014) Intergenerational transmission of postpartum haemorrhage risk – analysis of two Scottish birth cohorts. Am J Obstet Gynecol (in press).

- Sharp GC, Ma H, Saunders PTK, Norman JE (2013) A computational model of lipopolysaccharide-induced nuclear factor kappa B activation: a key signalling pathway in infection-induced preterm labour. PLoS ONE 8(7): e70180.

- Sharp GC, Saunders PTK, & Norman JE (2013) Computer models to study uterine activation at labour. Mol Hum Reprod. 19(11):711-7.

- Sharp GC, Stock SJ, Norman JE (2012) Fetal assessment methods for improving neonatal and maternal outcomes in preterm prelabour rupture of membranes. Cochrane Database of Systematic Reviews. Issue 11. Art. No.: CD010209.

PDFs of these papers are included on the CD accompanying this thesis.

Work included in this PhD was presented at the following conferences:

- Society for Epidemiologic Research, June 2013, Boston, MA – *poster presentation of work from Chapter 5*.
- Functional Genomics and Systems Biology, November 2011, Cambridge – *poster presentation of work from Chapter 4.*
- Royal College of Obstetricians and Gynaecologists Annual Academic Meeting, December 2011, London – *poster presentation of work from Chapter 4.*
- Eulerian and Lagrangian approaches towards quantitative prediction of premature labour, normal and dysfunctional full term labour, June 2011, Edinburgh – *oral presentation of work from Chapter 4.*
- Workshop on Systems Medicine, May 2011, Dundee – *poster presentation of work from Chapter 4.*

PUBLIC ENGAGEMENT ACTIVITIES

My PhD was funded by a University of Edinburgh Principal's Career Development scholarship, which provided a valuable opportunity to take part in training and activities to help develop non-research based skills that will benefit my future career. I chose to focus on public engagement skills, to help disseminate research to people outside of the University of Edinburgh and encourage a better public understanding of science.

Training

- BBC-run Research, Researchers and the Media – November 2010
- MRC Edinburgh Science Festival training – March 2011 and March 2012
- Popular Science Writing – April 2011
- Engaging Scotland conference – September 2011
- Edinburgh Beltane storytelling training – February 2012

Activities

- Egg and Sperm Race – November 2010 to September 2012 – Along with another PhD student at the Centre for Reproductive Health (Vicky Young), I designed and built an activity stall to communicate reproductive biology to music festival goers. We took the stall to Green Man festival in Wales in 2011 and 2012 where it was very popular.
- EUSci Science Magazine – November 2010 to May 2013 – I wrote several features and news articles on a range of scientific subjects for Edinburgh's Science magazine, EUSci. In 2011 I was selected as President of the EUSci student society.
- Edinburgh International Science Festival – April 2011 to April 2012 – I volunteered at the Science Festival for both the University and the MRC for two years. This involved using hands-on activities to explain biomedical science to children aged 4-12.
- I'm a Scientist Get Me Out of Here! – June 2011 – I took part in this online competition in which secondary school students ask scientists questions, and vote for their favourite scientist to win. I was the winner of my zone.

- Bright Club – August 2011 to August 2012 – Bright Club is a comedy night where academics perform stand-up routines about their research. I have performed six times in Edinburgh, Glasgow and at Green Man festival.
- Scientific Kitty – September 2011 to December 2012 – I set up a blog called Scientific Kitty where scientists review the science in popular media such as TV, films, books and games. The blog attracted reviews from scientists from all over the UK, received over 100 hits per day and was publicised by the British Science Association.
- London Student science correspondent – September 2012 – I wrote several articles for the science section of the London Student, the largest student-run newspaper in Europe.

Acknowledgements and awards

- National Co-ordinating Centre for Public Engagement (NCCPE) ambassador – My involvement in public engagement led to me being selected as an NCCPE ambassador in 2011.
- Edinburgh Beltane Annual Gathering 2011 – I was the only student member of a panel discussion about the importance of public engagement at the Edinburgh Beltane's annual gathering at the Botanic Gardens.
- Edinburgh Beltane Annual Gathering 2012 – I told a short "story" about how I had integrated public engagement with my PhD training to several audiences during the Edinburgh Beltane's annual gathering at the National Museum of Scotland.
- Society of Biology Science Communication Award – I was shortlisted for this award in August 2012.
- I'm a Scientist, Get Me Out of Here! – In this competition (described above) I won £500 to spend on science communication.
- EUSci funding – As president of EUSci, I secured £3500 from the MRC (£2500) and the Physiological Society(£1000) to fund printing costs of EUSci Magazine.
- Egg and Sperm Race funding – Vicky and I secured £7500 from the MRC (£6500) and the Society for Endocrinology (£1000) in 2011 and 2012 for the Egg and Sperm Race.

ABSTRACT

Our limited understanding of the molecular mechanisms driving the onset of normal human parturition makes it difficult to identify 'what goes wrong' in conditions such as preterm labour (PTL), preterm prelabour rupture of membranes (PPROM) and postpartum haemorrhage (PPH). This incomplete understanding seriously hampers the development of effective ways to predict, prevent and treat parturition complications, which are a cause of significant neonatal and maternal morbidity. Two principal barriers to improving our understanding are 1) the great complexity of both the molecular interactions initiating parturition and the aetiology of parturition complications, and 2) the difficulty in generating relevant high quality molecular and epidemiological data. To help make sense of this complexity, data should be analysed comprehensively to maximise the amount of useful information gleaned from it.

This thesis aimed to explore the use of specialist methods to analyse novel and previously published data to study the molecular mechanisms initiating human parturition and the epidemiology of parturition complications.

The molecular mechanisms initiating parturition were explored through a gene expression microarray of labouring and non-labouring myometrial tissue. This is the largest microarray of its kind to date. Functional analysis and a network graph approach were used to reveal genes and molecular pathways associated with labour. The first ever meta-analysis of similar myometrial microarray datasets was also conducted to assess the reliability and generalisability of the results. This work supported the hypothesis that labour is associated with inflammatory events in the myometrium. A computer model of an inflammatory signalling pathway associated with infection-induced PTL was then built to provide proof of concept that such models can be used to study parturition. The model was based on published data and described lipopolysaccharide-induced activation of the transcription factor Nuclear Factor kappa B (NF-κB). This is the first attempt to generate a dynamic kinetic model that has relevance to the molecular mechanisms of PTL, and the first model of this pathway to explicitly include molecular interactions upstream of NF-κB activation.

The epidemiology of complications at parturition was explored using three methods. Firstly, a novel approach was developed to use network graphs to visualise and analyse a dataset of nearly 50,000 birth records. The approach provided a quick and effective way to preliminarily explore relationships between exposures and pregnancy outcomes in an unbiased data-driven manner. Secondly, a record-linkage study of two datasets of birth records was conducted to determine risk factors for PPH, including intergenerational transmission of risk. This confirmed several known risk factors of PPH and showed that women whose mothers or grandmothers had PPH do not appear to be at increased risk themselves. Finally, a systematic review and meta-analysis of three randomised controlled trials investigated the effectiveness of fetal assessment methods in improving maternal and neonatal outcomes following PPROM. The review concluded that there is currently insufficient evidence on the benefits and harms of any method of fetal assessment, and further randomised controlled trials are required.

1 LITERATURE REVIEW: EFFECTIVE DATA ANALYSIS TO STUDY PARTURITION

1.1 INTRODUCTION

Our limited understanding of the molecular mechanisms driving the onset of normal human parturition makes it difficult to identify 'what goes wrong' in conditions such as preterm labour, preterm prelabour rupture of membranes and postpartum haemorrhage. This incomplete understanding also seriously hampers the development of effective ways to predict, prevent and treat parturition complications, which are a cause of significant neonatal and maternal morbidity. Two principal barriers to improving our understanding are 1) the great complexity of both the molecular interactions initiating parturition and the aetiology of parturition complications, and 2) the difficulty in generating relevant high quality molecular and epidemiological data. To help make sense of this complexity, data should be analysed comprehensively to maximise the amount of useful information gleaned from it (Figure 1.1).

This review explores barriers to data acquisition and analysis in parturition research and outlines some of the methods that can be used to ensure generated data is analysed appropriately.

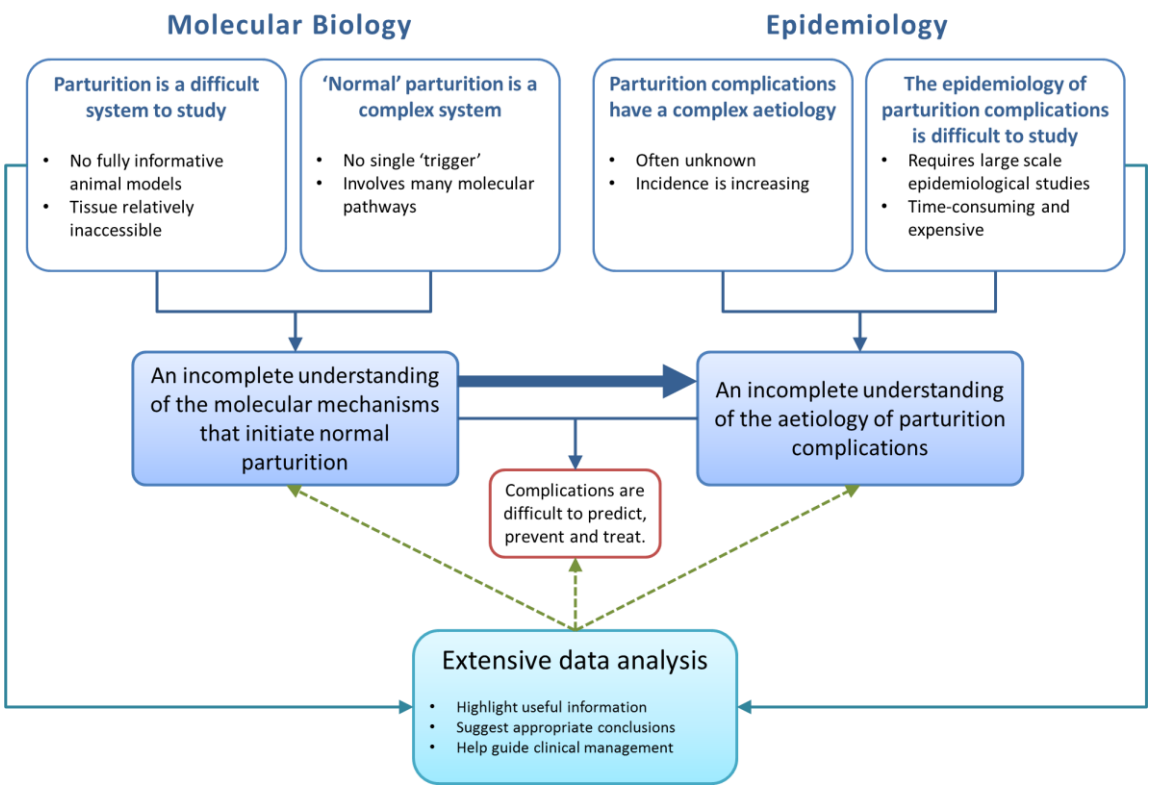


Figure 1.1. An overview of the argument for comprehensive data analysis in helping to overcome barriers to parturition research.

1.2 The molecular mechanisms driving the initiation of parturition are complex

At the molecular level, parturition research aims to identify genes, proteins and molecular pathways that play a role in initiating and regulating parturition. Through this work, many inflammatory and endocrine factors from the mother, fetus and placenta have been shown to play a role in human parturition (Figure 1.2).

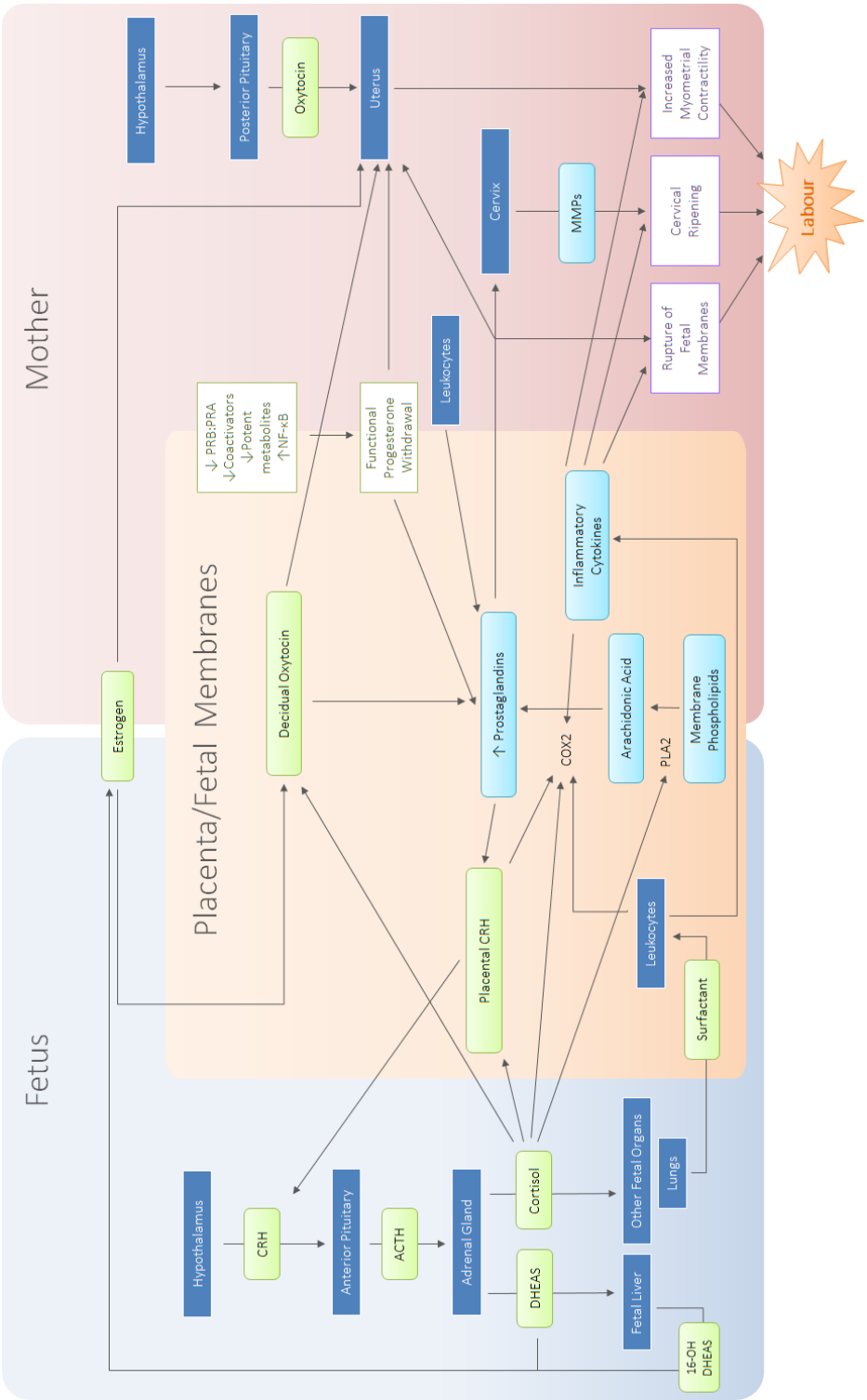


Figure 1.2. Key endocrine and inflammatory factors associated with the initiation of human parturition. ACTH, Adrenocorticotrophic Hormone; COX2, Cyclooxygenase; CRH, Corticotrophin-Releasing Hormone; DHEAS, Dehydroepiandrosterone; MMPs, Matrix Metalloproteinases; NF-κB, Nuclear Factor kappa B; PLA2, Phospholipase A; PRA, Progesterone receptor A; PRB, Progesterone Receptor B. Adapted from Snegovskikh.(357)

1.2.1 Inflammatory events associated with parturition

Increasingly, evidence suggests that the initiation of parturition is associated with an increase in expression of factors regulating inflammatory processes and changes in the

activation status of populations of immune cells. For example, studies conducted more than 10 years ago revealed that leukocytes (mostly neutrophils, macrophages and T cells) are attracted by increased tissue expression of chemokines and cell adhesion molecules (1,2) to invade the cervix, myometrium, and fetal membranes at parturition (3,4). This, in part (5), leads to increased cytokine expression in the myometrium and cervix (3). These cytokines attract yet more leukocytes (6), but also have a pronounced effect on the reproductive tissues, including remodelling of the cervix (7) and stimulating myometrial contraction through prostaglandin synthesis (8–13). Prostaglandins mediate inflammation, and prostaglandin production is increased in the fetal membranes and myometrium at labour. Chorionic prostaglandin dehydrogenase (PGDH) activity decreases in late pregnancy, which exposes the decidua and myometrium to pro-inflammatory prostaglandin F2 alpha (PGF2α) (14). PGF2α and prostaglandin E2 have a direct contractile effect; they bind to their receptors on myometrial myocytes to activate a signalling pathway, resulting in the activation of myosin light chain kinase and the release of intracellular calcium (15–17). Prostaglandins also mediate the release of metalloproteases, which weaken the placental membranes and facilitate membrane rupture and cervical ripening (18,19).

1.2.2 Endocrine events associated with parturition

Endocrine-dependent events associated with parturition include functional progesterone withdrawal, functional estrogen activation, myometrial stimulation by corticotrophin-releasing hormone, and regulation by oxytocin.

1.2.2.1 Functional progesterone withdrawal

Progesterone plays a critical role in maintaining myometrial relaxation throughout much of pregnancy (20). In many mammals, circulating progesterone levels fall prior to labour, and this is thought to promote increased myometrial contractility. However, it has been recognised for over three decades that in women, circulating progesterone levels remain constant in the late third trimester of pregnancy and at parturition (21–25). It has therefore been proposed that the onset of human parturition is associated with a 'functional progesterone withdrawal', whereby circulating concentrations of progesterone are maintained but they have a diminished effect on the uterus. Mechanisms proposed for this include changes in the ratios of the progesterone receptor isoforms (specifically, an increase in progesterone receptor A (PR-A) relative to B (PR-B)) (26), local metabolism of

progesterone (27), and changes in levels of cofactors affecting progesterone receptor function (28). Interestingly, prostaglandin PGF2α is able to stimulate preferential expression of progesterone receptor A (PR-A) in myometrial cells (29), which provides a potential link between inflammation, prostaglandins, progesterone and the onset of labour (30).

1.2.2.2 Functional estrogen activation

Estrogen stimulates uterine contractions by increasing the expression of contraction-associated protein genes such as connexin-43, the oxytocin receptor, PTGS2 (prostaglandin endoperoxide synthase 2) and the PGF2α receptor (26,31–35). This has recently been shown to be mediated by extranuclear signalling through estrogen receptor alpha (ERα) via activation of the ERK mitogen-activated kinase (MAPK) pathway (36). In women, there is a gradual increase in circulating estrogens throughout pregnancy (37), and this is accelerated in the final few weeks, but estrogen levels during parturition are not significantly different to those just before it (21,38). In a study conducted on a small number of myometrial samples from labouring and non-labouring women, Mesiano et al.(26) reported that both the PR-A/PR-B mRNA ratio, and ERα mRNA were significantly increased in labouring myometrium. They also showed that the expression of ERα mRNA was positively correlated with expression of COX2 and oxytocin receptor mRNAs in non-labouring myometrium and suggested the coordination of progesterone withdrawal and estrogen activation may be required for parturition.

1.2.2.3 Corticotropin-releasing hormone

Corticotropin-releasing hormone (CRH) is released from the placenta in human pregnancy (39), and this is associated with an increase in maternal plasma CRH levels that peaks at delivery (40,41). The bioavailability of CRH is limited by a binding protein (CRH-BP), but levels of CRH-BP decline in the last six weeks of pregnancy further increasing the circulating concentrations of unbound "active" CRH (42). Mclean *et al.* found associations between faster rises in CRH levels and preterm birth, and between slower rises and post-term birth (43). Subsequently, they put forward the 'placental clock hypothesis' – that CRH controls the timing of parturition. It is unclear exactly how CRH might exert this effect, however CRH can stimulate the activity of the COX2 that synthesizes PGF2α (44), which provides a potential link between placental production of CRH, prostaglandins and functional progesterone withdrawal.

1.2.2.4 Oxytocin and its receptor

Although not fully understood, oxytocin is known to play an important role in human parturition. Synthetic oxytocin is used to augment labour in a clinical setting (45), and oxytocin receptor antagonists show some success in treating preterm labour and are licenced for this purpose in the UK (46). Oxytocin concentrations in the maternal plasma (47) and gestational tissues (48) increase with labour, which causes prostaglandin formation in the chorion, decidua and amnion (49–52). Oxytocin receptor expression also increases in term myometrial tissue (33,53) and this leads to direct contractile effects (53).

1.2.3 The relative importance of these mechanisms is unclear

Despite decades of work elucidating several molecular mechanisms associated with the onset of human parturition, the relative importance of each of these factors or their interactions is still unclear. Crucially, this knowledge has also not led to a substantial improvement in the ability to predict, prevent or treat parturition complications and improve clinical outcomes. Clearly there are still large gaps in our knowledge, particularly regarding the potential complex interactions that may exist between the molecular pathways identified thus far.

1.3 The molecular biology of parturition is difficult to study

One of the main reasons we have a limited understanding of the molecular mechanisms initiating parturition is that it is a difficult system to study in humans. Attempts to improve our understanding are often restricted by the relative inaccessibility of human gestational tissues to study during pregnancy, and by the lack of fully informative animal models.

1.3.1 Experiments using humans

Ideally, the most effective experiments to study human parturition would use human participants. However, there are obvious ethical, clinical and economic considerations that make this impossible (54). Most human gestational tissues are inaccessible during pregnancy without causing harm. Some tissues such as the umbilical cord, amnion, chorion and placenta are relatively easy to collect post-partum. However collection of maternal tissues is more difficult. At Caesarean section, decidua and myometrium can be

extracted opportunistically, but ethical regulations regarding informed consent are necessarily stringent. Myometrial tissue is usually collected from the lower segment of the uterus as this offers least procedural complication whilst yielding useful amounts of starting material for a variety of experimental purposes. One has to be aware nonetheless that molecular pathways active in this region of the tissue may differ from those in other areas such as the fundus (54–56). Therefore the lower segment may not be an appropriate model of the whole uterus. Tissue from labouring women is particularly difficult to obtain because in these cases, the Caesarean section is performed as an emergency. Additionally, results from any investigations using these tissues may be affected by the clinical indications for Caesarean delivery (54).

1.3.2 Animal models

The use of animal models can overcome some of the problems inherent in human studies. Tissue can be collected at precise time-points throughout gestation and parturition, and the importance of certain molecular mechanisms can be investigated using genetic manipulation. Many of the molecular mechanisms that play a role in initiating parturition are likely to be common to humans and animal models. However, there are no fully informative animal models of parturition. In most mammals, such as the sheep, rat and mouse, a fall in maternal circulating progesterone appears to be the initiating mechanism (54,57,58). This differs from humans where, as discussed above, progesterone levels remain stable at the end of pregnancy and during parturition (21,22). In fact, in one comparison of genome-wide gene expression detected by microarray, surprisingly few genes that changed significantly throughout pregnancy were the same in the mouse and human (59). Nonhuman primates also show no maternal serum progesterone withdrawal and have a uterine anatomy closer to that of humans (54), there are stringent ethical regulations surrounding the use of nonhuman primates, and the studies are often financially and logistically impractical.

The molecular mechanism of parturition in these animal models appears to be far simpler than the mechanisms initiating and regulating human parturition. Human parturition does not appear to be associated with a single ‘trigger’, but rather a more gradual culmination of several different pathways (54). In other mammals, spontaneous preterm labour is rare but can be reliably induced pharmacologically, suggesting that the mechanisms initiating animal parturition are also more efficient than in the human (58). The limitations of

taking the human-only approach to parturition research (outlined in Section 1.3.1) highlight the need to also access information from appropriate non-primate animal models. Mitchell and Taggart argue that the guinea pig is a more relevant model for several reasons such as the absence of systemic progesterone withdrawal and the natural occurrence of spontaneous preterm labour in this species. However, further characterisation of the model is necessary (54).

1.4 The aetiology of parturition complications is complex

Largely due to the complexity of human parturition and the difficulties inherent in its study, the causes of complications related to parturition are often unclear. This limits the development of effective strategies to predict, prevent and treat such complications. However, it is important that we improve our understanding, because parturition complications are a cause of significant neonatal and maternal morbidity and mortality.

Epidemiological research aims to elucidate aetiology by identifying causative risk factors that may lead to the development of effective management strategies. A better understanding of the risk factors associated with parturition complications will help 1) identify at-risk patient and develop risk-specific treatments, 2) define a subset of women in which to study specific interventions, and 3) provide important insights into the mechanisms behind the complication. Table 1.1 outlines what is known about the epidemiology of some of the major parturition complications. Risk factors and current management strategies are discussed below.

Table 1.1. The epidemiology of some of the major complications associated with parturition.

Complication	Definition	Incidence	Clinical outcomes
Spontaneous preterm labour (PTL)	• Spontaneous labour (regular uterine contractions associated with cervical ripening) occurring before 37 weeks' gestation.	• 3% of all deliveries in England and Wales in 2011-2012 (60).	• Preterm birth, which is the leading cause of neonatal morbidity and mortality and often associated with long-term morbidity such as neurodevelopmental delays.
Preterm prelabour rupture of membranes (PPROM)	• Rupture of the fetal membranes more than one hour before the onset of labour, and before 37 weeks' gestation.	• 1-2% of all deliveries (61).	• 30-40% of PPROM cases are related to preterm birth. • Morbidity including cord prolapse placental abruption. • Increased susceptibility to intrauterine infection.
Dystocia	• Abnormal or difficult parturition.	• 3-8% of all deliveries(62).	• Assisted delivery (forceps, ventouse or Caesarean section). • Fetal asphyxia, brachial nerve damage, death.
Postpartum haemorrhage (PPH)	• ≥500ml blood loss from the genital tract in the first 24 hours after childbirth.	• 3-26% of all deliveries (63,64).	• Leading cause of maternal death worldwide. • Sixth highest direct cause of maternal death in the UK (65). • Maternal morbidity including coagulopathy, respiratory failure, renal failure, sepsis and hysterectomy.

1.4.1 Spontaneous preterm labour

Spontaneous preterm labour (regular contractions associated with cervical ripening, occurring before 37 weeks' gestation) is the cause of 40-45% of preterm births. The aetiology is unclear, and it is now considered to be a syndrome composed of multiple mechanisms (66). It may represent an early idiopathic activation of the normal mechanisms that initiate labour, or be associated with different mechanisms related to pathological insults. It is associated with many risk factors including intrauterine infection, previous preterm birth, black race, low socioeconomic status, low and high maternal age, short inter-pregnancy interval, low BMI, sociological and psychological stress and smoking during pregnancy (66–68). These risk factors are of only some use in predicting preterm labour (69). Similarly, biomarkers such as cervical length and cervicovaginal measurements of fetal fibronectin show a disappointing ability to predict preterm labour, especially in asymptomatic women (70). Current strategies to treat preterm labour once it

has begun involve tocolytic drugs to suppress uterine contractions. Such tocolytics include beta-agonists, calcium ion channel blockers, prostaglandin synthetase, nitric oxide donors and oxytocin antagonists (71). However, systematic reviews of randomised controlled trials (RCTs) have shown that there is no clear evidence tocolysis improves neonatal or maternal outcomes (72,73), but it may be useful if the small increased time to delivery allows in utero transfer to a hospital with specialist facilities, or completion of a course of corticosteroids to improve fetal lung function (71).

1.4.2 Preterm prelabour rupture of membranes

Twenty-five to 35% of preterm births are associated with preterm prelabour rupture of membranes (PPROM), which can precede preterm labour or indicated preterm delivery. Risk factors associated with PPROM are similar to those associated with preterm labour, including smoking, previous preterm delivery or PPROM, uterine overdistension, black race, and low socioeconomic status (74–77). The pathophysiology of PPROM is largely unknown, but excessive stretching of the membranes, decreased collagen content, placental abruption, and programmed amniotic cell death have all been proposed as mechanisms (78). However, 25 to 50% of cases are associated with intrauterine infection and inflammation (79). Methods to predict PPROM are also similar to those used to predict preterm labour, with short cervical length and PPROM in a previous pregnancy being the most effective predictors (80,81). Strategies to manage pregnancies with PPROM include tocolysis, antibiotics to treat intrauterine infection, and amnioinfusion to restore amniotic fluid volume. Although these interventions can increase time to delivery, there is limited evidence that they significantly reduce neonatal death and morbidity (82–84).

1.4.3 Dystocia

Dystocia is the result of weak or uncoordinated uterine contractility. It is one of the most common causes of prolonged labour and occurs most frequently in nulliparous pregnancies (85). The pathophysiological mechanisms behind dystocia are not fully understood, but could be myogenic, neurogenic, hormonal, or (most likely) multifactorial (86,87). There is some evidence of a genetic component to dystocia (62), with dystocia being more prevalent among women whose mothers had dystocic labours than among those whose mothers had normally progressing labours (88). Previous dystocia also appears to increase the risk of dystocia in a subsequent delivery (89). Other risk factors include a large interpregnancy interval, older maternal age, high maternal BMI, short

maternal stature, high caffeine intake, gestational diabetes, hypertension and hydramnios (90–95). Vaginal examination during labour can provide information on cervical dilatation and fetal head-to-cervix contact that could be useful in predicting dystocia in labouring women (96). Interventions to augment labour, such as amniotomy and oxytocin, are used to treat dystocia. These show some success in reducing the length of labour and the need for operative delivery (45,97,98), however dystocia is a large contributor to the rate of assisted delivery by forceps/ventouse or Caesarean section (99).

1.4.4 Postpartum haemorrhage

The annual incidence of postpartum haemorrhage (PPH) appears to be rising steadily, even in high resource countries (100). PPH can be associated with a failure of the uterus to contract adequately after birth (atonic PPH; 90% of cases), trauma to the genital tract (traumatic PPH; 7% of cases), or bleeding due to retention of placental tissue or failure in the coagulation system (3% of cases) (101). Known risk factors include young or old maternal age, Caesarean delivery, hypertension, polyhydramnios, intrauterine infection, multiple gestation and antepartum haemorrhage (64), however the aetiology is often unclear and PPH may occur in women with no identifiable risk factors. Few studies have investigated the power of these risk factors to predict PPH, but those that have found low predictive values (102,103). Treatments include drugs (such as oxytocin and prostaglandins) to increase uterine contractions, blocking the uterine artery, haemostatic drugs to stop the bleeding and active management of the third stage of labour. In severe cases, hysterectomy may be necessary. Few trials have assessed the efficacy of these interventions in improving outcomes (104–106), but two trials comparing misoprostol (synthetic prostaglandin E1) with placebo showed that treatment with misoprostol was not associated with any significant reduction in maternal mortality, hysterectomy, blood transfusion or rates of manual removal of retained placenta (107,108).

1.5 Epidemiological studies are difficult to conduct

A better understanding of the aetiology of parturition complications should lead to better prediction and treatment strategies, which will hopefully result in a reduction in incidence and improved neonatal and maternal outcomes. Population-level epidemiological studies are necessary to explore risk factors and the efficacy interventions for parturition complications. However, large clinical study designs (including epidemiological studies such as longitudinal cohort studies, case-control studies) require appropriate validated

datasets and randomised controlled trials are time-consuming, logistically difficult and expensive.

Although all are significant clinical problems, the complications discussed above occur in a relatively small percentage of pregnancies. This means large numbers of pregnant women need to be recruited to ensure the sample contains a sufficient number of cases for adequate power. Furthermore, to identify lifestyle risk factors for complications, it may be necessary to follow women for years prior to their pregnancy. Again, this means very large numbers of women need to be recruited. Women may be less likely to take part in an intervention trial during pregnancy, especially if they are asymptomatic or at low risk, so recruitment may be difficult. The drop-out rate may be high if pregnant women, who often have other commitments, are expected to attend a clinic or research centre to receive an intervention or provide data.

1.6 Attempting to overcome these barriers to parturition research

The complexity of normal and complicated parturition and the difficulties inherent in its study mean that relevant, high quality molecular and epidemiological data is precious. In order to maximise the amount of useful information gleaned, existing available data should be analysed comprehensively, using the most up to date analytical techniques. Table 1.2 outlines some useful strategies to conduct comprehensive analysis of existing data.

Table 1.2. An overview of some useful strategies in conducting comprehensive analysis of existing molecular and epidemiological data.

		Employing special techniques	
		Systems biology	**Combining data**
Using existing data	**Raw data made available through institutions and organisations**	• Analysis of global gene expression microarray data. • Computational epidemiology (hypothesis-generating techniques using unstructured data).	• Record-linking between different databases of epidemiological data.
	Raw data available in online repositories	• Analysis of global gene expression microarray data. • Protein pathway analysis using an online database of protein interactions.	• Meta-analysis of global gene expression data from different published microarray studies.
	Data reported in the literature	• Computational modelling of molecular pathways.	• Comparison of published lists of differentially expressed genes • Systematic review and meta-analysis of published randomised control trials.

1.6.1 Making efficient use of existing data

Molecular and epidemiological data can be made available to researchers in a variety of ways. Published papers are the most abundant source, although these generally provide the least amount of detail. Data is often presented in the form of a summary table or graph, so its interpretation may be limited by the chosen method of analysis and aims of the original work. Data in published papers is also open to reporting bias, whereby 'positive', confirmatory findings are more likely to be reported than important but negative findings that might refute current ideas (109). Increasingly however, raw data from published studies is being made available to download from online repositories as well as in the form of supplementary files linked to their publications. In ArrayExpress (http://www.ebi.ac.uk/arrayexpress/), the database of gene expression microarray data, there are currently 11 datasets relevant to human parturition. The raw data files are freely available to download for further analysis. Considering microarray experiments are

expensive to conduct and human gestational tissue is difficult to collect, online repositories such as ArrayExpress are extremely valuable resources. Other online repositories combine data from published studies to create useful, searchable catalogues of specific information. Examples include:

- KEGG Pathway (Kyoto Encyclopaedia of Genes and Genomes; http://www.genome.jp/kegg/pathway.html) – a collection of manually-drawn pathway maps representing knowledge of molecular interaction and reaction networks.
- BioModels (http://www.ebi.ac.uk/biomodels-main/) – a repository of computational models of biological processes, collected from the scientific literature and curated and cross-referenced to external data resources. The original model files are available for download.
- Uniprot (http://www.uniprot.org/) – a comprehensive, high-quality resource of protein sequence and functional information.

Alternatively, authors may be willing to provide raw data on request, or it may be made available for a fee through research institutions or other organisations. For example, the Health Informatics Centre (HIC) at the University of Dundee makes Scottish patient health records from the Information Services Division (part of NHS Scotland), and data from cohort studies available to researchers from other institutions. They maintain several databases that could be used to study the epidemiology of parturition complications in Scotland, including SMR02 (Scottish Morbidity Records 02), which contains detailed information on hospital maternity admissions in collected from January 1975 to present.

Improving access to quality data is critical to facilitating research, and could lead to advances in our understanding of normal and complicated parturition (110).

1.6.2 Using special analytical methods

Good access to high quality data is only beneficial if it is analysed appropriately and comprehensively. Special analytical methods should be employed to analyse data in the most efficient, useful way. These methods include, 1) using a data-driven 'systems' approach to generate as well as test hypotheses, and 2) combining data from different sources to provide extra information, increase the power and generalisability of results, and to compare and contrast different studies.

1.6.2.1 Systems biology approaches

Much of the research conducted in the field of obstetrics to date has applied the traditional 'reductionist' approach, with an emphasis on single cells or tissues focussing on the functions of certain molecules or key pathways, or the prevalence of certain risk factors. Although this approach has undoubtedly advanced our knowledge of the molecular and epidemiological mechanisms of parturition, it has not led to a significant improvement in neonatal or maternal outcomes. The novel paradigm of 'systems biology' provides a way to consolidate and extend current knowledge to characterise complex processes through describing the components of a system and their interactions at a global level. It considers the molecular biology and epidemiology of parturition from a holistic perspective and attempts to identify emergent properties.

1.6.2.1.1 Systems biology in molecular research

Research that uses a systems approach to study the molecular mechanisms initiating and regulating parturition is increasingly appearing in the literature thanks to recent advances in techniques that allow whole systems and emergent properties to be studied (55,110–128). These techniques fall broadly into two main categories: 1) methods to analyse high-throughput data, and 2) computational modelling. Both techniques involve large amounts of molecular data and existing knowledge of molecular pathways.

1.6.2.1.1.1 High-throughput data - Transcriptomics

High-throughput data can be generated to study biological systems at a global level. This means that rather than choosing certain genes/proteins/etc. to study in isolation, elements of all these entities in a biological system can be studied simultaneously. Examples of high-throughput techniques include:

- Transcriptomics – to study global gene expression through quantification of mRNA transcripts using microarray technology.
- Genomics – to study DNA variants between individuals using high throughput genome-wide sequencing.
- Proteomics – to study global protein expression using mass spectrometry and protein microarray technology.

- Metabolomics – to study metabolites (the products of cellular metabolism) at a global level using mass spectrometry.
- Epigenomics – to study epigenetic modifications to the genome such as DNA methylation using global methylation arrays.

Of these techniques, microarray experiments to study the transcriptome have been most commonly applied in parturition research.

Microarrays allow the expression of tens of thousands of genes to be monitored simultaneously in cells or tissues in a wide range of physiological or pathophysiological situations. They generate vast quantities of data that are analysed comprehensively using statistical and bioinformatics techniques to suggest biological processes and pathways associated with changes in gene expression (129). Microarray analysis of pregnant non-labouring and labouring gestational tissues aims to identify genes and pathways associated with the onset and regulation of labour. This will improve our understanding of the mechanisms initiating parturition, and expression of these genes may act as useful biomarkers to predict preterm labour (113). Several microarray experiments conducted in the last 10 years have reported labour-associated changes in gene expression in the cervix, myometrium and the placenta in both humans (55,59,112,114,115,119,120,130–143) and rodents (59,144–147). These have offered support for the observation that parturition is an inflammatory event associated with an upregulation of several inflammatory cytokines in the gestational tissues. One study also suggests there is a different gene expression signature associated with preterm compared to term labour in humans (115). Two microarray studies have compared gene expression in normal term labouring myometrium with labouring myometrium from women with dystocia. These have also identified genes that are differentially expressed in myometrium from complicated compared to normal parturition (148,149).

Thorough, appropriate analysis of microarray data is crucial. The mathematical method used to normalise expression values to allow direct comparison between samples can have a substantial effect on downstream analysis. Box-plot and principal component analysis of raw and normalised gene expression values can be useful in identifying biological samples that vary significantly from other samples and might be best excluded. Gene expression fold changes give a reasonable measure of effect size, but should be combined with a statistical method to define the level of ‘confidence’. Similarly, the

threshold beyond which fold changes are considered biologically interesting requires some consideration. Certain genes (for example a transcription factor) that might show a small change in expression may be overlooked, but even a small change in their expression could have a large downstream effect. Therefore, lists of up- and down-regulated genes can be informative, but are sometimes of limited value. Bioinformatics techniques can provide more interesting insights into the data by identifying clusters of genes with similar expression patterns over all samples, or mining gene lists for enrichment of genes involved in certain biological processes or KEGG protein pathways. Recently developed network graph approaches (in which proteins/genes are represented as nodes connected by edges representing interaction relationships or the degree of similarity in expression) are also a valuable tool in microarray analysis (150). These have identified patterns in gene expression data in other biological systems (151–154), but have been used relatively infrequently to investigate human parturition. However, since 2010 several investigators have applied this approach to microarray data and shown that the majority of the genes differentially expressed in labouring compared to non-labouring tissue are downstream targets of the transcription factor nuclear factor-kappaB (NF-κB) and pattern recognition receptors such as toll-like receptors (TLRs)(130,133,155). This supports the hypothesis that labour is associated with an inflammatory event.

1.6.2.1.1.2 Computational modelling

Computer models assemble complex experimentally-derived data into concise mathematical representations of biological systems that reveal more about their operation and structure than wordy descriptions and/or diagrams ever could. They can consolidate the current state of knowledge of a system, test that it is correct, highlight gaps in our understanding and can inform design of biological experiments to bridge those gaps. Furthermore, a well-validated model that describes the system accurately can be used to predict how the system will react to different stimuli in 'virtual' experiments that might be time consuming and expensive in the lab (156). Such models are particularly useful if they can reproduce or predict behaviours beyond those they were initially designed to describe. They can then be used to generate and test hypotheses about how the biological system might be organised and regulated. These hypotheses can be tested using wet lab experiments, and models can be updated with new data and used to generate further hypotheses.

It is important to note that the reliability of a model hinges on the availability and quality of the data obtained through such experiments. Models are also heavily dependent on decisions made regarding model design and which data to include or exclude. Wet lab validation helps to reduce this risk of bias, as does the availability of high quality published desciptions of the structure and dynamical behaviour of the system (157). Therefore, it is crucial that researchers have access to high quality data through online repositories. Model files should also be made available to the research community to allow testing by other groups and inform the design of further models (158). Online repositories of published models such as BioModels (159), and standardised file formats such as SBML (Systems Biology Markup Language) (160) facilitate this.

Computer modelling of the events involved in initiating parturition should help identify key pathways or mechanisms that might be altered in pathologies such as preterm labour or dystocia. Such models have several potential applications. For example, they could be used to test, *in silico*, hypotheses about drugs to treat or prevent such complications. Models could also help identify a sub-group of women who are likely to respond best to certain treatments. In this way, they could provide a useful tool to help design expensive time-consuming clinical trials more efficiently. With further development, an effective *in silico* model could be used by health care practitioners to develop personalised medicine for patients on a pregnancy-by-pregnancy basis.

Despite the explosion in the use of bioinformatics in the past decade, computer models have been used relatively infrequently to study pregnancy (110). Table 1.3 outlines the models that have been developed to date to describe uterine electrophysiology and propagation and the molecular mechanisms that initiate parturition (these are reviewed in more detail in Sharp *et al.*(161), which was written during the course of this PhD). Far more models focus on uterine electrophysiology than molecular pathways to uterine activation, but models of molecular mechanisms will be useful in identifying the key pathways and predicting the effects of drugs to alter such pathways. Computer models that integrate models of electrophysiological and molecular mechanisms will be particularly useful.

Table 1.3. Computer models developed to describe uterine electrophysiology and propagation and the molecular mechanisms that initiate parturition.

Authors	Model type	Model description	Data source
Sokolowski *et al.* (122)	Uterine physiology: anatomy	A statistical model of uterine wall thickness, volume and tension throughout gestation.	Published intrauterine pressure catheter data and ultrasound measurements of the shape of the uterus in 320 pregnant women.
Bursztyn *et al.* (162)	Uterine physiology: cell excitation	A model of uterine smooth muscle cell excitation with three elements: 1) intracellular calcium concentrations, 2) myosin light chain phosphorylation and 3) stress produced by the cell in response to depolarisation of the cell membrane.	Published results of voltage-clamp experiments using pregnant rat uterine smooth muscle cells, and data on myosin light chain phosphorylation and force production in non-pregnant human uterine smooth muscle cells.
Rihana *et al.* (163)	Uterine physiology: cell excitation	A detailed model involving all the main ionic mechanisms controlling uterine smooth muscle cell contraction.	Published data from voltage-clamp experiments, predominantly on rat uterine smooth muscle cells.
Tong *et al.* (124)	Uterine physiology: cell excitation	A comprehensive model of uterine smooth muscle cell contraction, incorporating 13 types of ionic current and cell size. The model included a reconstruction of cellular changes in response to estradiol.	Published and unpublished data from voltage-clamp experiments on pregnant rat uterine smooth muscle cells in the presence and absence of estradiol, and experimental recordings of spontaneous action potential, calcium ion and phasic force in rat myometrial tissue strips.
Andersen & Barclay (164,165)	Uterine physiology: tissue propagation	A model to explain how contractions spread throughout the uterus, in which the uterus was represented as a hollow ovoid shape composed of discrete contracting cells that send electrical impulses to their neighbours, generate tension, and have defined periods of contraction and refraction	No experimental validation, but it was able to reproduce complex contraction patterns similar to published reports of those observed in human labour.
Young (125)	Uterine physiology: tissue propagation	A tissue propagation model in which intercellular communication occurs via action potential and calcium wave propagation.	One *in vivo* recording of intrauterine pressure during human labour.
Vauge *et al.* (166)	Uterine physiology: tissue propagation	A model that uses physiologically significant parameters to simulate the changes in intrauterine pressure associated with a contraction during labour. The model takes into account the	*In vivo* recordings of intrauterine pressure during human labour.

		number of cells in each of three states (contracting, refractory and resting), the excitability of the cells in the resting state, the lifetime of the contraction state, and the duration of the refractory state.	
Laforet *et al.* (167)	Uterine physiology: multiscale	A preliminary multiscale model that combines a simplification of Rihana's cell excitation model with a model of electrical conductance from the myometrium to the skin. It aimed to describe the pattern of uterine electrical activity from its onset at the cell level, to its propagation throughout the myometrium, to its conduction to the abdominal surface.	Electrohystograms
La Rosa *et al.* (168)	Uterine physiology: multiscale	A multiscale model that combined a simplification of uterine anatomy with a cell excitation model and a tissue propagation model to simulate uterine electrical activity at the abdominal surface.	Published electromyography and magnetomyography data.
Aslanidi *et al.* (123)	Uterine physiology: multiscale	A current multi group effort to build a comprehensive multiscale model of the human uterus while pregnant. The model combines a model of the three-dimensional geometry of the uterus during pregnancy with cell excitation and tissue propagation models.	*In vivo* MRI and ex vivo DTMRI to reconstruct geometry of the uterus, and voltage-clamp experiments.
Wanner & Pliska (126)	Molecular mechanisms	An early model that described an increase in uterine tension in response to an oxytocin stimulus, mediated by intracellular calcium concentration.	Isometric tension of an oxytocin-stimulated rat uterus.
Bisits *et al.* (169)	Molecular mechanisms	Directed acyclic graph (DAG) approach to model the alternative hypotheses of pathways to the initiation of labour: functional progesterone withdrawal, inflammatory stimulation, and oxytocin receptor activation.	Information from the literature about different hypotheses, plus published qRT-PCR data of transcripts of key genes from labouring and non-labouring human myometrium.
Equils *et al.* (121)	Molecular mechanisms	An intracellular model of the immune-endocrine interactions associated with the initiation of labour. They also looked at the behaviour of the model in response to progesterone and COX2-inhibitor treatment.	Existing knowledge about immune-endocrine interactions in the myometrium, but no experimental validation.

1.6.2.1.2 Computational epidemiology

Computational techniques such as those discussed above have been increasingly applied to molecular research over the past decade. However, most epidemiological studies still use traditional statistical approaches, and 'computational epidemiology' is in its infancy (170,171). Traditional approaches rely on researchers generating hypotheses about the epidemiology of a disease, and then applying statistical techniques to test these hypotheses. In contrast, computational epidemiology looks for patterns in unstructured sources of data and serves as a useful means of generating hypotheses (171,172).

Computational epidemiology applies computational concepts to provide tools that aid in the prediction and analysis of disease manifestation and spread. Accordingly, much of the work that has been conducted in this field to date focusses on modelling the spatiotemporal spread of infectious disease throughout populations (171,173). A generic formal disease model (DisMod) has been developed that is applicable to a wide range of non-communicable diseases. DisMod is based on a set of differential equations that describe age-specific incidence, remission, case fatality and 'all other causes' mortality. The model returns predicted prevalence and case-specific mortality as outputs, and allows the impact of different scenarios to be assessed. It is particularly useful for analysing epidemiologic data that are incomplete (174). There are undoubtedly more ways in which computational methods may be useful for analysing epidemiological data.

There are no reports of computational epidemiology being applied to pregnancy research, although it could be a useful, novel approach that may help generate interesting hypotheses about the epidemiology of parturition complications.

1.6.2.2 Combining data from different sources

Data from different sources can be combined to aid analysis. Linking data on the same subjects from different studies can provide additional information that allows the scope of the analysis to be extended. Additionally, combining and meta-analysing data from separate studies that use similar methods can increase the power of analyses and the generalisability of results.

1.6.2.2.1 Record-linkage

Records from clinical databases or epidemiological studies can often be linked to other datasets to increase the breadth of data available for each patient. If the datasets were collected at different times, this allows later life health outcomes to be compared to earlier exposures without using an expensive and difficult to conduct longitudinal cohort design. For example, in Scotland each patient is assigned a unique identifier called a Community Health Identifier (CHI number) at birth or when they register with a general practitioner, which allows record-linkage between different databases of health records. Databases of birth records, such as Scottish Morbidity maternity admissions (SMR02) and the Walker database of 48000 births in Dundee between 1952 and 1966, can be record-linked via the CHI number to databases of other health information such as the Tayside prescriptions database.

To date, only two record-linkage studies have been conducted using the Walker database to investigate the impact of parturition complications on later life health outcomes. Murphy *et al.* (175) investigated the relationship between forceps delivery and epilepsy in adulthood by linking 21,441 Walker births to the Tayside prescriptions database and Scottish Morbidity records hospital admissions database (SMR01). The authors found that delivery by forceps was not associated with epilepsy compared with all other deliveries (adjusted odds ratio = 1.0, 95% confidence interval = 0.6-1.8), but preterm birth was associated with an increased risk of epilepsy (adjusted odds ratio = 2.0, 95% confidence interval = 1.2-3.2). Uma *et al.* (176) conducted a study to determine the influence of intrapartum care during a first delivery on the risk of pelvic floor surgery in later life. Primiparous women from the Walker cohort were linked to surgical data in SMR02. The authors found that delivery by Caesarean section was associated with a reduced risk of pelvic floor surgery in later life compared with spontaneous vaginal delivery (odds ratio = 0.16, 95% confidence interval = 0.05-0.55).

The benefits of carrying out record-linkage studies are clear, and there is scope to investigate the epidemiology of other parturition complications using this approach.

1.6.2.2.2 Meta-analysis

The statistical power and generalisability of results can be improved by combining data from different studies for meta-analysis. This technique is most commonly applied in

systematic reviews of RCTs, but it can also be useful in analysing molecular data, such as those produced by microarray experiments.

RCTs are useful for evaluating the efficacy of health interventions. Subjects are randomly allocated to receive one intervention or another/no interventions. Care is taken to reduce bias in the study design to increase the likelihood that any difference in outcome between the different treatment arms is due to the treatments being compared. However, RCTs are very time-consuming, expensive and difficult to conduct. Consequently, they are often performed using a small number of patients, and RCTs of the same interventions may disagree on their efficacy in different populations. Systematic reviews that attempt to identify and appraise RCTs are helpful in drawing a consensus on the efficacy of interventions. Most systematic reviews combine results from the RCTs under review and use statistical techniques to perform a meta-analysis which gives an idea of the overall effect size (for example, odds ratio or mean difference) of the intervention on specific outcomes, and also assesses heterogeneity. In evidence-based medicine, systematic review and meta-analysis of RCTs is considered the highest level of evidence used to inform clinical decision making. Systematic reviews help to highlight and reduce the impact of selective reporting of results from RCTs (reporting bias). A recent large systematic review of RCTs for women with breast cancer found that 33% of the 164 included trials selectively reported the primary endpoint and 67% selectively reported toxicity (177). Positive effects of the intervention were more likely to be reported than negative effects.

In order to reduce bias and maintain standards in systematic reviewing, precise guidelines on how to conduct systematic reviews are produced by the Cochrane Group, which is a group of health-care specialists that also publishes in their own journal, the Cochrane Library. The Cochrane Pregnancy and Childbirth Group, a subdivision of the Cochrane Group, focuses on systematic reviews of interventions involving the mother or baby during and after pregnancy and childbirth. To date, they have published over 480 reviews.

Less often, similar techniques are applied to meta-analyse molecular data. Individual microarray studies sometimes report findings that are not reproducible or robust to different forms of data analysis. It is often difficult to make direct comparisons between results obtained in different laboratories using different platforms and different methods of analysis. Additionally, small sample sizes in individual studies can lead to less robust results that are difficult to generalise. This problem is magnified when using human tissue

samples, which are highly susceptible to inter-patient variation (178). Meta-analysis of microarray data increases the statistical power to detect differentially expressed genes and allows the variation across studies (herterogeneity) to be assessed. This can result in more accurate, reproducible, and robust predictions (179–182).

There is one published meta-analysis of microarray data to study parturition (113). This study, which was published in 2007 combined lists of genes reported to be differentially expressed in five human studies of myometrial gene expression in pregnancy, term and preterm labour. The authors found that the remodelling and maturation processes associated with pregnancy last the full course of gestation, but genetic regulation of the onset of parturition was less well characterised. However, this study's use of published gene lists rather than raw microarray data ignores information on genes showing lower levels of differential expression, and does not account for the different methods used to generate and analyse the data. Similar meta-analyses using raw data and analytical techniques allowing the studies to be more directly comparable are needed. Initiatives such as MIAME (Minimum Information About a Microarray Experiment), which provide guidelines for the standard reporting of microarray experiments, specify that the raw data must be made available to the research community to allow different groups to analyse the data themselves, potentially using different techniques (183). Insistence on authors complying with such guidelines has greatly facilitated meta-analysis of existing data, and the validation of new analytical methods.

1.7 Conclusion

A limited understanding of the molecular mechanisms initiating human parturition stems from the great complexity of the system, and difficulties in studying it. This seriously hampers attempts to predict, prevent and treat parturition complications. Similarly, these complications are often idiopathic or have an unknown aetiology, and epidemiological studies to study aetiology are expensive and/or time-consuming. These barriers to parturition research can be partly overcome by making efficient use of existing data and by using special analytical methods to maximise the amount of useful information gleaned from it. Such methods include use of data-driven systems approaches, computational modelling, record-linkage and meta-analysis.

1.8 Aims and hypotheses

This work aims to explore the use of special methods in analysing novel and previously published data to study the molecular mechanisms initiating human parturition and the epidemiology of parturition complications. Specifically, the hypotheses are:

- Exploring myometrial microarray data using functional analysis and a network graph approach will reveal genes and molecular pathways associated with labour and other patient characteristics.
- The reliability and generalisability of these findings will be increased through meta-analysis with other, similar myometrial microarray datasets.
- Published knowledge of a signalling pathway relevant to preterm labour can be integrated to build a computational model that successfully simulates *in vitro* behaviour.
- A network graph approach can be used to visualise and explore epidemiological data relating to birth outcomes.
- Data from two birth cohorts can be record-linked to assess if the risk of PPH is higher in women whose mothers had PPH in their pregnancies.
- Systematic review and meta-analysis of RCTs can be useful in assessing the efficacy of fetal assessment following PPROM in improving neonatal and maternal outcomes.

2 METHODOLOGY

This chapter describes and provides a rationale for the techniques used to generate and analyse data throughout this PhD. A table of all computer applications and online tools used is provided in Appendix 1.

2.1 MICROARRAY ANALYSIS

The results of these analyses are discussed in Chapter 3.

Microarrays were first designed to improve our understanding of gene functions and molecular mechanisms (184–186). They are most commonly used to analyse expression of multiple genes simultaneously. Microarrays are particularly useful for comparing gene expression in different biological situations. For this PhD, a microarray experiment was conducted to investigate gene expression in labouring and non-labouring human myometrial tissue. Modern microarray techniques involve tens of thousands of 'probes' (short sections of genes) attached to microscopic beads or a 'chip' made of a solid surface, such as glass. Each probe is complementary to a cDNA or cRNA target transcript. Usually, the target is labelled with a fluorescent dye and then hybridised to the chip. The fluorescence intensities of each spot give a relative measure of transcript expression levels.

There are several microarray platforms available, and the choice of platform can have a large impact on analysis. Affymetrix and Illumina are the most common commercial microarray platforms and each version has a highly standardised protocol that facilitates inter-experiment comparisons. Here, the Illumina HT-12 v4.0 Expression BeadChip platform (Illumina, San Diego, CA, USA) was chosen because it offers the highest multiplexing capabilities for whole genome gene expression, is highly accurate and relatively inexpensive.

2.1.1 MYOMETRIAL TISSUE SAMPLES

Biopsy samples of full thickness lower segment specimens of human myometrium were selected from the Edinburgh Reproductive Tissue Biobank (ERTBB, http://www.crh.ed.ac.uk/biobank/), which is managed by staff within the MRC Centre for Reproductive Health. For each patient, samples are split and processed in two ways – wax

embedded for immunohistochemistry and frozen (with or without RNA*later*®) for RNA and protein extraction. Forty-eight ERTBB frozen myometrium samples were selected for RNA extraction. Forty-five of the samples selected had been transferred into RNA*later*® (Life Technologies, Invitrogen, Carlsbad, CA, USA) immediately after collection and then frozen at -80°C. Three labouring samples had not been transferred to RNA*later*® before freezing. The latter were included in the array analysis in order to maximise the number of labouring samples, and because a previous microarray experiment has shown that fresh or frozen myometrial tissue shows no systematic shift in quantitative RNA expression compared to myometrial tissue stored in RNA*later*® (187). To minimise any effect of this difference in storage, all three frozen samples were washed briefly in RNA*later*® before RNA extraction.

Labouring myometrium is usually collected during emergency Caesarean section. Such women are more difficult to recruit because of the unpredictable nature of the procedure. Therefore, the ERTBB contains fewer labouring myometrial samples compared to non-labouring samples. In order to maximise biological replicates on the array, all 22 labouring samples in the ERTBB as of August 2012 were selected. Non-labouring samples were then selected by loosely matching to labouring samples on certain patient characteristics. An effort was made to ensure mean values for key demographics (maternal BMI, gestational age, parity and maternal age) were similar between labouring and non-labouring groups.

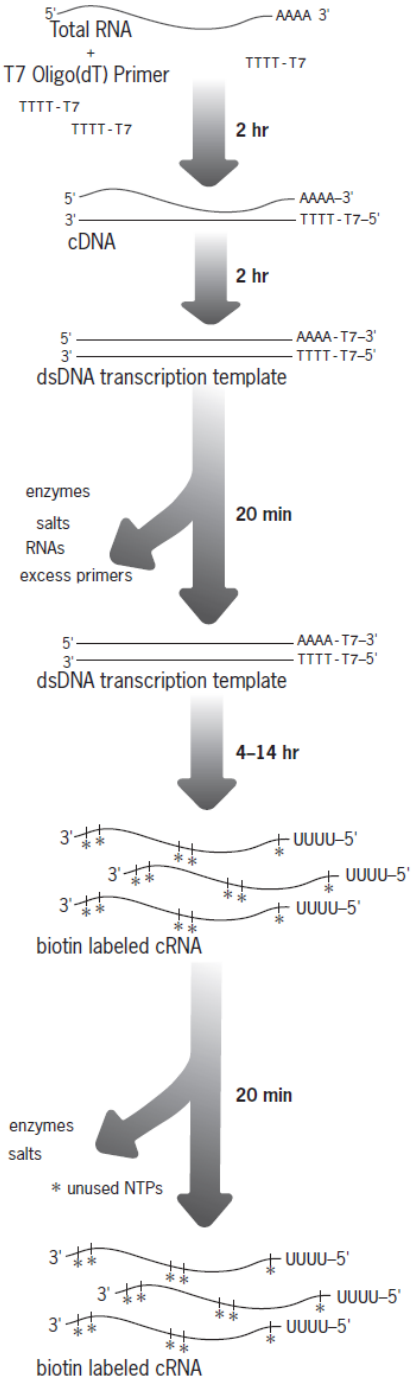
2.1.2 RNA PREPARATION

Total RNA was isolated using TRI Reagent® and the Qiagen RNeasy Lipid Tissue kit protocol (Qiagen, Valencia, CA, USA). This method exploits the ability of nucleic acids to pH hybridise onto silica within an RNeasy column (Qiagen) when tissue is homogenized in a high pH buffer. Briefly, tissue was homogenised in a solution of RNeasy lysis buffer. Chloroform was added and tubes were vigorously shaken before centrifuging for 15 minutes at 12000g and 4°C. The homogenate separates into three phases – the upper aqueous (RNA) phase, a white inter-phase and a lower organic phase. The aqueous phase was transferred to a new tube and ethanol was added. Samples were centrifuged for 15 seconds at 8000g at room temperature on RNeasy mini spin columns. The flow through was discarded and centrifugation was repeated. RW1 buffer was added to the spin columns and then centrifuged for 15 seconds at 78000g at room temperature. The flow through was discarded and the RNA was resuspended in 50µl RNase free water.

The quantity and quality of RNA was assessed using a NanoDrop 1000 (Thermo Scientific, Wilmington, DE, USA). This exposes the sample to ultraviolet light at 260nm and a photo-detector measures the light that passes through. The more light that is absorbed, the higher the concentration of RNA. The ratio of light absorbance at 260nm and 280nm is used to assess the purity of the RNA. A ratio of around 2.0 is considered "pure", but substantially lower readings may indicate contamination.

RNA quality was further assessed in the biotin-labelled samples by the Wellcome Trust Clinical Research Facility using the Bioanalyzer 2100 (Agilent Technologies, Wilmington, DE, USA). The Bioanalyzer reads a gel chip containing up to 12 RNA samples in micro-channels. Nucleic acid fragments within the samples are driven through the gel electrophoretically and this separates the fragments according to their size. The fragments are then detected by laser-induced fluorescence detection. This generates an electropherogram showing the concentration and ribosomal ratio.

To prepare samples for the microarray experiment, 450ng total RNA was amplified and biotin-labelled using the Illumina® TotalPrep™ RNA Amplification Kit (Ambion, Austin, TX, USA). This produces hundreds to thousands of biotinylated antisense RNA copies. Biotin is a small water-soluble B vitamin that is useful for labelling and isolating proteins for biochemical assays. The procedure consists of reverse transcription with a reverse transcriptase enzyme called ArrayScript. This enzyme catalyses the synthesis of virtually full-length cDNA, thus ensuring maximum production of reproducible microarray samples. After undergoing second strand synthesis and cleanup, the cDNA becomes a template for *in vitro* transcription with T7 RNA Polymerase. Ambion® MEGAscript *in vitro* transcription (IVT) technology and biotin-UTP is used to generate multiple biotinylated antisense RNA copies of each mRNA in a sample. Antisense RNA is referred to as cRNA (and sometimes aRNA in the literature). This kit was developed to produce labelled cRNA for hybridisation with Illumina arrays. The protocol is summarised in Figure 2.1.



II.C. Reverse Transcription to Synthesize First Strand cDNA

1. Bring RNA samples to 11 µL with Nuclease-free Water
2. Add 9 µL of Reverse Transcription Master Mix and place at 42°C
3. Incubate for 2 hr at 42°C

II.D. Second Strand cDNA Synthesis

1. Add 80 µL Second Strand Master Mix to each sample
2. Incubate for 2 hr at 16°C

Potential stopping point

II.E. cDNA Purification

Preheat Nuclease-free Water to 55°C
1. Add 250 µL cDNA Binding Buffer to each sample
2. Pass the mixture through a cDNA Filter Cartridge
3. Wash with 500 µL Wash Buffer
4. Elute cDNA with 20 µL of 55°C Nuclease-free Water

Potential stopping point

II.F. In Vitro Transcription to Synthesize cRNA

1. Add 7.5 µL of IVT Master Mix to each cDNA sample, and mix
2. Incubate for 4–14 hr at 37°C
3. Add 75 µL Nuclease-free Water to each sample

Potential stopping point

II.G. cRNA Purification

Before you begin: preheat Nuclease-free Water and assemble filter cartridges and tubes
1. Add 350 µL cRNA Binding Buffer to each sample
2. Add 250 µL 100% ethanol and pipet 3 times to mix
3. Pass samples through a cRNA Filter Cartridge(s)
4. Wash with 650 µL Wash Buffer
5. Elute cRNA with 200 µL 55°C Nuclease-free Water

Potential stopping point

Figure 2.1. A summary of the Illumina® TotalPrep™ RNA Amplification Kit protocol. Figure taken from the Illumina® TotalPrep™ RNA Amplification Kit handbook (Ambion).

RNA extraction and labelling was carried out by myself under the guidance and training of Dr Lawrence Hutchinson.

2.1.3 Illumina HT-12 v4.0 BeadChip Expression microarray platform

The Illumina HT-12 v4 Expression BeadChip is a direct hybridisation whole-genome expression assay used to analyse intact complementary strand RNA samples. The set-up is slightly different to that of an Affymetrix microarray. Oligonucleotide probes are immobilized to beads held in microwells on the surface of an array substrate (Figure 2.2). The beads are randomly distributed across the surface of the substrate. Each bead contains a 29 nucleotides long (29-mer) address sequence, used to identify the location of the bead and validate its hybridisation. Biotin-labelled sample cRNA is hybridised to 50-mer probes on the BeadChip, and after washing and staining with a fluorescent protein that binds to biotin the BeadChips are scanned. There are 47231 probes in total, with 28,688 of those coding a transcript with well-established annotation. Probes were designed by Illumina using the HumanRef-8 beadchip, with additional probe designs based on UniGene release 188. There are 12 microwells on each chip (Figure 2.2), which allows up to 12 samples to be processed on the same chip. Not only does this multi-sample format increase throughput, but it also reduces sample-sample variability.

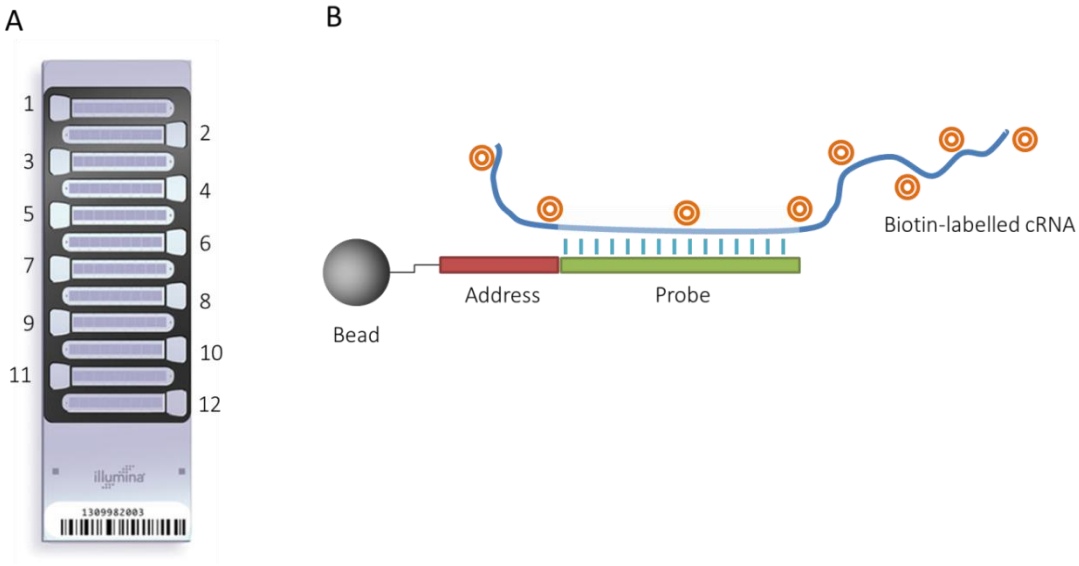


Figure 2.2. Illumina ht-12 v4.0 beadchip expression array technology. A) the chip contains 12 microwells that are each loaded with a single sample. Each microwell contains oligonucleotides immobilized to beads. B) each bead is attached to a 29-mer address sequence and a 50-mer probe. Biotin-labelled cRNA is hybridised to the probe (redrawn from the Illumina literature).

The Wellcome Trust Clinical Research Facility (WTCRF) at the Western General Hospital in Edinburgh conducted the microarray experiment and scanned the chips. Samples were split randomly over the four chips to minimise any effect of inter-chip variability interfering with differential expression between labouring and non-labouring groups. The chips were imaged using a BeadArray Reader and raw data was obtained with Illumina BeadStudio software. The raw gene expression data was returned as tab delimited files.

2.1.4 Data preprocessing

Raw data was preprocessed by the Cardiovascular Sciences (CVS) Bioinformatics team using the Lumi package (188–191) in R. They subjected the data to Robust Multichip Average (RMA) background correction before quantile normalisation to remove non-biological systematic variation. Background correction removes signal caused by nonspecific binding of the fluorescent protein. Quantile normalisation is used to remove any 'batch effect' due to sample-sample and chip-chip variation (192). Quantile normalisation is a widely used standard method, based on the assumption that the distributions of expression values are similar in every sample. For each sample, the probes are arranged in ascending order of expression value. These arranged values are then averaged over all the samples. The original expression values for the features for each sample are then replaced by their mean over all samples. This brings expression values into the same register so that meaningful comparisons can be made. CVS Bioinformatics provided box plots and intensity plots to visualise the distribution of expression values across all probes and all samples before and after normalisation. Large variations may indicate poor quality RNA.

2.1.5 Annotation

Also using the Lumi package, CVS Bioinformatics annotated the data according to the Illumina probe ID. They incorporated information on the gene name, official gene symbol, Unigene ID, Entrez gene ID, and Gene Ontology (GO) terms.

CVS Bioinformatics returned the normalised data as comma delimited files in which each row represents a probe with its ID in the first column and annotation information in subsequent columns followed by expression values for each sample.

2.1.6 Background cut-off determination

Most of the probes in an array are not expressed in most experiments but have low 'background' level expression values. Removing these probes is important because they may mask any true changes in gene expression between groups (for example, it is not uncommon for an unexpressed gene to appear to be differentially expressed between groups because its low expression value varies slightly but the fold-change appears significant).

The median expression value over all probes and all samples can be taken as the cut-off value used to define background expression. Another method to define this cut-off is to take the mean expression value of negative control probes included on the array. These negative control probes contain sequences not expected to hybridise to most genomes, so they provide a measure of non-specific hybridisation, non-specific dye signal and scanner background.

Here, the mean expression value over all probes was 168.78 (median=64.05) and the mean expression value of negative control probes was 81.4. The final background cut-off value was rounded up to 100, as advised by the CVS Bioinformatics team. Probes for which the labouring and non-labouring group mean expression values were both below background (≤100) were removed. All further analysis was performed using this background-removed data.

2.1.7 Determining sample relations

To assess the similarity of expression values between samples and sample groups (labouring/non-labouring), a correlation matrix was created using the cor() function in R (see Appendix 3). The correlation values in the matrix are computed using the Pearson's method and describe the between sample expression value similarity on a scale of 0-1 (0=perfectly uncorrelated, 1=perfectly correlated). The minimum and mean correlation values between all samples, and between labouring and non-labouring sample groups were observed to assess whether or not there are differences between sample groups.

To help with the interpretation of this correlation matrix, the data were also visualised using two methods. Firstly, the matrix was used to draw a heatmap using the heatmap.2() function in the gplots R package (193) (see Appendix 3). The heatmap represents the

strength of each correlation value using colours on a scale. Secondly, the matrix was loaded into BioLayout *Express*^{3D} as a .matrix file to create a network graph where each sample is represented as a node (sphere) connected to others by 'edges' weighted according to the strength of the correlation. All correlation values above 0.95 were used to draw the graph. As a rule of thumb, a correlation of at least 0.95 is usually expected from samples that are biological replicates. Markov Chain clustering (MCL) was performed to assess how the samples cluster into groups (inflation value=7). Biolayout Express3D and MCL clustering are discussed in more detail below in section 2.1.10.

Additionally, the nodes were coloured by key patient characteristics (labouring/non-labouring, BMI group, preterm/term, maternal age category and para=0/para≥1) to visualise the effect of these characteristics on between-sample correlation. The association between cluster membership and membership to these patient characteristic groups was assessed using Fisher's exact or χ^2 tests. Another network graph was built in the same way but using only labouring samples. MCL clustering was performed in the same way and nodes were coloured according to whether or not the patient was administered syntocinon/prostaglandins and the rough stage of labour (assessed as 'early' if cervical dilation was ≤3cm, and 'late' if cervical dilation was ≥6cm). Again, the association between MCL cluster and these categories was assessed using Fisher's exact or χ^2 tests.

2.1.8 Assessing differential expression between groups

CVS Bioinformatics used R package Limma (194) (part of Bioconductor) to provide a means through which I could assess differential expression between groups. I requested that two comparisons could be made: labouring versus non-labouring samples (in line with the aims of the experiment), and para=0 versus para≥1 samples (chosen after further analysis of results). Limma calculates the fold-change in mean expression between groups for each probe. It is reported as a ratio where a value of 1 indicates no change, 2 indicates a doubling in expression and -2 indicates a halving in expression. Fold change values are in relation to the base level group (non-labouring or para=0). Confidence intervals for the fold changes were calculated using Fieller's theorem (195) via an R function (see Appendix 3). Limma was also used to perform statistical testing between the groups. The method relies on linear modelling and empirical Bayesian methods to give modified t-test p-values for group comparisons for each probe. Because of the large number of t-tests being performed, there is a high risk of false-positive significant values. Therefore, an adjusted

P-value is calculated that corrects for the false discovery rate. Smallest p-values (most significant) are associated with probes that have similar expression values within a group, but a clear difference between groups. Absolute P-values should be interpreted cautiously because they describe not only the biological significance, but also the sum of the biological effects and other sources of variability or noise in the data. Therefore, CVS Bioinformatics also reported the significance of differential expression using a non-parametric statistical method called Rank Product (196). This method is based on the ranks of fold-changes rather than individual values, which reduces the chance of false negative results. This is particularly useful in noisy data. It is corrected for multiple testing using the percentage of false positives (PFP) and reported as an RP-PFP value. This value is roughly equivalent to a P-value, so values ≤0.05 can be considered statistically significant.

CVS Bioinformatics built data into a filterable database that could be used to filter differentially expressed genes by specifying fold-change and adjusted-P-value/RP-PFP criteria. In general, I used fold-changes >1.2 or <-1.2 and RP-PFP ≤0.05 to define differential expression. There is no 'correct' way to select a fold-change/significance level cut-off, so these liberal values were chosen in an attempt to maximise true positive findings.

Differentially expressed genes were also visualised using a heatmap. The heatmap represents expression values as colours on a colour scale. Columns represent samples and rows represent differentially expressed genes. To make the list small enough to enable clear visualisation, a fold change cut-off of >2 or <-2 and an RP-PFP cut-off of ≤0.0001 were used. A hierarchical clustering algorithm is used to sort samples according to the similarity of their expression values for all the genes included in the map. The heatmap shows 1) which genes are up/down regulated in which samples, and 2) how samples cluster.

2.1.9 Resolution of the many-to-many issue between probes and genes

Where multiple probes exist for the same gene, searches were carried out over all probes. To investigate single genes in more detail, the gene was 'summarised' by selecting the probe with the smallest adjusted P-value for its fold change, because it is the least likely to occur by chance. This summary method was assumed to be more accurate than simply

taking an average of the probe expression values (which would be inaccurate because different probe sequences have different binding affinities and therefore different measurement scales) or selecting a probe at random (which would lose information and risk under/overestimating the degree of differential expression).

2.1.10 NETWORK GRAPH ANALYSIS

Further analysis and visualisation of the data was conducted using a network graph approach. Biolayout *Express*^{3D} was used to draw a weighted, non-directed graph in which probes are represented as 'nodes' (spheres) connected by 'edges' (lines) weighted according to similarity in expression value. Representing data visually in this way makes it easier for humans to infer relationships and identify structural features such as clusters, it also allows interactive analysis of the data. The approach uses no prior assumptions about the likely structure within the data or the questions to be addressed.

Biolayout Express3D uses the Pearson correlation coefficient as a measure of similarity between expression profiles. An expression profile is the expression of each probe over the range of samples analysed. Pairwise pearson correlation coefficients are calculated for every probe on an array and correlations beyond a threshold are used to draw edges between nodes (probes). It is necessary to impose a correlation threshold to avoid generating an overly large graph dominated by uninformative edges.

Probes with similar expression profiles appear closer to each other in the graph, creating clusters of probes. These clusters are expected to be biologically significant, containing genes with similar biological functions or involved in similar mechanisms. Biolayout Express3D uses the Markov cluster algorithm (MCL) to define clusters within the graph. The MCL algorithm is appropriate because it avoids incorrectly defining clusters in data that contains a lot of spurious (false positive) correlations. MCL clustering is partly dependent on the "inflation value" used to run the algorithm. Lower inflation values result in fewer, larger, noisier clusters and higher inflation values result in many, smaller, cleaner clusters.

Data was loaded to Biolayout Express 3D as a tab delimited .expression file. A range of Pearson correlation coefficient thresholds and MCL inflation values were explored. A correlation cut off of 0.8 and an MCL inflation value of 2.2 were determined empirically to be optimal to fragment the data into meaningful clusters.

2.1.11 FUNCTIONAL ANALYSIS

Functional analysis aims to infer biological meaning from expression data. DAVID (The Database for Annotation, Visualization and Integrated Discovery) (197,198) is a popular web-based tool used to analyse gene lists according to different functional categories and pathways. The Functional Classification Tool within DAVID assesses whether certain functional categories appear more frequently in the gene list than over the rest of the genome. It uses over 75000 terms from 14 functional annotation sources to generate generates a gene-to-gene similarity matrix. The clustering algorithm classifies highly related genes into functionally related groups and assigns each group an 'enrichment score' to gauge how important the group is in the overall list.

DAVID offers a large choice of functional annotation categories. A widely used category is Gene Ontology (GO) (199). One of the aims of the GO project is to develop and maintain standardised annotations for genes and gene products. The GO terms cover three domains: 1) cellular component, 2) molecular function, and 3) biological process. DAVID's Functional Annotation Clustering tool groups similar annotations together to make the biology clearer and reduce the redundancy of multiple similar annotations.

KEGG (Kyoto Encyclopedia of Genes and Genomes) Pathways (200,201) is also useful for annotating gene lists in DAVID. KEGG Pathways is a database of canonical molecular pathways. Gene lists can be tested for enrichment to certain KEGG pathways.

In this analysis, lists of official gene symbols representing 1) down-regulated genes (fold change <-1.2 and RP-PFP ≤0.05), 2) up-regulated genes (fold change >1.2 and RP-PFP ≤0.05), and 3) MCL cluster specific genes were uploaded to DAVID. The lists were assessed for biological process GO term enrichment and KEGG pathway enrichment. DAVID integrates with KEGG to help visualise gene expression data by mapping it onto KEGG pathway diagrams.

2.1.12 IMMUNOHISTOCHEMISTRY

Immunohistochemistry refers to the process of localising antigens in a tissue section through binding to a specific primary antibody. This technique was performed to detect the leukocyte marker CD45 in labouring and non-labouring myometrial tissue samples.

The same technique was also applied to a human tonsil sample, acting as a positive control known to contain an abundance of leukocytes.

2.1.12.1 SAMPLES

Of the patients used in this microarray experiment, five labouring and five non-labouring tissue samples were chosen at random (patients were first allocated and then sorted by a random number, the top five patients in each category were chosen). The non-labouring samples were NL3095, NL2275, NL2158, NL2410 and NL3114. The labouring samples were L3189, L0190, L3132, L3197 and L3226.

2.1.12.2 EMBEDDING AND SLIDE CUTTING

At the time of collection, myometrial tissue samples from the ERTBB were fixed immediately in 4% NBF for 24 hours and transferred to 70% ethanol. Tissue was subsequently wax embedded using standard protocols.

Tissue sections were cut at 5μM sections and adhered to glass slides by flotation in warm water. Slides were then incubated at 37°C for at least 24 hours to ensure complete adherence to the slide. This work was performed by the SuRF@QMRI Histology department.

2.1.12.3 STAINING FOR CD45

The CD45 antibody was supplied by Dako (CD45, Leucocyte common antibody Clone 2B11+PD7/26; Santa Clara, CA, USA) and was used at a dilution of 1:1000. Slides were dewaxed in xylene for 5 minutes, rehydrated in ethanol (absolute, 90%, 80% and then 70%), and washed in water. Antigen retrieval was performed using Novocastra pH6 retrieval buffer (Leica microsystems, Wetzlar, Germany). Staining was carried out using the Bond-max automated immunostaining machine (Leica microsystems, Wetzlar, Germany), via a service supplied by the SuRF@QMRI Histology department.

Slides were washed five times in Bond Wash Solution for a total of 15 minutes. Following this, endogenous peroxidase activity was blocked by incubating the slides for 10 minutes in 3% hydrogen peroxide in PBS. Slides were washed in Bond Wash Solution for two minutes prior to incubating the sections for 2 hours at room temperature in primary antibody diluted in 1 part normal goat serum to 4 parts bond wash buffer. Slides were

then washed in Bond Wash buffer for 10 minutes before incubation with the pre-polymer for 15 minutes at room temperature. Following further washes slides were incubated with an HRP polymer. Binding of the antibody was visualised by incubating in 3,3'diaminobenzidine (DAB) substrate for peroxidase for a maximum of 5 minutes, resulting in positive brown staining. Slides were then counterstained in haematoxylin, dehydrated in increasing concentrations of ethanol, incubated in xylene for 5 minutes and coverslips were mounted using pertex mounting medium.

A negative control, where the primary antibody was replaced with normal goat serum/Bond wash buffer, was included in each run.

Images of stained sections were obtained using the Olympus PROVIS microscope (Shinjuku, Tokyo, Japan) and AxioVision Rel 4.8 software (Carl Zeiss microscopy, Oberkochen, Germany).

Figure 2.3 shows positive staining for CD45 in the positive control (human tonsil) tissue and no staining for CD45 in myometrial tissue incubated without the antibody.

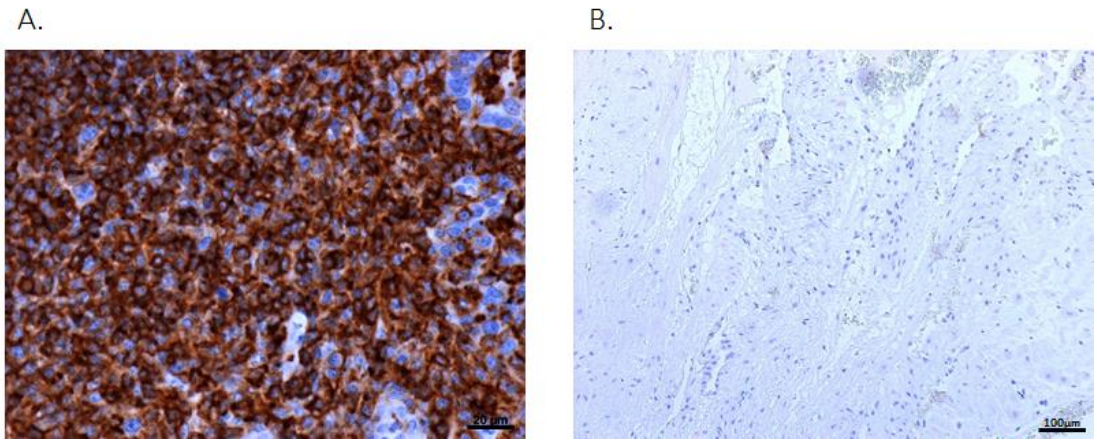


Figure 2.3. Immunohistochemistry controls. A. Positive control, human tonsil stained for CD45. B. Negative control, human myometrial tissue incubated without the antibody.

2.1.13 Validation with TaqMan real-time PCR

Microarrays are an excellent tool for initial target discovery, but the relatively limited dynamic range of fluorescent microarrays limits the sensitivity and specificity of the technology. It is therefore essential to validate differential expression of some genes using an alternative method.

TaqMan Gene Expression Assays were chosen to validate this microarray because they have a high specificity, sensitivity and large dynamic range. A TaqMan probe is labelled with a fluorescent reporter dye at the 5' end and a second 'quencher' dye at the 3' end, which quenches light emitted from the reporter dye. The probe hybridises to the target DNA sequence between specific forward and reverse primers. During the amplification phase of PCR, the Taq polymerase enzyme moves from the 5' to 3' end, thereby cleaving the probe and releasing the reporter dye from the proximity of the quencher. The reporter dye emits light when excited by a laser, and the amount of fluorescence is measured as directly proportional to the amount of target present in the PCR.

2.1.13.1 SAMPLES

qRT-PCR was performed using the original RNA samples used in the microarray experiment. Generally, more reliable validation is achieved using samples from different patients, however, this was impossible because such samples were unavailable. The large number of samples (48) used in this experiment means that this is less of a limitation than it would be for smaller experiments.

2.1.13.2 THE RELATIVE QUANTIFICATION STANDARD CURVE METHOD

Relative gene expressions were determined using the standard curve method. Relative quantification is based on the expression levels of a target gene versus a reference gene and is considered appropriate in this scenario. The standard curve uses multiple serial dilutions of sample to plot a curve from which relative concentrations can be read. The advantage of this method over other relative quantitative RT-PCR methods is that it requires no assumptions about amplification efficiency. Expression values of target genes from the standard curve are then normalised against those from a reference gene.

Here, the standard curve method was used to evaluate the relative expression of levels of target genes between labouring (test) samples and non-labouring (calibrator) samples.

2.1.13.3 cDNA SYNTHESIS

cDNA synthesis was performed by Nanette Hibbert.

1µg of total RNA of each sample was reverse transcribed to cDNA using a cDNA synthesis kit (4368813-Applied Biosystems, Carlsbad, US). A "mastermix" was prepared according

to Table 2.1. For each sample, 1µg (10µl) of total RNA was added to 10µl of mastermix. A negative control containing 4.2µl RNase free water instead of the reverse transcriptase was included to measure residual genomic DNA, and another negative control containing 13.2µl of RNase free water instead of the total RNA was included to assess contamination. Transcription was performed using a G-Storm GS1 thermal cycler (G-Storm, Somerton, Somerset, UK) with a cycle of 25˚C for 10 minutes, 37˚c for 120 minutes and 85˚C for 5 minutes. Samples were stored at -20˚C until needed.

Table 2.1 cDNA synthesis mastermix.

cDNA synthesis mastermix component	Volume/sample (µl)
10x Reverse transcriptase buffer	2
25x dNTPs	0.8
10x random primers	2
Reverse transcriptase	1
RNase inhibitors (1unit/µl)	1
RNase free water	3.2
Total	***10***

2.1.13.4 HOUSEKEEPING AND TARGET GENES

Expression of the target genes was normalised to the amount of the housekeeping gene, 18s ribosomal RNA. 18s expression was found to be consistent between labouring and non-labouring myometrial samples, and its use as a housekeeping gene in these tissues is well described in the literature.

Pre-designed, inventoried TaqMan gene expression assays (Applied Biosystems, Foster City, CA, USA) were used to measure the expression of ten genes. These assays are optimised for use in quantitative real time polymerase chain reaction (qRT-PCR). The gene names, corresponding assay IDs and the reasons for choosing these genes are given in Table 2.2.

Table 2.2. Information about the gene assays used to validate the microarray.

Gene ID	Gene name	TaqMan assay ID	Reason gene chosen for validation
RBM42	RNA binding motif protein 42	Hs00225667_m1	Not differentially expressed according to the array.
SHROOM4	Shroom family member 4	Hs00393349_m1	Not differentially expressed according to the array.
FABP4	Fatty acid binding protein 4	Hs01086177_m1	Downregulated in labour according to the array.
IGFBP5	Insulin-like growth factor binding protein 2	Hs00181213_m1	Downregulated in labour according to the array.
MYH11	Myosin heavy chain 11	Hs00224610_m1	Downregulated in labour according to the literature. Smooth muscle marker.
TPM1	Tropomyosin 1	Hs00165966_m1	Downregulated in labour according to the literature. Smooth muscle marker.
IL6	Interleukin 6	Hs00985639_m1	Upregulated in labour according to the array and the literature.
IL8	Interleukin 8	Hs00174103_m1	Upregulated in labour according to the array and the literature.
MT1E	Metallothionein 1E	Hs01938284_g1	Upregulated in labour according to the array and the literature.
OXTR	Oxytocin receptor	Hs00168573_m1	Upregulated in labour according to the literature.
18s (RN18s1)	18S ribosomal RNA 1	Hs03928985_g1	Housekeeping gene classically used in previous studies comparing labouring vs. non-labouring human myometrium.

2.1.13.5 qRT-PCR

qRT-PCR was conducted by Nanette Hibbert and myself (under Nanette's guidance).

qRT-PCR was performed using an ABI 7900HT (Applied Biosystems, Carlsbad, US), on a 384 well plate with inventoried gene-specific Taqman assays (Table 2.2). The reaction mix is shown in Table 2.3.

Table 2.3. qRT-PCR mastermix.

qRT-PCR mastermix component	Volume/reaction (µl)
TaqMan mastermix	5
Primers and probes	0.5
RNase-free water	2.5
Total	*8*

The standard curve was obtained by diluting a pool of equal volumes of all labouring and non-labouring samples (dilutions: neat, 1:2, 1:4, 1:8, 1:16, 1:32, 1:64, 1:128, 1:256). All samples and RT negative controls were diluted 1:10 using RNase-free water. This ensures that the expression values correspond with the linear part of the standard curve, which

allows relative values to be compared more accurately. For each sample/point of the standard curve, 8µl of mastermix was added to 2µl (diluted) cDNA. All standard curve points, RT negative and water controls were assayed in technical triplicate for each gene.All non-labouring and labouring samples were assayed in technical duplicate. All targets and controls were run in separate wells. The reaction involved three stages, the third of which was repeated 40 times: 1) incubation at 50˚C for 2 minutes, 2) incubation at 95 ˚C for 15 seconds, and 3) incubation at 95 ˚C for 15 seconds, then at 60˚C for 1 minute.

Where the mean cycle threshold (CT) values of technical duplicates had a standard deviation over 0.2, the reaction was repeated. If this did not resolve the difference, the results were included in further analysis but the implications of this were noted in the discussion of the results.

2.1.13.6 DATA ANALYSIS

All data analysis was carried out by myself.

For each target gene and 18s, a nine-point standard curve was plotted of the log total cDNA (ng) against the raw CT (cycle threshold) value. For each dilution, there were three technical replicates. Figure 2.4 shows an example of a standard curve used in this study. The slope of the standard curve is used to estimate PCR amplification efficiency. A slope of -3.32 indicates a reaction with 100% efficiency. Slopes more negative than -3.32 are less than 100% efficient and slopes more positive than -3.32 may indicate problems with pipetting or sample quality. Slopes should be between -3.3 and -3.6 (90-100% efficient). Regression analysis was used to draw a regression line for the data and to calculate its slope and Y intercept. These values were used to determine the quantity of cDNA generated for each sample by the PCR reaction. For each target gene and 18s, the log total cDNA (ng) was interpolated by dividing the sample's CT value minus the intercept, divided by the slope.

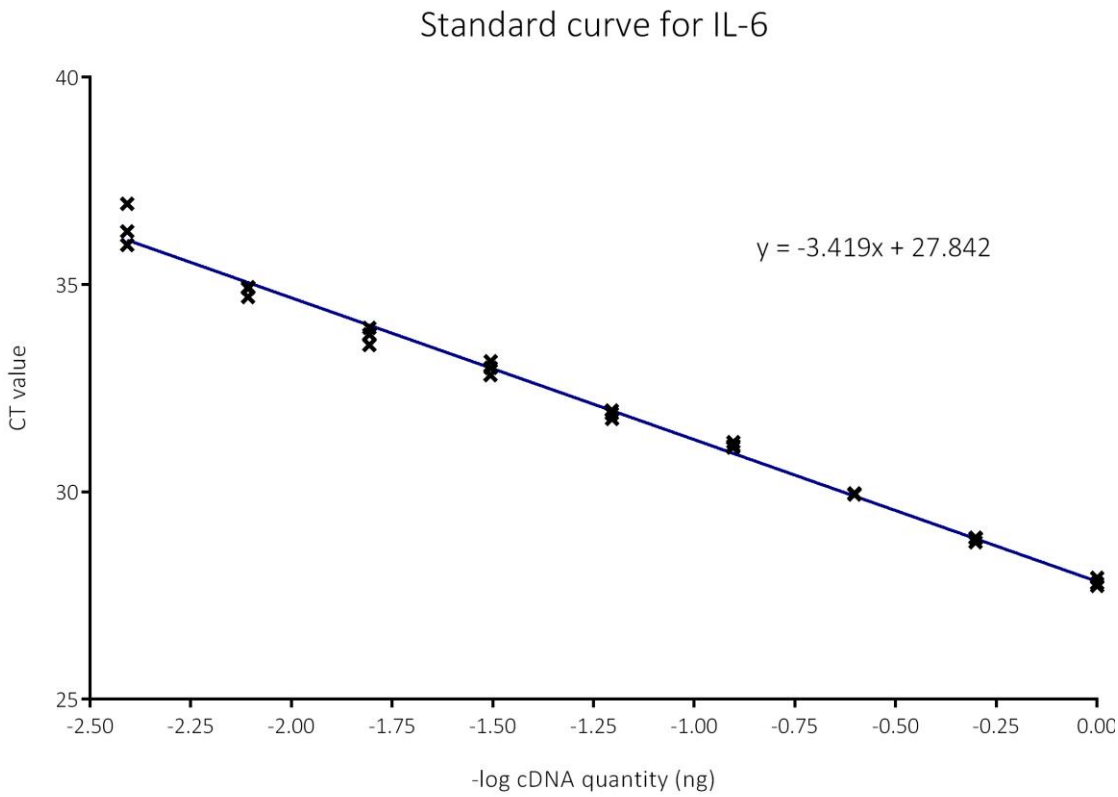


Figure 2.4. An example of a standard curve used to interpolate cDNA concentrations from CT values.

The cDNA quantity mean and standard deviation of the technical duplicates was calculated, followed by the cDNA quantity mean and standard deviation of the biological replicates (i.e. all labouring samples and all non-labouring samples). The mean values for target genes were divided by the mean values for 18s to give normalised cDNA quantities of target genes. The standard deviations of these normalised values were calculated using the coefficients of variation. The fold change in normalised target gene cDNA quantity from non-labouring to labouring samples was calculated by dividing the normalised target gene cDNA quantity for labouring samples by the normalised target gene cDNA quantity for non-labouring samples. The standard deviation of this fold change was calculated using the coefficient of variation for labouring samples.

2.1.14 MICROARRAY META-ANALYSIS

In addition to the wet-lab validation of this microarray, a meta-analysis was also conducted to assess how the results compare to the published results of similar microarrays. Ramasamy *et al.* (178) provided useful guidelines for the conduct of this

meta-analysis. The chosen meta-analysis technique was similar to that used by Hu *et al*.(182). The microarray experiment described in this thesis is referred to as "Sharp 2013" for the purposes of the meta-analysis.

2.1.14.1 SEARCH STRATEGY

A systematic approach was used to search the literature to identify all eligible studies and minimise the effect of bias that would arise from selectively choosing which studies to include. The online repositories of abstracts PubMed (http://www.pubmed.gov/), Google Scholar (http://scholar.google.com/), Web of Science (http://wos.mimas.ac.uk) and SCOPUS (http://www.scopus.com), and the microarray data repositories Array Express (http://www.ebi.ac.uk/arrayexpress/) and Gene Expression Omnibus (GEO)(http://www.ncbi.nlm.nih.gov/geo) were searched using the following search terms: "myometrium" and "microarray", "uterine smooth muscle" and "microarray", "myometrium", "myometrial", "uterine smooth muscle".

2.1.14.2 INCLUSION-EXCLUSION CRITERIA

Any human studies investigating myometrial gene expression in labouring or non-labouring pregnant women were included in the meta-analysis. Criteria for excluding studies were as follows:

- No human samples.
- No myometrium samples.
- Myometrium collected from a location other than the lower uterine segment.
- No pregnant samples.
- No opportunity to compare labouring to non-labouring myometrial samples.
- Study does not report the results of a microarray experiment, or is a review.
- Samples were prepared from cells in culture rather than tissue.
- Tissue samples were cultured and/or treated before RNA extraction.

2.1.14.3 DATA IDENTIFICATION AND ACQUISITION

Where the raw data had not been made available via an online repository, authors were contacted. Despite this effort, the data for only three studies (Bukowski 2006 (55), O'Brien 2008 (119), Weiner 2010 (115)) could be included in the meta-analysis, along with Sharp 2013. O'Brien provided normalised data on request via email. The Bukowski (accession number E-MEXP-106) and Weiner (accession number GSE9159) raw data were downloaded from ArrayExpress.

2.1.14.4 PREPROCESSING

Where possible, arrays were preprocessed consistently to remove any systematic differences. The CVS Bioinformatics team subjected the Sharp 2013, Weiner 2010 and Bukowski 2006 arrays individually to RMA background correction before quantile normalisation to remove non-biological variation. Due to the small sample size, the O'Brien 2008 array did not undergo consistent preprocessing. Instead, the preprocessed data (method not reported) supplied by the authors was analysed.

2.1.14.5 ANNOTATION MATCHING

The CVS Bioinformatics team used the annotation packages in BioConductor to map probe IDs to Entrez Gene IDs, Ensemble IDs, UniGene IDs, UniProt IDs, Gene symbols, gene descriptions and GO Terms, as described for Sharp 2013 above.

2.1.14.6 DETERMINING BACKGROUND CUT-OFF FOR EACH STUDY

The background expression level for each study was determined by comparing the study-specific expression of certain genes to their expression in Sharp 2013, which acted as a reference. The background level in Sharp 2013 was set at 100 (explained above), so all genes with a mean expression (over all samples) of between 99.9 and 100.1 in Sharp 2013 were selected as "at background level". Official Gene IDs of these genes were used to search datasets from the other studies. Six of these genes appeared on all four arrays. The mean expression values of these genes were roughly rounded to give the background level for each array. Figure 2.5 shows a plot of the expression values of each of these genes for each study and indicates that there is not much variation between the "at background" genes. The graph includes two genes expected to have low expression (DCX, double cortin, which is known to be expressed exclusively in brain tissue, and LIPC, hepatic lipase, which is known to be expressed in the liver and adrenals), and two genes expected to have high expression (TAGLN, transgelin, which is expressed in fibroblasts and smooth muscle, and HBB, beta globin, which is a constituent of the most common form of haemoglobin in adult humans). These genes show the range of expression values exhibited in each array, and fall the expected side of the background cut-off. For each array, the background cut-off levels were set at:

- Bukowski 2006 = 5

- O'Brien 2008 = 10,000
- Weiner 2010 = 150
- (Sharp=100).

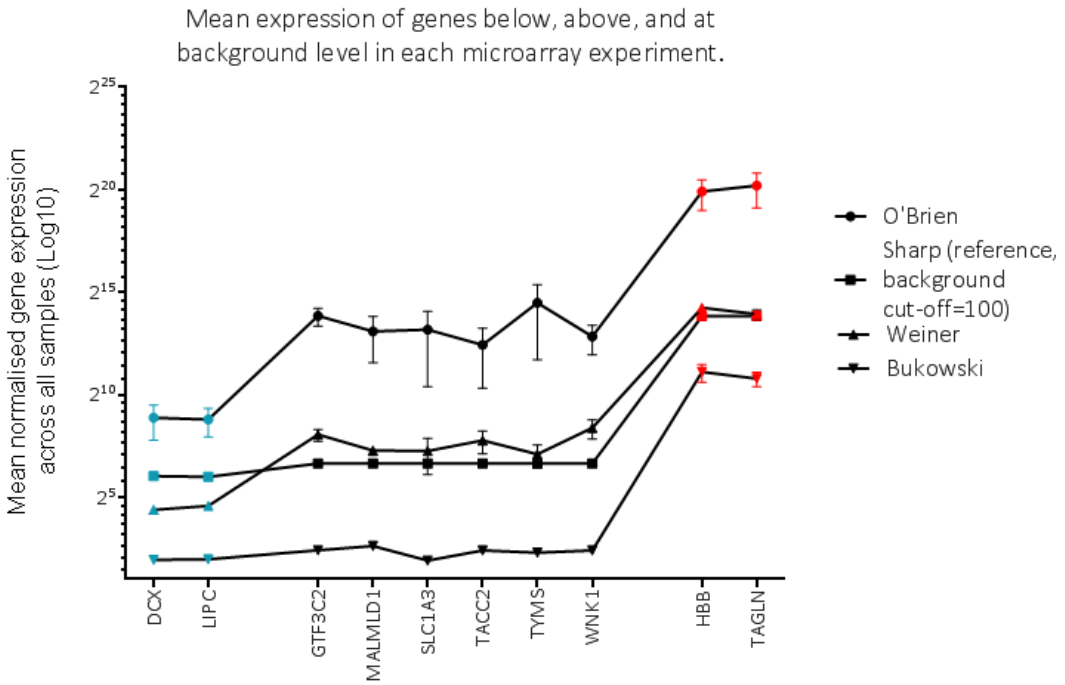


Figure 2.5. Determining the background expression cut-off for each study. A plot of each study's expression values for two genes not expected to be expressed (DCX, LIPC), six genes expected to be expressed at around background level (GTF3C2, MALMLD1, SLC1A3, TACC2, TYMS, WNK1) and two genes expected to show high expression (HBB, TAGLN), compared to their expression in Sharp 2013.

2.1.14.7 Assessing differential expression between groups for individual and combinations of studies

For each study individually, CVS Bioinformatics used R package Limma to assess differential expression between groups, as described above for Sharp 2013. CVS Bioinformatics built this data into a filterable database that could be used to filter differentially expressed genes by specifying fold-change and adjusted-P-value/RP-PFP criteria for each study individually. Searches can be combined to give lists of genes for which criteria is true across studies, for example, a search could identify all genes upregulated > 1.2 with a P-value of <0.05 in Sharp 2013 and Weiner 2010.

I identified differentially expressed genes by filtering out genes below study-specific background cut-offs in both labouring and non-labouring samples, and using an arbitrary but liberal fold-change cut-off of >+/<- 1.2 and a rp-pfp value <0.05.

Lists of differentially expressed genes were produced for each study individually and for agreed genes in each combination of studies (i.e. Sharp 2013 and Weiner 2010, Sharp 2013 and Bukowski 2006, Weiner 2010 and Bukowski 2006, and Sharp 2013, Weiner 2010 and Bukowski 2006). O'Brien 2008 could not be included in these searches because no RP-PFP value could be calculated for this study, and all adjusted P-values were non-significant (>0.05).

To visualise the number of differentially expressed genes in each study and how well studies agree, a proportional Venn diagram was created using an online tool at http://apps.bioinforx.com/bxaf6/tools/app_venndiagram.php. This is described as a valid technique in Ramasamy et al.(178).

2.1.14.8 FUNCTIONAL ANALYSIS

Functional analysis was performed on lists of official gene symbols representing 1) down-regulated genes (fold change <-1.2 and RP-PFP ≤0.05) and 2) up-regulated genes (fold change >1.2 and RP-PFP ≤0.05) for each study individually. The same technique was also applied to lists of genes each combination of studies agreed to be differentially expressed. Again, these analyses could not be conducted using the O'Brien data. Gene lists were uploaded to DAVID and assessed for biological process GO term enrichment and KEGG pathway enrichment.

Lists of GO terms found to be associated with labour by each study were also compared using a proportional Venn diagram, as described above.

2.1.14.9 RESOLUTION OF THE MANY-TO-MANY ISSUE BETWEEN PROBES AND GENES

Where multiple probes exist for the same gene, searches were carried out over all probes. To investigate single genes in more detail, genes were 'summarised' as described above for Sharp 2013.

2.1.14.10 META-ANALYSIS TECHNIQUE

In order to understand whether or not individual studies agreed on the direction (up regulated or down regulated) of the difference in expression between labouring and non-labouring samples, and to give an idea of heterogeneity between studies, Forrest plots of the effect sizes were created. Forrest plots list the names of the studies on the left-hand side and plot the measure of the effect (for example, a fold-change) as a square with confidence intervals on the right. The area of each square is proportional to the weight of the study in the meta-analysis. The overall 'summary' meta-analysed measure of effect is also plotted, often plotted as a diamond with the lateral point indicating the confidence intervals. A vertical line is plotted to represent the point of no effect. Forrest plots were drawn for each gene that was agreed to be significantly differentially expressed in Sharp 2013, Weiner 2010 and Bukowski 2006. The O'Brien 2008 results were also included on the Forrest plots. Forrest plots were also drawn for genes that have previously been associated with labour, and genes used to validate Sharp 2013 by q-PCR.

Two approaches to create the forrest plots were tried. Firstly, GraphPad Prism was used to plot the study-specific effect sizes as the $\log^2$ ratio of the mean expression values for labouring and non-labouring samples. Confidence intervals for the ratios were conducted using Fieller's theorem (195) and by writing an R function to do the calculation (see Appendix 3). The overall 'summary' effect size was calculated as the mean of the other study-specific effect sizes, and the confidence intervals for these were calculated as the 95% confidence intervals of that mean. This approach provided easy to interpret plots, but there was no way to change the size of the square to indicate the weight of the study, and the summary effect size was not adjusted for the weight of the studies. Additionally, when the confidence interval of the denominator (non-labouring) mean includes zero, the confidence interval of the mean ratio could not be calculated.

Instead, a more robust approach was taken utilising Review Manager, which was developed specifically for performing Cochrane systematic reviews and meta-analyses. Study-specific effect sizes were plotted as the standardised mean difference between labouring and non-labouring samples. The standardised mean difference was used because all the studies assess the same outcome (differential expression) but assess it in a variety of ways (using different platforms, leading to different expression value scales). The results of each study must be standardised to a uniform scale before they can be

combined. A random effect model was then used to integrate the study-specific effect sizes and give an overall summary measure of effect that takes into the weight of the study into account. A random effects model was used rather than a fixed effects model to allow for both within-experiment sampling error (variance) and between-studies variation to be included in the calculation of the confidence interval. To assess heterogeneity, Review Manager also reports the results of a Chi^2 test (to test whether the observed differences are compatible with chance alone) along with the Tau^2 statistic (an estimate of the between-study variance in the random-effects meta-analysis). If the Chi^2 statistic is significant at $P<0.1$ then there is likely to be substantial heterogeneity. Similarly if Tau^2 is more than 1 there is likely to be substantial statistical heterogeneity.

2.2 IN SILICO MODEL

The results of these analyses are discussed in Chapter 4.

In silico models translate biological systems into mathematical representations that can be used to simulate the behaviour of that system. Models can be built at several levels using various approaches. For this PhD, an *in silico* model of a key signalling pathway in preterm birth (lipopolysaccharide (LPS)-induced activation of Nuclear Factor kappa B (NF-κB)) was built using published data and an ordinary differential equation (ODE) approach.

2.2.1 CHOICE OF MODELLING TARGET

When modelling a system as complex as human parturition, it is especially important to clearly define, and limit, the scope for the model. So rather than attempt to model all the elements of a system concurrently, it is more sensible to focus on a single aspect, such as a functionally relevant intracellular signalling pathway, and model that accurately before building in more elements (202). Intracellular signalling pathways allow cells to respond to extracellular stimuli in order to perform their functions and survive. They are particularly appropriate targets for computational models, because they are relatively simple (in comparison to systems at the tissue or organ level for example) and they can be observed and manipulated *in vitro*. Models of such pathways are increasingly appearing in the literature thanks to recent advances in the quality, quantity and availability of modelling software and experimentally-derived parameter data (202).

Here, to maximise the likelihood that sufficient published data would be available, a well characterised pathway was chosen to model: lipopolysaccharide (LPS)-induced translocation of nuclear factor kappa B (NF-κB) p65-p50 heterodimer to the nucleus. The reasons for choosing this pathway are discussed in more detail in Chapter 4.

2.2.2 Choice of modelling approach

Dynamic models that describe the behaviour of the system over time are most useful for studying signalling pathways (202), and the most well-established way to do this is using deterministic models based on ordinary differential equations (ODEs) (203). ODEs have a well-established biophysical basis and straightforward molecular interpretation. ODE models are very comprehensive because they include all the known molecular entities in a system and quantitatively describe the kinetics of each physical interaction between them. This increases the likelihood that the model will allow us to fully understand and manipulate the complex behaviour of the system (203).

Furthermore, ODE models are highly detailed and require specific information about initial concentrations of reactants and rates of reactions (kinetic parameters). Therefore attempting to build a kinetic ODE model using solely published data is a challenge and a good test of the current level of knowledge and data accessibility.

An approach using partial differential equations (PDEs) was also considered, because this approach often allows more in-depth kinetic analysis. However, the PDE approach requires even more specific data (for example, on spatial distribution of molecules within cells). Therefore, due to limited data availability, we favoured an ODE approach.

2.2.3 Literature search

Comprehensive searches of the current literature are necessary to provide specific information on the structure of the pathway, its expected behaviour and reaction kinetic parameters such as initial concentrations of reactants and reaction rates. Searching databases of abstract repositories can be frustrating because papers often depict signalling pathways using simplified diagrams, with little information about the specific interactions between the components of the pathway. Curated pathway databases are more useful for gathering specific information, particularly annotated databases that link to the primary literature.

For this PhD, searches were conducted of the abstract repository PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and the curated pathway databases KEGG (http://www.genome.jp/kegg/), Nature Pathway Interaction Database (http://pid.nci.nih.gov/) and Reactome (http://www.reactome.org/). Previously described computational models of NF-κB activation in scenarios outwith pregnancy(204–206) were also useful for finding reaction kinetics and were accessed via Biomodels (http://www.ebi.ac.uk/biomodels-main/), the online database of peer-reviewed published models. Figure 2.6 shows terms that were used to conduct these searches.

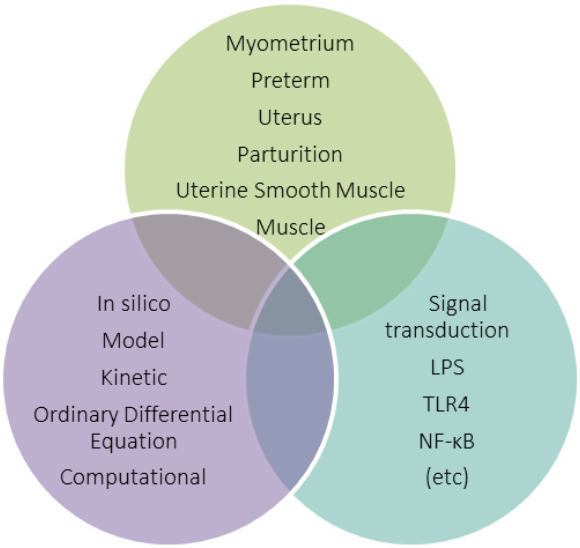


Figure 2.6. Search terms can be broadly grouped into three categories: 1) parturition-related terms, 2) pathway components and related terms, and 3) computational biology terms. Searches were carried out with individual search terms and various combinations.

2.2.4 Graphical representation

Graphical representations of all the reactions in a model are useful when building or interpreting the model. They depict every entity (for example a protein or a complex), the reactions they are involved in, and in what way they react (binding, phosphorylation, etc.). The graphical depiction acts as a blueprint to the static structure of the model.

It is helpful to use a standard graphical notation that can be interpreted precisely and unambiguously by scientists from different disciplines. Examples include Systems Biology Graphical Notation (SBGN)(207) and modified Edinburgh Pathway Notation (mEPN) (208), which have both been developed over several years by biochemists, modellers and

computer scientists. These notations use a set of symbols to represent the different components, reactions, relationships and cellular compartments in the pathway. A range of tools exists to build and edit topological maps of pathways. SBGN maps can be built using dedicated pathway editing tools such as CellDesigner (209) and mEPN maps can be loaded using general network and visualisation software such as yED (210). Both approaches were used in this PhD. Reference keys for SBGN (process description) and mEPN are shown in Appendix 4.

2.2.5 Model development

A graphical representation is useful for visualising a pathway, but to turn it into a quantitative computational model differential equations must be generated to describe each reaction. There are many tools designed to help researchers create, simulate and analyse computational models. Copasi (211) is a particularly intuitive, purpose-built tool that is freely available for academic use. It can import and export models in systems biology mark-up language (SBML), which is particularly beneficial because SBML is a universal programming language that is software-independent. SBML models built in Copasi are compatible with different programs and can be made available to other researchers via online repositories (160).

Reaction equations are generated to describe each reaction in the topological map. These follow a simple format where a binding reaction could be represented by:

"SUBSTRATE A" + "SUBSTRATE B" = "PRODUCT C"

- Where "SUBSTRATE A" binds to "SUBSTRATE B" to produce a complex called "PRODUCT C". The reaction is reversible, as indicated by the equals sign.

An enzymatic reaction could be represented by:

"SUBSTRATE A" -> "PRODUCT C" ; "ENZYME D"

- Where "PRODUCT C" is converted from "SUBSTRATE A" via a reaction catalysed by "ENZYME D". Reaction modifiers are separated from the main reaction by a semi-colon. The reaction is irreversible, as indicated by the unidirectional arrow, "->".

Entering these reaction equations into Copasi generates a list of all the necessary 'species' (components of the pathway). Each reaction is automatically assigned an ordinary differential equation to describe its rate. As a default this is generated according to mass-action kinetics, but this can be changed to one of the default rate equations in Copasi (such as Michaelis-Menten, which is more commonly used to describe enzymatic reactions) or a custom, user-defined rate equation.

Copasi was used to develop the LPS–NF-κB pathway model described in this PhD. The structure of the IκB-NF-κB signalling portion of the model is largely based on a high quality model developed by Hoffmann *et al.* (204). All other portions of the model were based on information derived from the literature. The final model was developed through successive rounds of model building and simulation within Copasi. Model development was performed by myself, with guidance from Dr Hongwu Ma.

Throughout the modelling process, model refinements were made to ensure the model is accurate but simple:

- Initially the formation of the receptor complexes was modelled using separate reactions. For example:
    - `LPS + LBP = LPS:LBP`
    - `LPS:LBP + CD14 = LPS:LBP:CD14`
    - `LPS:LBP:CD14 + TLR4 = LPS:LBP:CD14:TLR4`
    - `LPS:LBP:CD14:TLR4 + TIRAP = LPS:LBP:CD14:TLR4:TIRAP`
    - `LPS:LBP:CD14:TLR4:TIRAP + Myd88 = LPS:LBP:CD14:TLR4:TIRAP:MyD88`
    - `LPS:LBP:CD14:TLR4:TIRAP:MyD88 + IRAK4 = LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4`
    - `LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4 + TRAF6 = LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4:TRAF6`

    This represents the order of pathway events described in the literature. However, these are a linear series of mass-action reactions, which are now modelled more appropriately as a single mass action reaction:

    - `LPS + LBP + CD14 + TLR4 + TIRAP + MyD88 + IRAK4 + TRAF6 = LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4:TRAF6`

- All the enzymatic phosphorylation reactions are reversible, meaning that the dephosphorylation step is incorporated in the Michaelis-Menten ODE. The reference to the phosphatase is omitted because its total concentration can be considered constant and, unlike the kinase, it does not need to be activated by the pathway. Furthermore, less is known about phosphatase concentrations, specificities and kinetic parameters than is known about the kinase enzymes, so incorporating more detailed information about the dephosphorylation step would be difficult.
- Iterative rounds of model development and simulation (described below) were completed until the model was able to recapitulate the timecourse of active NF-κB reported in the literature (205). For example, early versions of the model showed the TRIF-dependent pathway was having more of an effect on IKK phosphorylation and downstream NF-κB activation than the MyD88-dependent pathway. This is the opposite of what is reported in the literature. The problem was traced to the TRIF-dependent pathway producing very high levels of TNFα, which meant none of the reactions downstream of TNFα were rate-limited. Therefore, we altered the values of parameters involved in the production of TNFα.

2.2.6 Model simulation

When all the reactions have been added, the model's behaviour can be analysed. The stability of the model can be ascertained by its ability to reach a steady-state (at which components are no longer changing over time). The behaviour of the components of the model over time can be analysed by calculating a time course. Plots of the time course trajectory are useful for analysing model behaviour. If the model is not behaving as expected it may be necessary to deconstruct the model to find which parameters to amend.

The final model was produced after successive rounds of simulation and model parameter adjustments to fit published data. To run the final model, initial concentrations are first updated to reflect the steady-state concentrations with LPS = 0 μmol/ml. This represents the initial steady-state of the cell before treatment. LPS treatment is then simulated by running a time course simulation with LPS arbitrarily set at 1 μmol/ml. The model was run for 250 minutes (model time). This produces graphs showing the concentration of each non-fixed model species over time.

The behaviour of the model was compared to published behaviour of the system *in vitro*. An attempt was made to test the ability of the *in silico* model to recapitulate the *in vitro* behaviour of the system when either of two proteins (Myeloid Differentiation Primary Response Gene 88 (MyD88) and Tir-Domain-Containing Adapter-Inducing Interferon-β (TRIF)) in the pathway are knocked out. The *in vitro* behaviour is published in Covert *et al*. (205) 'Knock-out' models were created by fixing the concentration of either of the proteins to 0 (thus removing their contribution) and running simulations as described for the full model. The model was developed under "wild-type" conditions , and tested under "wild-type", MyD88- and TRIF- "knock-out" conditions. Several "tests" and subsequent changes to the wild-type model were made during model development.

2.3 A NETWORK GRAPH OF EPIDEMIOLOGICAL DATA

The results of these analyses are discussed in Chapter 5.

As mentioned above, representing large datasets visually as network graphs makes it easier for humans to infer relationships and identify structural patterns. Network graphs also allow interactive analysis of the data and provide an unbiased insight into how data points cluster. The network graph tool Biolayout *Express*^{3D} was first developed to visualise and analyse gene expression data; its use for this purpose is described above. In theory however, the tool can be used with any large datasets to explore complex systems of interaction. These include patient health record databases to explore interactions between different health characteristics and disease outcomes.

For this PhD, a network graph approach to study epidemiological data was conducted for the first time. Using birth data collected for the Walker cohort, several options for using this approach were explored.

2.3.1 THE WALKER COHORT

The Walker cohort (212) is a dataset of 48985 birth records that contains meticulously recorded details of pregnancy, labour and care before discharge for births in hospital in Dundee between 1952 and 1966. These babies accounted for 75% of all births in Dundee at this time (212).

Data was collected by Professor James Walker and his senior obstetric colleagues on purpose-written record cards according to a predefined dataset. Three types of card were

used (Card 0 for 1952-58, Card 1 for 1959-63, and Card 2 for 1964-66), with an increasing number of details being recorded on each.

The written data from the cards were computerised between 2000 and 2001. The date of birth and sex of each baby was used to identify the Community Health Number (CHI number), which is a unique personal identifier assigned to all people registered with a general practitioner in Scotland. These CHI numbers allow linkage of Walker data to demographic data such as postcode, Scottish Index of Multiple Deprivation score, and date of death.

2.3.2 Data preparation

2.3.2.1 Raw data preparation

The Health Informatics Centre (HIC) at the University of Dundee stores and maintains the Walker data. They provided access to the data via a secure online environment. The data was provided as a separate comma delimited file for each Walker card, in which each row represented a delivery, usually identified by the CHI number of the baby. Card 0 contained 20683 deliveries, Card 1 contained 15132 deliveries and Card 2 contained 13170 deliveries. All delivery records were unique (i.e. there were no erroneous duplicates). CHI numbers were used to record link to demographic and additional birth information provided in separate comma delimited files. Records were linked using the =lookup() function in Microsoft Excel.

All three cards with additional variables obtained through linkage were merged to create one spreadsheet containing all 49585 deliveries. The variables present in this file are listed in Appendix 5.

2.3.2.2 Data cleaning

Large datasets often contain a proportion of inaccurate datapoints, introduced through human error either at data collection or computerisation. Most of the Walker data is binary (1=characteristic is present, 0=characteristic is absent), so inaccurate data is difficult to identify if it is simply recorded as the opposite value.

Some of the binary variables listed in Appendix 5 were derived from continuous variables (see below) that first had to be 'cleaned' to identify and remove inaccurate data. If not

removed, this data will introduce bias and distort the results of analyses. A useful, objective way to identify these data points is to use a statistical method to detect 'outliers' – measurements that significantly deviate from what one would expect based on the rest of the data.

Here, outliers were usually defined as any value falling outside of a certain percentage of the data. This was considered to be a more accurate method than using the standard deviation or variance to define outliers, because these statistics are calculated using the entire data set and are therefore "dirtied" to some degree by the outliers themselves, whereas the median and percentiles are more robust to errors in the data (213).

Before defining outliers, all 0 values were deleted. Table 2.4 lists the method of outlier identification for each continuous variable.

Table 2.4. The method of outlier identification for each continuous variable.

Variable name	Outlier definition	Comment
Crown heel length (cm)	<1st percentile (<33.02) and >99th% (>58.105). 19 values removed. 438 values removed.	Variable seemed particularly noisy so a smaller range was used to identify non-outliers.
Crown rump length (cm)	<1st percentile (<24.8) and >99th% (>58.42). 168 values removed.	Variable seemed particularly noisy so a smaller range was used to identify non-outliers.
Gestation (weeks)	<23 and >44 weeks. 10 values removed.	Determined using knowledge of the usual range of gestational ages at delivery.
Head circumference (cm)	<1st percentile (<27.5) and >99th% (>39.3). 233 values removed.	Variable seemed particularly noisy so a smaller range was used to identify non-outliers.
Birth weight (g)	<0.1 percentile (<623.7) and >99.9% (>5274.6). 95 values removed.	
Discharge weight (g)	<0.1 percentile (<2012.8) and >99.9% (>4918.3). 19 values removed.	
Paternal age	<13 and >72. 9 values removed.	Determined using an estimate of the usual range of paternal ages.
Maternal age	<13 and >60. 17 values removed.	Determined using an estimate of the usual range of maternal ages.
Paternal height (cm)	<0.1 percentile (<149.86) and >99.9% (>194.74). 18 values removed.	
Maternal height (cm)	<0.1 percentile (<135.255) and >99.9% (>180.34). 44 values removed.	
Placenta weight (g)	<0.1 percentile (<141.75) and >99.9% (>3316.9). 18 values removed.	

2.3.2.3 Converting continuous data to binary

Variables containing imperial measurements were first converted to metric. Continuous variables were then converted to binary variables according to ranges of values. For example, information in the continuous variable "B_BIRTHWEIGHT_G" (birth weight in grams) was used to create two discrete variables "B_BIRTHWEIGHT_OVER4KG" (birth weight over 4kg) and "B_BIRTHWEIGHT_UNDER2.5kg" (birth weight under 2.5kg). This discretising of continuous variables was necessary to ensure all variables were in the same register. If not converted, a continuous variable would have a disproportionately high influence on the overall topology of the network graph.

The methods used to create binary variables from continuous variables are explained in the table in Appendix 5.

2.3.3 Obstetric characteristic graph

A graph was produced in which each node represented a Walker variable (i.e. an obstetric characteristic such as small placenta weight or extreme preterm birth). The aim of this was to produce a graph to visualise how variables are associated with each other, and explore how informative MCL clusters might be in generating hypotheses about these relationships.

To create this graph, the prepared Walker data (48985 births, 148 variables) was loaded into R and converted to a correlation matrix using the cor() function (Appendix 3). The correlation matrix contains Pearson correlation values for each pairwise comparison of variables. The matrix was exported from R to MS Excel where missing values (where it was impossible to calculate a Pearson correlation value due to there being no cases with information on both variables) were assigned a correlation coefficient of 0. It was also necessary to convert negative values to 0 to avoid crashing Biolayout *Express*^{3D}. Pairwise comparisons for which all correlated values were 0 were deleted, leaving 108 variables. The matrix was saved as a .matrix file and loaded into Biolayout *Express*^{3D}. Graphs were inspected visually at varying correlation coefficient threshold levels to assess their informativeness. Eventually a threshold value of 0.2 was chosen. Any correlation values below this were not incorporated into the graph. This produced a graph in which the structure could be viewed easily and the number of meaningful (higher) correlations was maximised while less meaningful (lower) correlations were not included.

The MCL algorithm was run with an inflation value of 2.2 and a minimum cluster size of 3 to define the boundaries of clusters in an objective, unbiased manner. Other values were explored, but these produced the cleanest, most informative clusters.

2.3.4 Delivery episode graph

Larger graphs were also produced in which each node represented a different delivery episode (case). The aim of this was to produce graphs to visualise how deliveries cluster according to certain similar characteristics. Again, this allows exploration of how variables are associated with each other.

Data was loaded into Biolayout Express3D as .expression files, which cannot contain any missing values. In .expression files, each row represents a different delivery episode and each column contains data on a different variable. Missing values cannot simply be replaced with zeros because this would skew the binary data. Replacing all 0 values (characteristic absent) with -1 and changing all missing values to 0 also skewed the data. Therefore, because some variables were unique to certain cards, missing values were minimised by splitting the data by card number. For each card, complete variables were retained and incomplete and/or less informative variables were formatted as "class sets", whereby the information can be overlaid onto the graph by colouring nodes by class, but the information does not influence the graph topology. Data for each card was saved as a separate .expression file and loaded separately into Biolayout Express3D.

A threshold value of 0.8 was set. Any correlations below this value were not incorporated into the graph. The MCL algorithm was run with an inflation value of 2.2 and a minimum cluster size of 3. Again, graphs were explored at a variety of threshold, inflation and minimum cluster size values, but these values provided the most informative, clear graphs.

2.4 THE EPIDEMIOLOGY OF POSTPARTUM HAEMORRHAGE

The results of these analyses are discussed in Chapter 6.

The historical nature of the Walker cohort allows birth information to be linked with a large number of current health-outcome databases and allows linkage of records across siblings and over generations. For this PhD, the Walker cohort was linked to a current database of maternity information, SMR02, to investigate intergenerational transmission of postpartum haemorrhage.

2.4.1 THE COMMUNITY HEALTH IDENTIFIER (CHI)

Since the 1970's, people living in Scotland have been allocated a unique Community Health Identification (CHI) number, which is used on all their health records. This allows record linkage across different health databases and can be used to track the health of relatives across generations.

2.4.2 The Walker cohort

The Walker cohort is described above in section 2.31. 34183 (73%) of Walker babies can be identified through their CHI number, and this presents the opportunity to link this birth information with a large number of current health-outcome datasets covering both primary and secondary care.

2.4.3 Scottish morbidity records – maternity admissions (SMR02)

The Information Services Division (ISD) Scotland provides SMR02 data to HIC.

The Scottish Morbidity Records Maternity Admissions dataset (SMR02) contains detailed information on hospital maternity admissions in Tayside, Scotland collected from January 1975 to present.

2.4.4 Data extraction and record linkage

The Health Informatics Centre (HIC) at the University of Dundee provided anonymised data from Walker and SMR02 in separate delimited text files. I then used the CHI number to record-link between the two datasets using the merge function in SPSS.

From Walker, HIC extracted data on the births of all Walker babies who could be identified in SMR02 as having had a child themselves. Each Walker record was identified by the baby's CHI, with 65.5% (n=2532) also having information about the mother's CHI. From SMR02, HIC extracted all records that could be linked to the extracted Walker records using the CHI number. Each birth was identified by the mother's CHI, with 70% (n=15043) of births also having information about the baby's CHI. They extracted all records matching the CHI number of the Walker babies, and records matching the CHI number of the daughters of the Walker babies.

HIC also provided a separate delimited text file containing demographic information for each CHI number. This file was also record linked to the birth information.

2.4.5 The generations

I identified data on three generations of women, defined as follows:

- Generation 1 - Walker Mothers – Women who appear in the Walker cohort as mothers.

- Generation 2 - SMR02 Mothers – Women who appear in the Walker cohort as babies, and the SMR02 cohort as mothers.
- Generation 3 - SMR02 Daughters – Women who appear in the SMR02 cohort as babies, and also as mothers if they have had children themselves.

2.4.6 Classification of postpartum haemorrhage

In the Walker dataset, there is information about postpartum haemorrhage (PPH) stored as a dichotomous variable (no PPH = 0, PPH = 1) for the earliest Walker births only (born 1952-1958). For later Walker births, there exists a variable called "RET_PLAC_PPH". Unfortunately no further information was available to define this variable. It could be interpreted as indicating the occurrence of PPH associated with retained placenta, however the prevalence would have been much higher than that reported in the literature and in SMR02. Because of this ambiguity, this variable was not used to classify PPH. This greatly reduced the number of records that could be included in the analyses.

In SMR02, outcomes are coded according to the International Classification of Diseases (ICD9 and ICD10). Table 2.5 shows the codes used to indicate PPH as an outcome. These were used to create a dichotomous variable to indicate PPH in the same way as in the Walker cohort.

Table 2.5. International Classification of Diseases (ICD) 9 and 10 codes used to identify postpartum haemorrhage (PPH) in SMR02 birth records.

Cause of PPH	ICD-9 code	ICD-10 code
Third stage (associated with retained, trapped or adherent placenta)	666.0	O72.0
Atonic (after placenta delivery)	666.1	O72.1
Delayed and secondary PPH (associated with retained portions of placenta)	666.2	O72.2
Coagulation defects	666.3	O72.3

2.4.7 Data cleaning

It was necessary to prepare the data before analysis to remove records and datapoints that would otherwise have introduced bias. All data preparation was performed using SPSS version 19 (IBM).

In SMR02, admissions that did not result in delivery of a child (for example, abortions or early miscarriages) were removed. Stillbirths of a baby >500g were included. Deliveries

where no further information about the delivery was provided were also excluded, because it was impossible to determine whether or not the delivery was associated with PPH. Variables were checked, and unlikely values were changed to missing values (for example, missing data was often coded as 99, and therefore at risk of being analysed as a real value). For continuous variables, any value falling outside a 'sensible' range based on clinical knowledge. Outliers were changed to missing values.

In Walker, 13 duplicate baby CHIs with differing birth information were identified and deleted. Variables containing imperial measurements were converted to metric. For continuous variables, again any data under the 1% percentile or over the 99th percentile were changed to missing values.

2.4.8 Converting continuous variables to categorical variables

Continuous variables were converted to categorical variables to aid interpretation of results. This was performed to create the following categorical variables:

- PPH in a previous pregnancy – calculated by identifying all deliveries by the same mother and using the date of birth of the child and the PPH status for previous deliveries.
- High birth weight – all birth weights recorded as ≥4kg = yes (1), all birth weights recorded as <4kg but not <2.5kg = no (0). <2.5kg = blank.
- Low birth weight – all birth weights recorded as <2.5kg = yes (1), all birth weights recorded as ≥2.5kg but not ≥4kg = no (0). ≥4kg = blank.
- Maternal age under 20-years-old – all maternal ages recorded as <20 = yes(1), all maternal ages recorded as ≥20 but not ≥40 = no (0). ≥40 = blank.
- Maternal age over 40-years-old – all maternal ages recorded as ≥40 = yes(1), all maternal ages recorded as <40 but not <20 = no (0). <20 = blank.
- Nulliparity – all deliveries where parity is 0 = yes (1), all deliveries where parity is >0 = no (0).
- Preterm birth – all deliveries where gestation is ≤37 weeks = yes (1), all deliveries where gestation is >37 weeks but not ≥42 weeks = no (0). ≥42 = blank.
- Post-term birth – all deliveries where gestation is ≥42 weeks = yes (1), all deliveries where gestation is <42 weeks but not ≤37 weeks = no (0). ≤37 = blank.

- Caesarean delivery – all deliveries via Caesarean section = yes (1), all normal, vaginal deliveries = 0, all other modes of delivery = blank.
- Instrumental delivery – all deliveries by ventouse or forceps = yes (1), all normal, vaginal deliveries = 0, all other modes of delivery = blank.

2.4.9 Data analysis

2.4.9.1 The odds ratio

Estimates of the size of the effect of exposure variables on occurrence of PPH were reported as odds ratios (OR), which describe the ratio of odds of PPH given exposure status. The precision (an indication of the degree of random error in the estimate) of these ORs was indicated using 95% confidence intervals.

The OR was favoured over the relative risk (RR) because the OR is deducible from the R output of multivariate logistic regression models.

2.4.9.2 Unlinked data

Obstetric and demographic risk factors for PPH were assessed in the Walker and SMR02 datasets, separately and as a pooled dataset.

IBM SPSS Statistics version 19 (IBM Corp., Armonk NY), was used to perform a univariate analysis of the data. Unadjusted ORs were calculated to assess the effects of each of: PPH in a previous pregnancy, multiple pregnancy, high birth weight (using the World Health Organisation (WHO) definition of ≥4kg), low birth weight (using the WHO definition of <2.5kg), maternal age under 20-years-old (previously identified as a risk factor for PPH (64)), maternal age over 40-years-old (previously identified as a risk factor for PPH (64)), parity, preterm birth (using the WHO definition of ≤37 weeks' gestation), post-term birth (using the WHO definition of ≥42 weeks' gestation), delivery by Caesarean section, instrumental delivery and smoking status on risk of PPH in the index pregnancy. 95% confidence intervals were reported in the output from SPSS.

To calculate adjusted ORs, factors identified as significant in the univariate analyses were built into a multivariate logistic regression model. The function glm() in the R package lme4 (214) (see Appendix 3) was used to run these models. The ORs were calculated as the exponential of the fixed effects coefficients reported in the R output, and 95%

confidence intervals were calculated as the exponential of these coefficients ± 1.96 multiplied by the standard error (see Appendix 3).

Results from the separate and pooled analyses were compared for differences that might indicate a cohort effect.

2.4.9.3 LINKED DATA

To assess intergenerational transmission of PPH, the CHI number was used to link records across generations as follows:

- Generation 1 was linked to Generation 2
- Generation 2 was linked to Generation 3
- Generation 1 was linked to Generation 3
- Pooled mother-daughter analysis: mothers from Generations 1 and 2 were linked to daughters in Generations 2 and 3.

For each of these comparisons, unadjusted ORs were calculated through logistic regression models using glm() in R package lme4 (214). These models assess the relationship between PPH in the younger generation (the dependent variable) and PPH in the older generation (the independent variable) without adjusting for any other potential covariates.

A particular challenge was presented when statistically analysing this data, because the mother-daughter units were not independent of each other. The same women can appear in different mother-daughter/grandmother-granddaughter pairs (for example a woman could be a mother in the comparison of generations 2 and 3, but a daughter in the comparison in generations 1 and 2. Additionally, one woman could be a mother to more than one daughter). It is important to adjust for this non-independence, because non independence invalidates the assumptions of many statistical tests and can introduce bias that can mask exposure effects. Therefore, generalised linear mixed models (GLMMs) were built using the glmer() in R package lme4 (214). These models incorporated any other covariates found to be significantly associated with PPH in the univariate analyses, and also adjusted for the 'random effects' introduced through non-independence of the mother-daughter units. These random effects were included by building younger and

older generation CHI numbers into the model. The models follow a binomial distribution with a logit link because the presence of PPH is a dichotomous variable.

Again, the ORs were calculated as the exponential of the fixed effects coefficients reported in the R output, and 95% confidence intervals were calculated as the exponential of these coefficients ± 1.96* the standard error.

Professor Graham Horgan of Biomathematics and Biostatistics Scotland, and Daniel Ayoubkhani of the UK Office for National Statistics provided advice during model design.

2.4.10 Power analysis

Where low power was suspected, the fe.mdor() function in the R package clinfun (215) was used to calculate the smallest effect size (OR) that the analyses would have been able to detect at 80% power based on the actual sample sizes. This is more correct than carrying out a 'post-hoc' power calculation using results generated through the analysis. Such analyses are misleading – if an effect is not significant, then by definition the power to detect that effect is very low. No new information is gained from carrying out the power analysis (216).

2.5 A Cochrane systematic review

The results of these analyses are discussed in Chapter 7.

The Cochrane Collaboration Pregnancy and Childbirth Group publishes systematic reviews of health interventions and diagnostic tests involving the mother or baby during and after pregnancy and childbirth. For this PhD, I carried out a Cochrane systematic review of fetal assessment methods to improve neonatal and maternal outcomes following preterm prelabour rupture of membranes (PPROM).

2.5.1 Cochrane workflow

The Cochrane Collaboration have precise guidelines on the process of preparing and maintaining Cochrane systematic reviews. These help to standardise reviews to make them easier to interpret and compare. The guidelines also act to minimise bias and improve the reliability of the review findings. Authors wishing to conduct a review are asked to choose from a list of available titles. They then complete a title registration form (Appendix 6), outlining the background, interventions and outcomes of the proposed

review, and giving details of the team of authors. There must be at least two authors, at least one of whom should be experienced at writing Cochrane reviews. For this PhD, the review team consisted of myself, and two experienced Cochrane reviewers, Dr Sarah Stock and Professor Jane Norman. Considering the expertise of this team and the other data presented in this thesis, the title "Fetal assessment methods for improving neonatal and maternal outcomes in preterm prelabour rupture of membranes" was chosen from the list as the most appropriate title to address.

2.5.1.1 Author contributions

I wrote the title registration form, protocol and review. I also performed all data extraction and analysis. In accordance with the Cochrane guidelines, Sarah Stock also independently extracted data, and I assessed our results for any differences. Sarah Stock and Jane Norman provided a clinical perspective and revised the protocol and review before submission.

2.5.2 Protocol

In order to minimise the potential for bias in the review process, Cochrane require authors to produce a detailed protocol prior to conducting the review. This protocol outlines the methods to be used, which promotes transparency of methods and processes and reduces the impact of review authors' bias. The protocol undergoes peer review and is published in the Cochrane Library. Any changes to the protocol must be justified in the final review. The published protocol(217) for the review carried out here is available at http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD010209/full. The methods used to conduct the review are outlined below.

2.5.3 Criteria for considering studies for this review

2.5.3.1 Types of studies

All relevant published and unpublished randomised controlled trials (RCTs), quasi-RCTs and cluster-randomised trials were included. Crossover trials were excluded because this is an unsuitable design to study fetal assessment during pregnancy. It was planned that abstracts would be included if sufficient details were available and authors of abstracts would be contacted to obtain further information where necessary.

2.5.3.2 Types of participants

Women with PPROM before 37+0 weeks' gestation with no specific maternal or fetal contraindications to expectant management (defined by trialists).

2.5.3.3 Types of interventions

All methods of fetal assessment that could detect fetal compromise and provide indication for early delivery. These include fetal movement counting, fetal cardiotocography, biophysical profiling, fetal and umbilical artery or venous Doppler ultrasound, fetal MRI lung volumetrics and amniocentesis for fetal lung maturity. Comparisons were made of any intervention versus no intervention and one intervention versus another intervention.

2.5.3.4 Types of outcome measures

Outcomes of interest are listed in Table 2.6.

Table 2.6. Primary and secondary outcomes addressed in the review.

Primary outcomes
Fetal death (antenatal).
Neonatal death (in the first 28 days of life).
Maternal death.
Serious maternal morbidity defined as: 1) septicaemia, 2) need for intensive care, 3) organ failure/need for ventilation, 4) need for hysterectomy.
Secondary outcomes
Maternal
Maternal chorioamnionitis (as defined by trialists).
Major postpartum haemorrhage.
Maternal endometritis (as defined by trialists).
Mode of delivery.
Induction of labour.
Postpartum maternal pyrexia (as defined by trialists).
Days of antenatal hospitalisation.
Days of postnatal hospitalisation.
Total days hospitalisation.
Breastfeeding initiated in hospital.
Breastfeeding at hospital discharge.
Maternal satisfaction.
Caesarean delivery.
Spontaneous labour before 34+0 weeks.
Maternal antibiotics after delivery.
Time of maternal antibiotic treatment before labour (hours).
Fetal
Gestational age at birth.
Days from randomisation/rupture of membrane to birth.
Birth within 48 hours after rupture of membranes.
Birth within seven days of rupture of membranes.
Birth before 37+0 weeks' gestation.
Birth before 34+0 weeks' gestation.
Birth before 28+0 weeks' gestation.
Use of corticosteroids.
Fetal sepsis.
Neonatal
Postneonatal mortality - death after 28 days of life but before one year.
Infant mortality - death at or after 12 months of age.
Respiratory distress syndrome.
Use of surfactant.
Use of mechanical ventilation.
Days of mechanical ventilation.
Days of oxygen therapy.

Oxygen treatment greater than 28 days.
Oxygen therapy at 36+0 weeks' gestation.
Birthweight.
Birthweight less than 10th centile for gestational age.
Birthweight less than 2500 grams.
Birthweight less than 1500 grams.
Admission to neonatal intensive care unit.
Length of stay in neonatal intensive care unit.
Days from birth to discharge home from hospital.
Major cerebral abnormalities on ultrasound prior to discharge.
Necrotising enterocolitis.
Neonatal encephalopathy (as described by trialists).
Postural deformities (as defined by trialists).
Disability at time of childhood follow up (as defined by trialists).
Serious disability (as defined by trialists) after two years.
Diagnosis of fetal distress in labour (as defined by trialists).
Cardiotocographic abnormality in labour (as defined by trialists).
Cord pH less than 7.00.
Apgar scores less than seven at five minutes.
Apgar scores over five at five minutes.
Apgar scores.
Intraventricular haemorrhage.
Use of antibiotics.
Days of antibiotic use.
Neonatal infection.
Neonatal sepsis.
Presumed neonatal sepsis.
Transient tachpnea.
Pneumothorax.
Pneumonia.
Persistent ductus arteriosus.
Seizures.
Hyperbilirubinaemia.
Metabolic problems.
Apnea.
Feeding difficulties.
Days of postnatal hospitalisation (after discharge of mother).
Retinopathy of prematurity.
Delivery for maturity.
Other
Caregiver satisfaction.

2.5.4 Search methods for identification of studies

2.5.4.1 Electronic searches

The Cochrane Trials Search Co-ordinator searched the Cochrane Pregnancy and Childbirth Group's Trials Register on 28th November 2012. The Trials Register is maintained by the Trials Search Co-ordinator and contains trials identified from:

- monthly searches of the Cochrane Central Register of Controlled Trials (CENTRAL);
- weekly searches of MEDLINE;
- weekly searches of EMBASE;
- handsearches of 30 journals and the proceedings of major conferences;
- weekly current awareness alerts for a further 44 journals plus monthly BioMed Central email alerts.

Trials identified through these searches are each assigned to a review topic (or topics). The Trials Search Co-ordinator searches the register for each review using the topic list rather than keywords. No language restrictions were applied.

2.5.5 Data collection and analysis

2.5.5.1 Selection of studies

Two review authors (GS, SS) independently assessed for inclusion all the potential studies identified as a result of the search strategy. Any disagreement was resolved through discussion or, if required, consultation with a third person (JEN).

2.5.5.2 Data extraction and management

A data extraction form was designed (Appendix 7) and used to extract data from the studies. For eligible studies, two review authors (GS, SS) extracted the data independently. Discrepancies were resolved through discussion or, if required, consultation with a third author (JEN). Data was entered into The Cochrane Collaboration's statistical software, Review Manager, and checked for accuracy.

2.5.5.3 Assessment of risk of bias in included studies

Two review authors (GS and SS) independently assessed risk of bias for each study using the criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* (218). Any disagreement was resolved by discussion or by involving a third author (JEN).

2.5.5.3.1 1. Random sequence generation (checking for possible selection bias)

For each included study, the review describes the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.

The method was assessed as:

- low risk of bias (any truly random process, e.g. random number table; computer random number generator);
- high risk of bias (any non-random process, e.g. odd or even date of birth; hospital or clinic record number); or
- unclear risk of bias.

2.5.5.3.2 2. Allocation concealment (checking for possible selection bias)

For each included study, the review describes the method used to conceal allocation to interventions prior to assignment and assesses whether intervention allocation could have been foreseen in advance of, or during recruitment, or changed after assignment.

The methods were assessed as:

- low risk of bias (e.g. telephone or central randomisation; consecutively numbered sealed opaque envelopes);
- high risk of bias (open random allocation; unsealed or non-opaque envelopes, alternation; date of birth); or
- unclear risk of bias.

2.5.5.3.3 3.1. Blinding of participants and personnel (checking for possible performance bias)

For each included study, the review describes the methods used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Studies were considered to be at low risk of bias if they were blinded, or if it was judged

that the lack of blinding would be unlikely to affect results. Blinding was assessed separately for different outcomes or classes of outcomes.

The methods were assessed as:

- low, high or unclear risk of bias for participants;
- low, high or unclear risk of bias for personnel.

2.5.5.3.4 3.2. Blinding of outcome assessment (checking for possible detection bias)

For each included study, the review describes the methods used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Blinding was assessed separately for different outcomes or classes of outcomes. The method of outcome assessor blinding was assessed as low, high or unclear risk of bias.

2.5.5.3.5 4. Incomplete outcome data (checking for possible attrition bias due to the amount, nature and handling of incomplete outcome data)

For each included study, and for each outcome or class of outcomes, the review describes the completeness of data including attrition and exclusions from the analysis. The review states whether attrition and exclusions were reported and the numbers included in the analysis at each stage (compared with the total randomised participants), reasons for attrition or exclusion where reported, and whether missing data were balanced across groups or were related to outcomes. Where sufficient information was reported, or could be supplied by the trial authors, it was planned that missing data would be re-included in the analyses.

Methods were assessed as:

- low risk of bias (e.g. no missing outcome data; missing outcome data balanced across groups);
- high risk of bias (e.g. numbers or reasons for missing data imbalanced across groups; 'as treated' analysis done with substantial departure of intervention received from that assigned at randomisation); or
- unclear risk of bias.

2.5.5.3.6 5. Selective reporting (checking for reporting bias)

For each included study, the possibility of selective reporting bias was assessed by comparing outcomes reported in the results section with those mentioned in the methods section.

We assessed the methods as:

- low risk of bias (where it is clear that all of the study's pre-specified outcomes and all expected outcomes of interest to the review have been reported);
- high risk of bias (where not all the study's pre-specified outcomes have been reported; one or more reported primary outcomes were not pre-specified; outcomes of interest are reported incompletely and so cannot be used; study fails to include results of a key outcome that would have been expected to have been reported); or
- unclear risk of bias.

2.5.5.3.7 6. Other bias

Any important concerns about other possible sources of bias were considered, and the risk was assessed as:

- low risk of other bias;
- high risk of other bias; or
- unclear whether there is risk of other bias.

2.5.5.3.8 7. Overall risk of bias

Explicit judgements were made about whether studies were at high risk of bias, according to the Cochrane guidelines (218). The likely magnitude and direction of the bias and whether it was likely to impact on the findings was considered. It was planned that the impact of the level of bias would be explored through undertaking sensitivity analysis if necessary.

2.5.5.4 Measures of treatment effect

2.5.5.4.1 Dichotomous data

For dichotomous data, results are presented as summary risk ratios (RRs) with 95% confidence intervals (CIs).

2.5.5.4.2 Continuous data

For continuous data, the mean difference (MD) was used if outcomes are measured in the same way between trials. The standardised mean difference (SMD) was used to combine trials that measure the same outcome, but using different methods.

2.5.5.5 Dealing with missing data

For included studies, levels of attrition were noted. It was planned that the impact of including studies with high levels of missing data in the overall assessment of treatment effect would be explored using sensitivity analysis.

For all outcomes, analyses were carried out, as far as possible, on an intention-to-treat basis, i.e. all participants randomised to each group in the analyses were included, and all participants were analysed in the group to which they were allocated, regardless of whether or not they received the allocated intervention. The denominator for each outcome in each trial was the number randomised minus any participants whose outcomes are known to be missing.

2.5.5.6 Assessment of heterogeneity

Heterogeneity of treatment effects was measured between trials using the T^2, I^2 and Chi^2 statistics. Heterogeneity was regarded as substantial if I^2 was greater than 30% and either T^2 was greater than zero, or there was a low P value (less than 0.10) in the Chi^2 test for heterogeneity.

2.5.5.7 Data synthesis

Statistical analysis was carried out using the Review Manager software. Fixed-effect meta-analysis was used for combining data where it was reasonable to assume that studies were estimating the same underlying treatment effect: i.e. where trials were examining the same intervention, and the trials' populations and methods were judged sufficiently

similar. If there was clinical heterogeneity sufficient to expect that the underlying treatment effects differed between trials, or if substantial statistical heterogeneity was detected, a random-effects meta-analysis would have been used to produce an overall summary.

2.5.6 Writing the review

The review was produced using the Review Manager software developed by The Cochrane Collaboration for preparing and maintaining Cochrane reviews. Review Manager facilitates the preparation of protocols and full reviews, including text, characteristics of studies, comparison tables and study data. It can perform meta-analysis of the data entered and present the results graphically.

Once written, the review was submitted to the Cochrane Collaboration for peer review and publication.

3 USING MICROARRAY DATA TO STUDY THE MYOMETRIAL TRANSCRIPTOME IN PREGNANCY AND PARTURITION

3.1 INTRODUCTION

Over the past decade, microarray technology has been increasingly applied to study parturition (219). Such experiments provide large amounts of gene expression data, which is combined with existing knowledge of molecular pathways to study the molecular mechanisms underlying the initiation of labour and labour complications. The myometrium is arguably the most relevant tissue in which to conduct such studies, being the tissue that contracts to expel the baby. Several research groups have published microarray data comparing gene expression in labouring and non-labouring myometrium (55,112,114,115,119,120,140,141). Although there is some agreement between these studies on the general mechanisms that appear to underlie the initiation of labour, no clear agreed labour-associated myometrial gene expression signature has been identified. This may be partly due to variation introduced by the use of different microarray platforms. It may also be attributable to the small number of samples used in these studies - a problem that often arises due to the relative inaccessibility of myometrial tissue and efforts to select only those patients with similar key characteristics (such as gestational age, parity, BMI, maternal age and indication for Caesarean delivery). Although the ideal experimental design would control for these characteristics in this way, perhaps excluding gestational age (115), there is little evidence to suggest that the molecular mechanisms initiating parturition are affected by these factors. Therefore, it may be more useful to prioritise maximising the number of samples over minimising inter-patient variability.

This chapter describes the results of the largest-to-date microarray experiment to compare gene expression in labouring and non-labouring myometrium. All labouring samples in the Edinburgh Reproductive Tissue BioBank (ERTBB) were selected regardless of patient characteristics and then loosely matched to similar non-labouring samples (of which there were far more). The data is analysed using robust statistical techniques and a modern unbiased network graph approach, before functional enrichment to provide further insight into the biological meaning. To further increase the statistical power and allow inter-study heterogeneity to be assessed, a meta-analysis of similar published microarray raw data is also conducted. This is the first time such a meta-analysis has been

carried out. This approach will likely lead to more robust, accurate and generalisable conclusions and could be widely adopted.

3.2 SUMMARY OF METHODS

The methods used in this chapter are described in detail in Chapter 2. Briefly, 22 labouring and 26 non-labouring lower segment myometrium samples were selected from the Edinburgh Reproductive Tissue Biobank (described in section 2.1.1). RNA was isolated, amplified and labelled with biotin (described in section 2.1.2). The microarray experiment was conducted at the Wellcome Trust Clinical Research Facility using the Illumina HT-12 v4 Expression BeadChip platform (described in section 2.1.3). Raw data was subjected to RMA background correction and quantile normalisation to remove non-biological systemic variation (described in section 2.1.4). Probes were annotated and genes expressed, on average, below background in both labouring and non-labouring groups were removed (described in section 2.1.5 and 2.1.6). Expression over all probes was used to assess relationships between samples (described in section 2.1.7). Differential expression between groups was explored by observing fold-changes and P- or RP-PFP values as a measure of statistical significance (described in section 2.1.8). Biolayout *Express*^{3D} was also used to visualise the data and provide further insights (described in section 2.1.10). Functional analysis was performed by uploading gene lists to DAVID and comparing biological GO terms and KEGG pathways (described in section 2.1.11). The microarray was validated using TaqMan qRT-PCR for selected up-, down- and non-differentially regulated genes (described in section 2.1.13). The relative abundance of leukocytes in labouring and non-labouring tissue was assessed by staining myometrial tissue sections for CD45 (described in section 2.1.12). In addition to the wet-lab validation of this microarray, a meta-analysis of the complete data from previously published similar microarray experiments was conducted (described in section 2.1.14). A systematic search of the literature identified comparable studies and data was preprocessed and analysed using similar methods to those described above. Differential expression between labouring and non-labouring groups was assessed for individual studies, and agreement between studies was also investigated visually and using Forrest plots of the standardised mean difference.

3.3 RESULTS

3.3.1 RNA QUALITY

Before including RNA samples in the microarray experiment, the quality was assessed. Total RNA extracted from myometrium samples was analysed using a NanoDrop 1000. Pure RNA is expected to show a 260nm/280nm ratio of around 2.0. The samples assessed for use in this study had a mean 260nm/280nm ratio of 2.07 (standard deviation = 0.04), so the quality of the RNA was considered high enough to undergo biotin labelling in preparation for the array. After biotin labelling, the Wellcome Trust Clinical Research Facility used a Bioanalyzer 2100 to assess the quality of the RNA again. For biotin-labelled mRNA, Bioanalyzer traces are not associated with an RNA Integrity Number (RIN), however the experienced staff at the WTCRF confirmed that all the traces suggested the samples were of satisfactory quality so all samples were included on the chips. The CD accompanying this thesis contains all the Bioanalyzer results obtained by the WTCRF.

Using data obtained from the microarray experiment, Figure 3.1 shows boxplots and intensity plots to visualise the distribution of expression values across all probes and all samples before (Figure 3.1a and Figure 3.1c) and after (Figure 3.1b and Figure 3.1d) quantile normalisation. There are no large variations in expression values between samples, which indicates that the RNA was of acceptable quality.

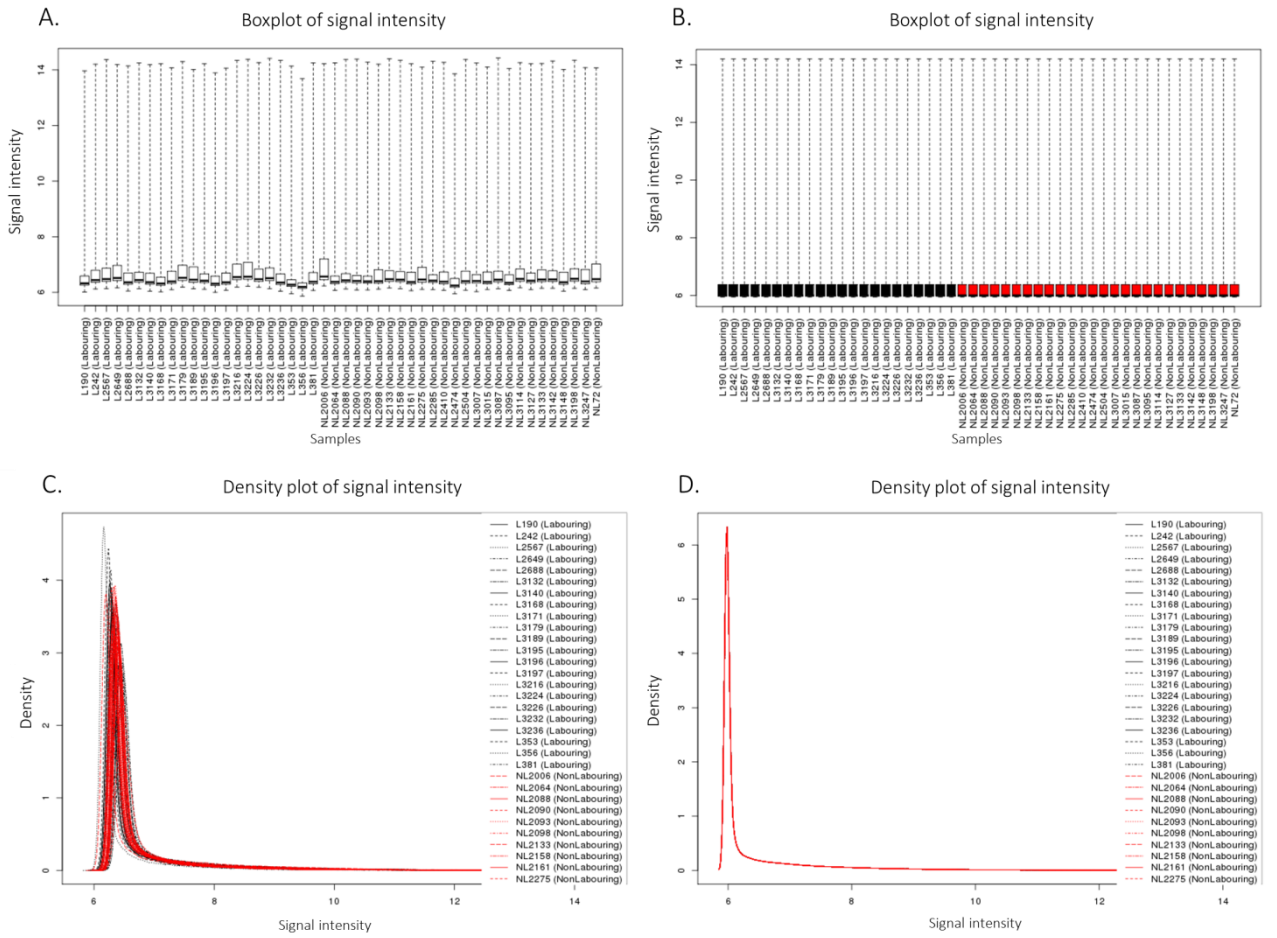


Figure 3.1. Boxplots and density plots of signal intensity (expression value) before (A, C) and after (B,D) quantile normalisation.

3.3.2 Sample characteristics

Table 3.1 shows the patient characteristics for the samples included in the microarray. There are 22 labouring and 26 non-labouring samples. Table 3.2 summarises patient characteristics (e.g. BMI, maternal age, etc.) for both groups. The samples were not selected to minimise inter-patient variability, so there is a reasonably high degree of variation in patient characteristics within both groups. Within the labouring group, there is a high degree of variation in cervical dilation, suggesting that the women range from early to late stages of labour. There are no significant between-group differences at $P<0.05$ in mean/median values of parity, maternal age, maternal BMI or gestational age at delivery, so the groups are reasonably well matched and any difference in gene expression between the groups is likely to be the result of labour status rather than variation in these characteristics.

Key for Tables 3.1 and 3.2:

RL = RNA later

FD = Fetal distress

FTP = Failure to progress

N/R = Not recorded

PPROM = Preterm prelabour rupture of membranes

B = Breech

PP = Placenta praevia

O/H = Obstetric history

IUGR = intrauterine growth restriction

PIH = Pregnancy-induced hypertension

PCS = Previous Caesarean section

Hip = Hip problems

PG = Prostaglandin

S = Syntocinon

Table 3.1. Characteristics for samples included in the microarray experiment.

Sample ID	Tissue storage	Indication for CS	Maternal age	BMI	Method of induction	Gestation (weeks +days)	Parity	Cervical dilation (cm)
L0190	RL	FD	44	42.4	PG & S	40+3	0+2	9
L0242	RL	FTP	37	21.6	S	41+3	0+0	N/R
L0353	RL	N/R	32	39.8	PG	39+0	1+0	N/R
L0356	No RL	N/R	21	43.7	PG	39+6	0+0	N/R
L0381	No RL	N/R	28	20.7	None	36+4	N/R	N/R
L2567	RL	FD	28	30.1	None	41+2	0+0	10
L2649	RL	FTP	39	22.2	None	40+5	1+1	6
L2688	RL	PPROM/B	36	25.8	None	35+0	1+2	1
L3132	RL	FTP	32	22.7	S	40+4	0+0	4.5
L3140	RL	B	28	19.1	None	33+4	0+0	5
L3168	RL	N/R	32	39.3	S	39+2	0+0	8
L3171	RL	FTP	34	20.7	None	41+0	1+0	1
L3179	RL	FTP	22	26.2	None	40+6	1+0	3
L3189	RL	FTP	28	21.2	S	40+3	0+0	5
L3195	RL	B	28	25.7	None	41+3	0+1	N/R
L3196	RL	FTP	29	21.6	S	39+6	0+0	5.6
L3197	RL	B	39	21.7	None	40+1	0+0	2
L3216	RL	FTP	23	23.1	S	41+4	0+0	9
L3224	RL	FTP	33	27	None	41+6	0+0	8
L3226	RL	FD	36	25.6	None	42+2	0+1	4
L3232	No RL	FTP	39	40.3	PG	41+6	2+0	8
L3236	RL	FTP	30	36	S	36+6	N/R	9
NL0072	RL	N/R	30	23.9	None	39+0	≥1	n/a
NL2006	RL	B	33	23.4	None	39+1	0+1	n/a
NL2064	RL	PP	43	27.2	None	38+4	0+0	n/a
NL2088	RL	B	31	20.5	None	39+5	0+0	n/a
NL2090	RL	O/H	32	27.7	None	39+5	2+0	n/a
NL2093	RL	B	33	41.2	None	39+0	1+1	n/a
NL2098	RL	B	30	22	None	39+3	0+0	n/a
NL2133	RL	N/R	35	29.4	None	41+4	0+0	n/a
NL2158	RL	B	23	21.1	None	40+1	0+1	n/a
NL2161	RL	O/H	36	23	None	37+0	2+0	n/a
NL2275	RL	B	38	23.2	None	39+1	0+0	n/a
NL2285	RL	B	34	22.8	None	40+0	0+2	n/a
NL2410	RL	PP	32	23.1	None	38+5	0+0	n/a
NL2474	RL	IUGR	36	38.1	None	31+2	0+1	n/a
NL2504	RL	PIH	34	26.6	None	36+4	1+0	n/a
NL3007	RL	B	34	46.8	None	39+3	0+0	n/a
NL3015	RL	B	40	26.5	None	39+2	0+1	n/a
NL3087	RL	PCS	31	43	None	39+4	1+2	n/a
NL3095	RL	PCS	26	26.3	None	39+1	1+1	n/a
NL3114	RL	PCS	33	32.3	None	39+2	1+0	n/a
NL3127	RL	B	31	31.7	None	39+3	0+0	n/a
NL3133	RL	N/R	35	44.4	None	39+4	0+0	n/a
NL3142	RL	N/R	38	24.6	None	39+5	1+0	n/a
NL3148	RL	Hip	22	24.3	None	39+2	0+0	n/a
NL3198	RL	B	29	23.9	None	39+5	2+0	n/a
NL3247	RL	O/H	32	22.2	None	39+0	1+0	n/a

Table 3.2. Summary of patient characteristics for the labouring and non-labouring groups. P-values represent the significance of the difference between the groups.95% confidence intervals are shown in parentheses.

Tissue storage	Indication for CS	Mean maternal age	Mean BMI	Method of induction	Mean gestation (weeks +days)	Median parity	Mean cervical dilation (cm)
Labouring Group							
RL = 19 No RL = 3	FD = 3 FTP = 11 B = 3 CS/OH = 0 Other = 1 N/R = 4	32 (29.1-34.4)	28.0 (24.4-31.6)	PG = 3 S = 7 Both = 1 None = 11	40 (39-41)	0 (0-1)	5.8 (4.3-7.3)
Non-labouring Group							
RL = 26 No RL = 0	FD = 0 FTP = 0 B = 11 CS/OH = 6 Other = 5 N/R = 4	32 (30.8-34.6)	28.4 (25.3-31.6)	None = 26	39 (38-40)	0 (0-1)	n/a
P-value							
n/a	n/a	0.52	0.86	n/a	0.15	0.45	n/a

3.3.3 Sample relations

Expression values for all probes were used to calculate Pearson's correlation values to describe how the gene expression values of each sample correlate with those of each of the other samples. Figure 3.2 shows the mean correlation values for three comparisons: labouring versus labouring samples, non-labouring versus non-labouring samples, and non-labouring versus labouring samples. Non-labouring samples are significantly more highly correlated with other non-labouring samples than they are with labouring samples (one way ANOVA with Tukey's multiple comparisons test $P<0.0001$). Labouring samples are not more highly correlated with each other than they are with non-labouring samples (one way ANOVA with Tukey's multiple comparisons test $P=0.12$). The error bars show that the standard deviation for labouring samples is also higher than that for non-labouring samples. This suggests that non-labouring samples show a similar gene expression signature, but that labouring samples are more varied.

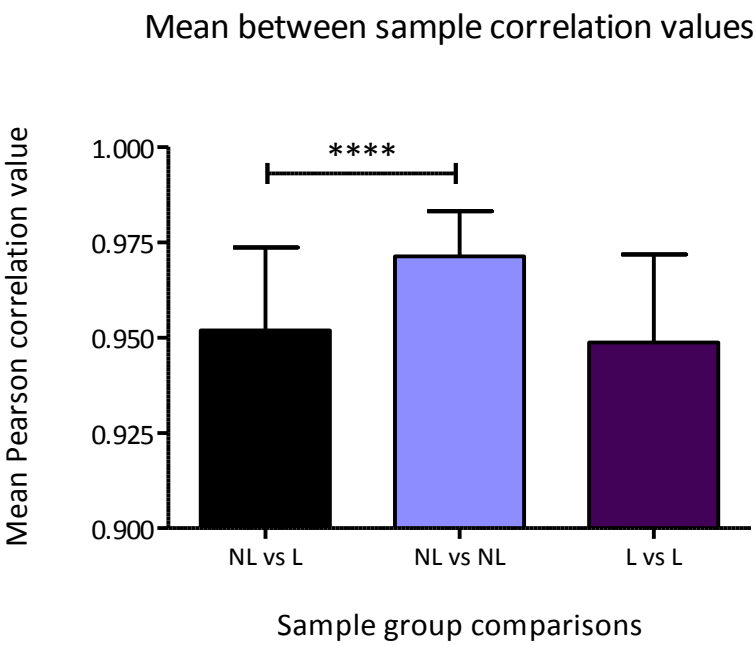


Figure 3.2. Mean between-sample Pearson's correlation values. Non-labouring samples (NL) are significantly more highly correlated with each other than with labouring samples (L) ($P<0.0001$). Labouring samples are not significantly more or less correlated with each other than they are with non-labouring samples ($P=0.12$). Error bars indicate standard deviation.

Figure 3.3 shows a heatmap that illustrates this further by indicating the strength of the correlation between each sample. Non-labouring samples show high correlation with each other (indicated by lighter colours), and lower correlation with labouring samples (indicated by darker colours). When compared with each other, labouring samples show a degree of correlation similar to that observed in comparison to non-labouring samples, in other words, in contrast to the scenario with non-labouring samples, labouring samples are not more correlated to each other than they are to non-labouring samples.

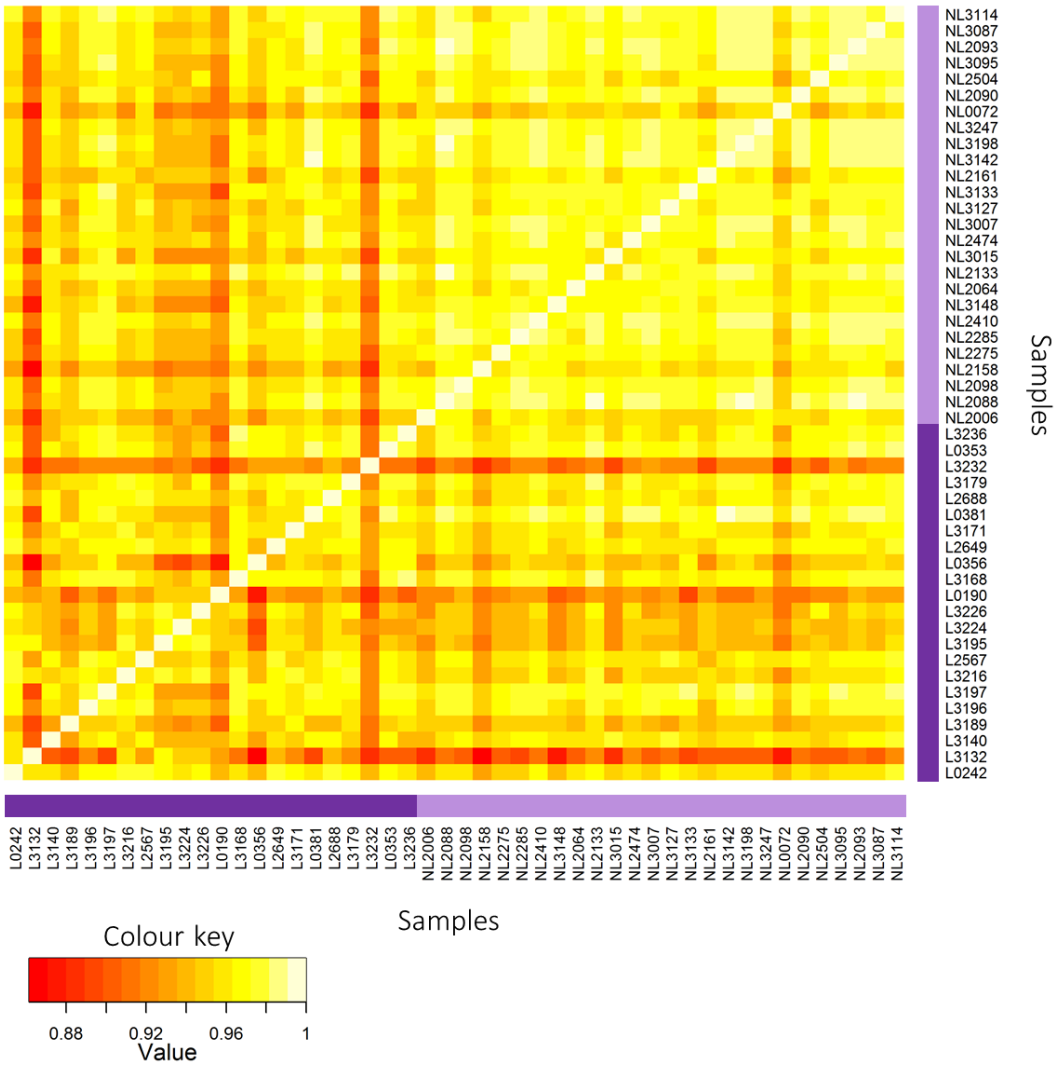


Figure 3.3. A heatmap illustrating the degree of correlation of gene expression between each sample on the array. The lighter the colour, the stronger the correlation. Samples are listed along the axes by their IDs. Samples beginning with L are from labouring women and samples beginning with NL are from non-labouring women.

In the network graph shown in Figure 3.4, the nodes represent samples and the edges are weighted according to the strength of the correlation. Only Pearson's correlation values above 0.95 are included when building the graph. Markov Chain clustering (MCL) was performed to assess how the samples cluster into groups. It produced two clusters. When nodes are coloured according to labour status it is clear that MCL cluster 1 (42 samples) contains all 26 of the non-labouring samples and MCL cluster 2 (5 samples) contains labouring samples only. A Chi-squared test confirms that this association is significant

(P=0.01). This suggests that the samples cluster by labour status, and the appearance of labouring samples in both clusters supports the above finding that labouring samples show more variation than non-labouring samples.

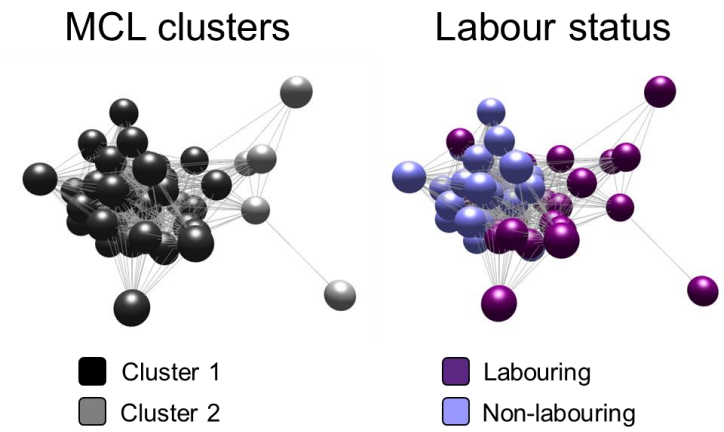


Figure 3.4. A network graph in which each node represents a different sample on the array. The thickness of the edges indicates the strength of the correlation. The same graph is coloured according to Markov chain (MCL) cluster number and labour status.

3.3.4 Assessing differential expression between labouring and non-labouring samples

A complete list of differentially expressed genes at a fold change of <- or > 1.2 and a RP-PFP value of <0.05 is presented on the CD accompanying this thesis. There are 960 genes upregulated and 801 genes downregulated in labour. The 40 most significant differentially expressed genes are summarised in Table 3.3 (upregulated) and Table 3.4 (downregulated).

Table 3.3 The 20 genes most significantly upregulated in labour.

Gene Symbol	Ilumina Probe ID	Gene Name	Fold Change (NL to L)	Adjusted P-value
IL8	ILMN_2184373	interleukin 8	9.6	4.2E-05
MT2A	ILMN_1686664	metallothionein 2A	6.9	8.1E-08
MT1A	ILMN_1691156	metallothionein 1A	6.9	6.3E-08
RARRES1	ILMN_1800091	retinoic acid receptor responder (tazarotene induced) 1	4.2	3.8E-07
CCL2	ILMN_1720048	chemokine (C-C motif) ligand 2	4.0	3.8E-05
ANGPTL4	ILMN_1707727	angiopoietin-like 4	3.7	6.9E-07
MT1G	ILMN_1715401	metallothionein 1G	3.3	2.8E-07
MT1X	ILMN_1775170	metallothionein 1X	3.2	1.8E-06
C19orf59	ILMN_1762713	chromosome 19 open reading frame 59	3.1	1.0E-05
CHI3L1	ILMN_3307868	chitinase 3-like 1 (cartilage glycoprotein-39)	3.0	4.3E-05
MT1IP	ILMN_2136089	metallothionein 1I pseudogene	3.0	1.8E-06
IL6	ILMN_1699651	interleukin 6 (interferon beta 2)	2.9	2.3E-06
LILRA5	ILMN_2266595	leukocyte immunoglobulin-like receptor subfamily A (with TM domain) member 5	2.8	6.6E-07
FCN1	ILMN_1668063	ficolin (collagen/fibrinogen domain containing) 1	2.8	3.1E-05
AQP9	ILMN_1715068	aquaporin 9	2.8	1.5E-04
LILRA3	ILMN_1661631	leukocyte immunoglobulin-like receptor subfamily A (without TM domain) member 3	2.7	1.2E-04
MMP3	ILMN_1784459	matrix metallopeptidase 3 (stromelysin 1 progelatinase)	2.7	6.0E-04
CCL23	ILMN_1686109	chemokine (C-C motif) ligand 23	2.7	1.8E-06
SLC11A1	ILMN_1741165	solute carrier family 11 (proton-coupled divalent metal ion transporters) member 1	2.6	4.2E-06
HBEGF	ILMN_2121408	heparin-binding EGF-like growth factor	2.4	2.1E-04

Table 3.4. The 20 genes most significantly downregulated in labour.

Gene Symbol	Ilumina Probe ID	Gene Name	Fold Change (NL to L)	Adjusted P-value
MAMDC2	ILMN_1679391	MAM domain containing 2	-2.1	9.8E-06
APCDD1L	ILMN_1689431	adenomatosis polyposis coli down-regulated 1-like	-2.1	2.7E-04
FABP4	ILMN_1773006	fatty acid binding protein 4 adipocyte	-2.0	2.5E-05
SVIL	ILMN_1690754	supervillin	-1.9	2.6E-06
HCFC1R1	ILMN_1757877	host cell factor C1 regulator 1 (XPO1 dependent)	-1.9	4.8E-06
SEPP1	ILMN_1785071	selenoprotein P plasma 1	-1.8	9.3E-05
CD34	ILMN_1732799	CD34 molecule	-1.8	2.8E-06
GPR161	ILMN_1773940	G protein-coupled receptor 161	-1.8	2.8E-06
RERG	ILMN_1746359	RAS-like estrogen-regulated growth inhibitor	-1.8	8.0E-06
CD34	ILMN_2341229	CD34 molecule	-1.8	2.3E-04
SORBS1	ILMN_1749792	sorbin and SH3 domain containing 1	-1.8	4.4E-05
PPP1R3C	ILMN_1736670	protein phosphatase 1 regulatory (inhibitor) subunit 3C	-1.8	8.7E-05
LMOD1	ILMN_1680948	leiomodin 1 (smooth muscle)	-1.7	7.1E-07
DAAM1	ILMN_1787251	dishevelled associated activator of morphogenesis 1	-1.7	1.2E-06
MYH11	ILMN_2368834	myosin heavy chain 11 smooth muscle	-1.7	2.9E-06
SNORD13	ILMN_1892403	small nucleolar RNA C/D box 13	-1.7	3.5E-07
PLCL1	ILMN_2206953	phospholipase C-like 1	-1.7	6.3E-06
LINGO2	ILMN_1695978	leucine rich repeat and Ig domain containing 2	-1.7	2.6E-07
KLRAQ1	ILMN_1667356	KLRAQ motif containing 1	-1.7	5.5E-05
C7orf58	ILMN_1677038	chromosome 7 open reading frame 58	-1.6	6.0E-08

Differentially expressed genes are visualised in Figure 3.5. This heatmap represents expression values as colours on a scale. Columns represent samples and rows represent differentially expressed genes. To make the list small enough to enable clear visualisation, a fold change cut-off of >2 or <-2 and an RP-PFP cut-off of ≤0.0001 were used. The heatmap shows that most genes are upregulated at labour. Additionally samples cluster almost perfectly according to labour status.

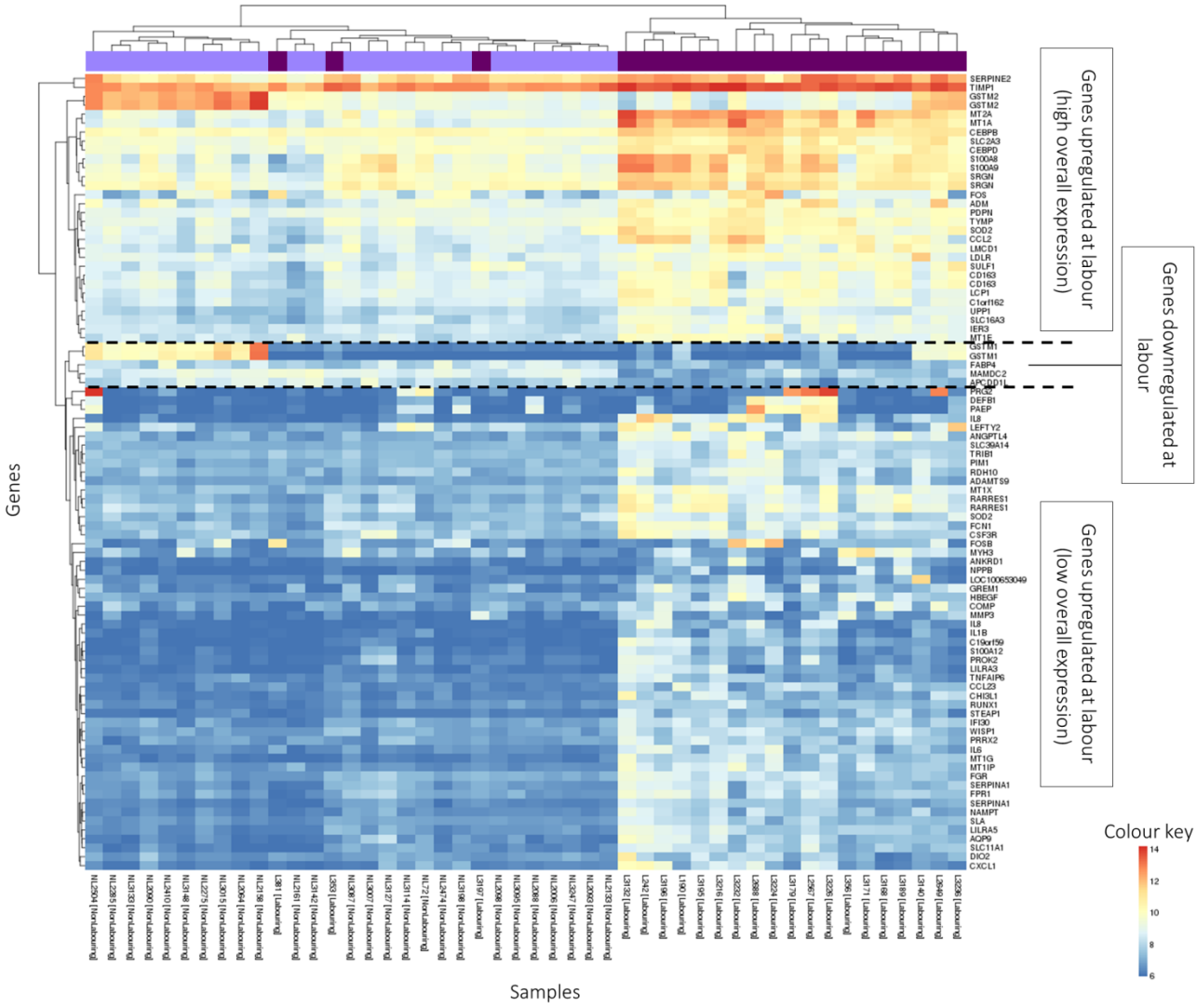


Figure 3.5. A heatmap to show how genes and samples cluster based on similar expression levels. The bar at the top indicates the sample group (dark purple = labouring, light purple = non-labouring). Normalised expression values are indicated on a colour scale with red indicating high expression and blue indicating low expression. Genes with a non-labour to labour fold change of >2 or <-2 and an RP-PFP ≤0.0001 are included in the heatmap.

3.3.4.1 Network graph analysis

Further analysis and visualisation of the data was conducted using a network graph approach. In Figure 3.6a each node is a gene probe and the edges connecting them are based on similarity of expression profiles across the samples. Probes with a correlation value above 0.8 were included in the graph. The graph is comprised of 2304 nodes and

20984 edges. The nodes are coloured according to the MCL cluster they belong to. There are 227 clusters in total (calculated using an inflation value of 2.2). To aid visualisation of the graph, Figure 3.6b shows a 2D view where each cluster is condensed to a single node. The size of the node is proportional to the size of the cluster. The largest clusters are numbered.

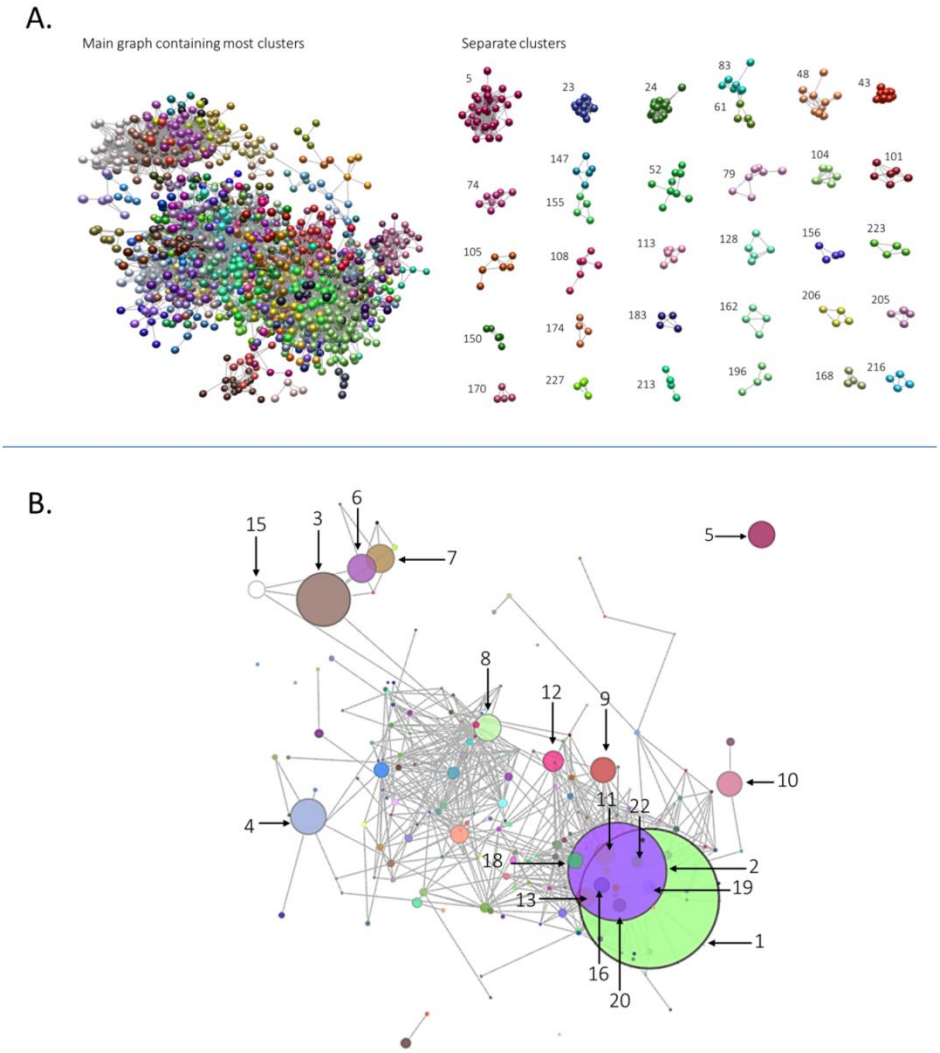


Figure 3.6. (A) A network graph of the microarray data coloured according to MCL cluster. (B) A 2D condensed view of the graph, where each cluster is represented as a single node.

MCL clusters are formed based on similar probe expression profiles. Figure 3.7 shows four example cluster expression profiles. Figure 3.7a shows a cluster comprising probes that are generally upregulated in labouring samples. Figure 3.7b shows a cluster of probes that

are not differentially expressed according to any patient characteristics. Figure 3.7c shows a cluster of probes that are generally downregulated in labouring samples. Figure 3.7d shows a cluster of probes that are generally upregulated in labouring samples, but only for first deliveries (para=0).

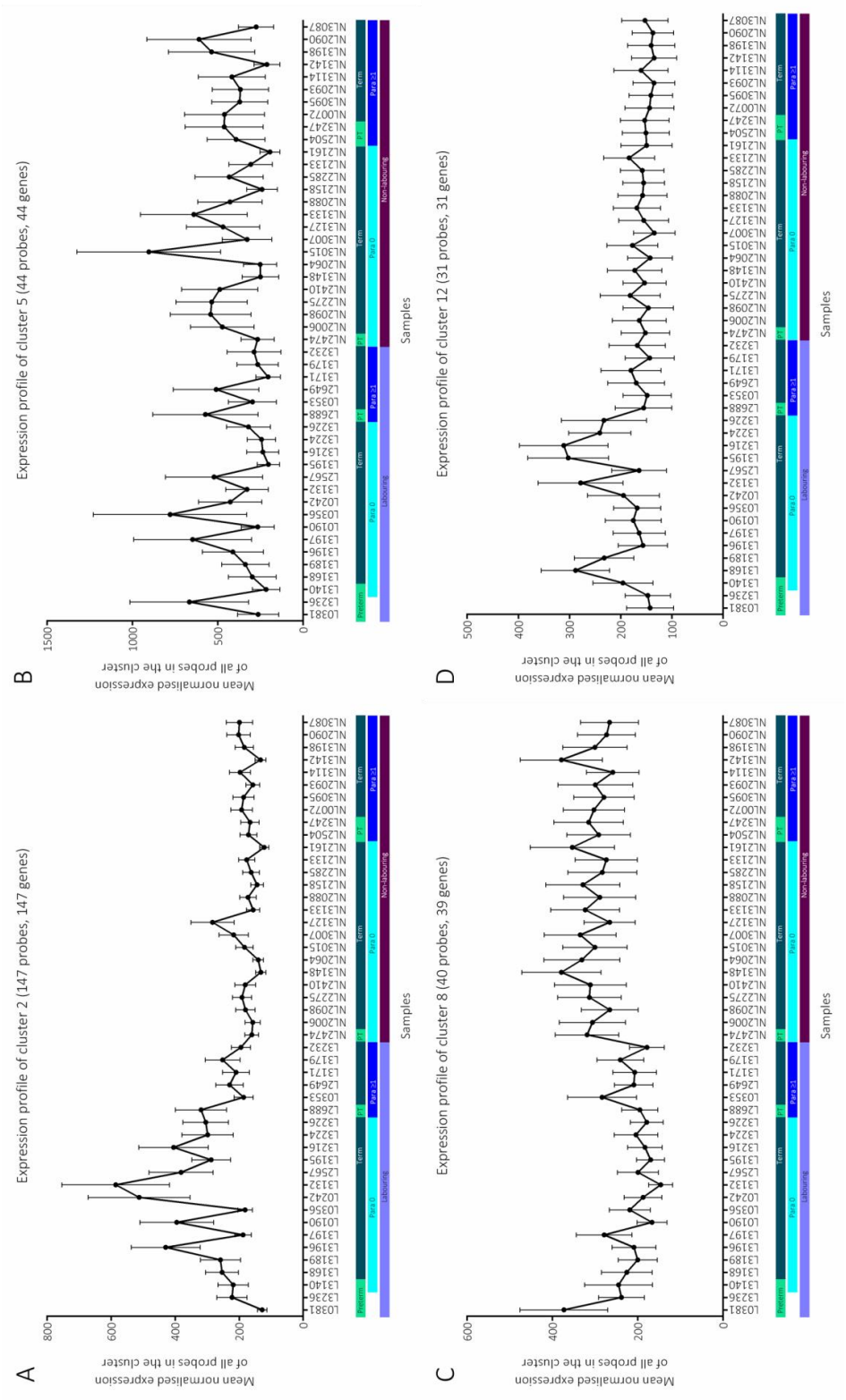


Figure 3.7. Expression profiles for four clusters. Error bars indicate 95% confidence intervals.

3.3.5 qRT-PCR VALIDATION

qRT-PCR was performed using all samples in the array to validate the expression of ten genes. I chose a selection of genes based on specific changes or no change in labour, as identified from the microarray and/or the literature (the rationale behind choosing these genes is explained in Chapter 2, Table 2.2).

The standard curve method was used to assess gene expression relative to an endogenous control (18s). Standard curve slopes should be between -3.6 and -3.3 to confirm reactions are 90-100% efficient. The mean standard curve slope for these experiments was -3.46 (standard deviation= 0.16). Despite repeating reactions, one target gene (RBM42) was consistently around -3.63, and another target gene (FABP4) was consistently around 3.06, indicating issues with the reaction. Therefore the results for these target genes should be considered less reliable than the other results.

For each target gene, three samples (L0353, L2649 and NL2504) were excluded from analysis because CT values were over 35 or undetermined. Values over 35 are considered unreliable. All samples were run in technical duplicate (standard curves were run in triplicate). Where the mean CT values of technical duplicates had a standard deviation over 0.2, the reaction was repeated. This did not resolve the difference for MTE1, so the results were included in further analysis but the implications of this are noted. After the three exclusions there were 25 non-labouring biological replicates and 20 labouring biological replicates.

A table showing the mean raw, normalised and delta CT and quantity values for each gene, as well as the fold change in gene expression between non-labouring and labouring samples is provided in Appendix 2. T-tests were used to analyse whether the delta CT values (raw CT value for the target gene minus the raw CT value for 18s) were significantly different in labouring and non-labouring samples. This data is plotted alongside results from the microarray experiment in figures 3.8-3.10.

Interleukin-8 (IL-8), Interleukin-6 (IL-6) and Metallothionein1E (MT1E) were confirmed to be upregulated in labour (Figure 3.8). RNA binding motif protein 42 (RBM42) and Shroom family member 4 (SHROOM4) were confirmed to be not differentially expressed between groups (Figure 3.9). Oxytocin receptor (OXTR), Myosin heavy chain 11 (MYH11) and Fatty acid binding protein 4 (FABP4) were confirmed to be downregulated in labour.

However, contrary to the microarray results, Tropomyosin 1 (TPM1) and Insulin-like growth factor binding protein 5 (IGFBP5) were not significantly downregulated in labouring samples (Figure 3.10).

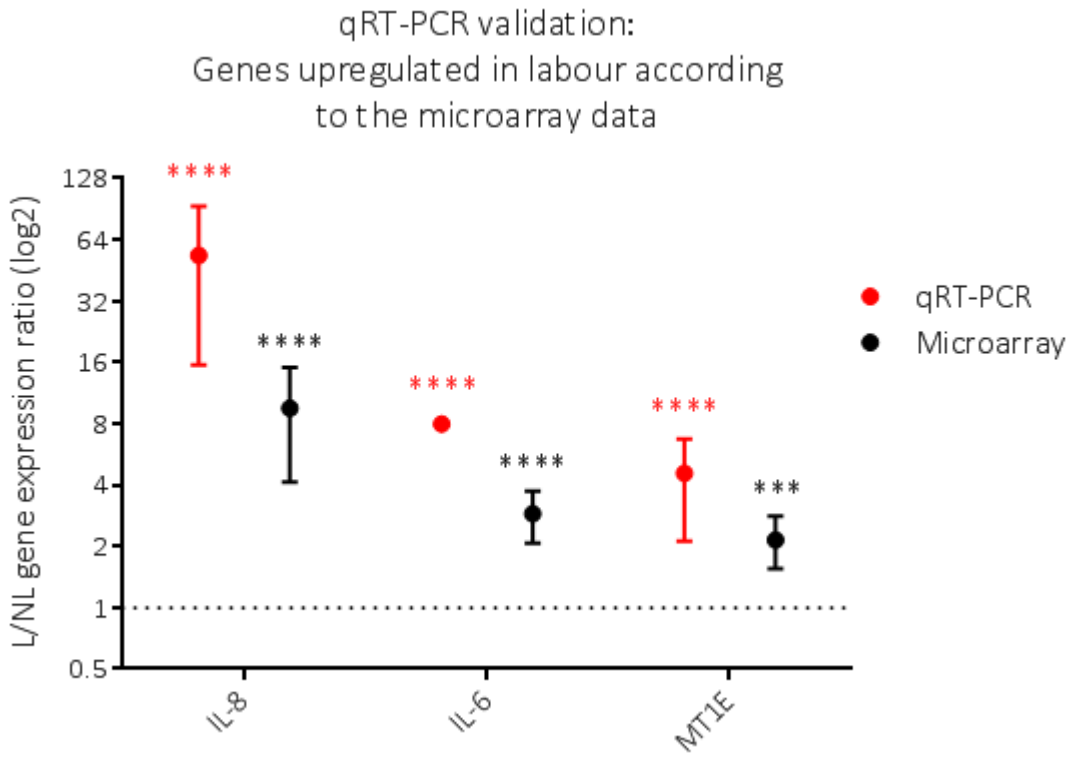


Figure 3.8. qRT-PCR validation of Interleukin-8, Interleukin-6 and Metallothionein 1E. PCR results are shown in red and microarray results are shown in black. Error bars indicate 95% confidence intervals of fold changes. Stars indicate the level of significance according to t-tests of the delta CT values.

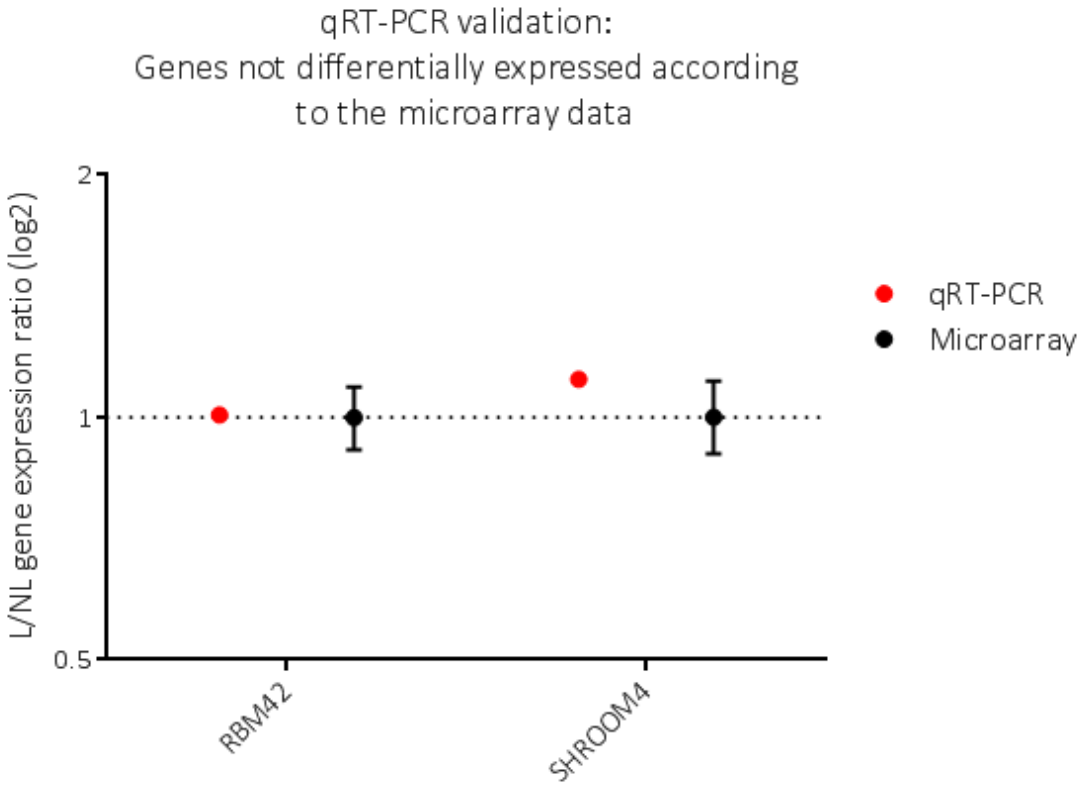


Figure 3.9. qRT-PCR validation of RNA binding motif protein 42 and Shroom family member 4. PCR results are shown in red and microarray results are shown in black. Error bars indicate 95% confidence intervals of fold changes. Stars indicate the level of significance according to t-tests of the delta CT values.

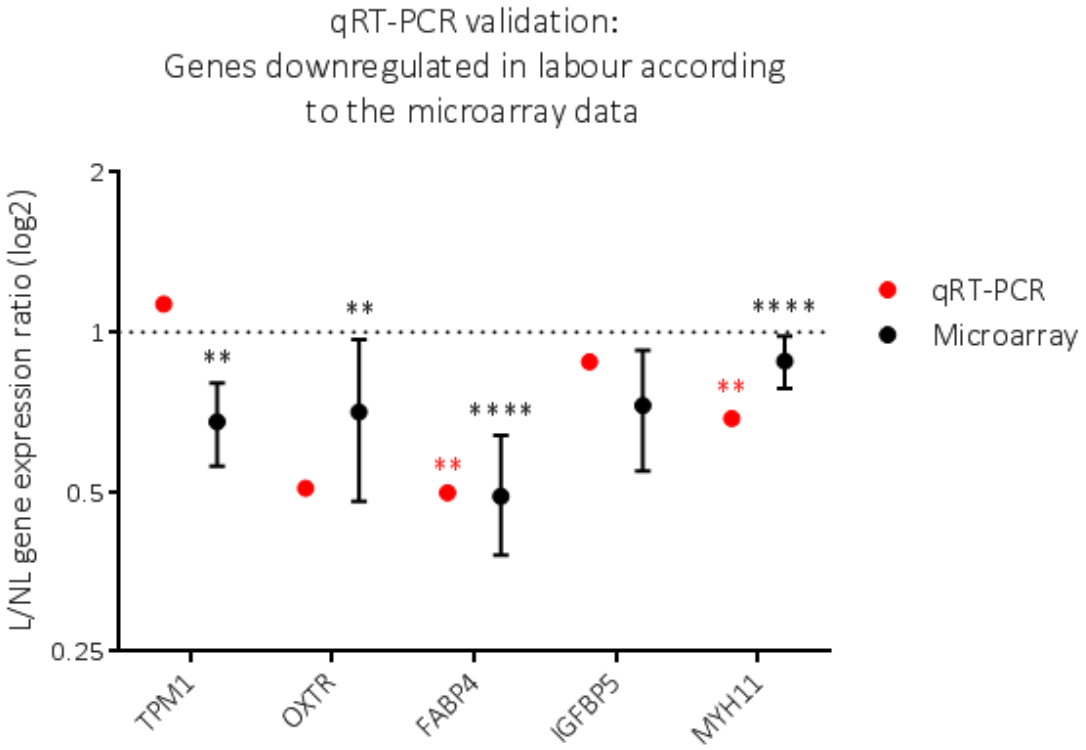


Figure 3.10. qRT-PCR validation of Tropomyosin 1, Oxytocin receptor, Fatty acid binding protein 4, Insulin-like growth factor binding protein 5 and Myosin heavy chain 11. PCR results are shown in red and microarray results are shown in black. Error bars indicate 95% confidence intervals of fold changes. Stars indicate the level of significance according to t-tests of the delta CT values.

3.3.6 Functional analysis

The Functional Classification Tool within DAVID was used to assess whether the up and down regulated gene lists (on the accompanying CD) were enriched for any biological process gene ontology (GO) terms and KEGG pathways. The same method was also used to analyse lists of genes comprising each MCL cluster in the network graph.

3.3.6.1 Gene ontology

3.3.6.1.1 Full gene lists

Table 3.5 shows the top five clusters of GO terms associated with the genes upregulated in labour. Upregulated biological processes are mainly associated with inflammation, cell (probably leukocyte) movement/migration, response to hormone stimulus, regulation of cell death and angiogenesis. Table 3.6 shows the top five clusters of GO terms associated with the genes downregulated in labour. Downregulated biological processes are mainly

associated with the response to hormone stimulus, muscle development, cytoskeleton organisation, ion homeostasis, neuron development and cell adhesion.

Table 3.5. The top five clusters of GO terms associated with genes found to be upregulated in labour.

DAVID cluster	Cluster enrichment score	GO Term	Genes linked to this term	% of total genes in the list (923)	P Value
1	15.07	GO:0006952~defense response	91	9.86	7E-18
		GO:0009611~response to wounding	78	8.45	4E-15
		GO:0006954~inflammatory response	57	6.18	2E-14
2	8.5	GO:0042330~taxis	35	3.79	8E-12
		GO:0006935~chemotaxis	35	3.79	8E-12
		GO:0007626~locomotory behavior	38	4.12	5E-07
		GO:0007610~behavior	52	5.63	3E-06
3	5.87	GO:0010033~response to organic substance	96	10.40	1E-15
		GO:0048545~response to steroid hormone stimulus	28	3.03	8E-06
		GO:0009719~response to endogenous stimulus	41	4.44	3E-04
		GO:0009725~response to hormone stimulus	38	4.12	3E-04
		GO:0043627~response to estrogen stimulus	14	1.52	5E-03
4	5.16	GO:0043066~negative regulation of apoptosis	51	5.53	9E-10
		GO:0043069~negative regulation of programmed cell death	51	5.53	2E-09
		GO:0060548~negative regulation of cell death	51	5.53	2E-09
		GO:0042981~regulation of apoptosis	81	8.78	2E-07
		GO:0043067~regulation of programmed cell death	81	8.78	3E-07
		GO:0010941~regulation of cell death	81	8.78	3E-07
		GO:0006916~anti-apoptosis	31	3.36	1E-06
		GO:0043065~positive regulation of apoptosis	38	4.12	5E-03
		GO:0043068~positive regulation of programmed cell death	38	4.12	6E-03
		GO:0010942~positive regulation of cell death	38	4.12	6E-03
		GO:0006917~induction of apoptosis	27	2.93	3E-02
		GO:0012502~induction of programmed cell death	27	2.93	3E-02
5	4.92	GO:0046651~lymphocyte proliferation	13	1.41	2E-06
		GO:0070661~leukocyte proliferation	13	1.41	3E-06
		GO:0032943~mononuclear cell proliferation	13	1.41	3E-06
		GO:0045321~leukocyte activation	33	3.58	4E-06
		GO:0046649~lymphocyte activation	29	3.14	5E-06

	GO:0001775~cell activation	36	3.90	1E-05
	GO:0042110~T cell activation	19	2.06	2E-04
	GO:0042098~T cell proliferation	8	0.87	5E-04

Table 3.6. The top five clusters of GO terms associated with genes found to be downregulated in labour.

DAVID cluster	Cluster enrichment score	GO Term	Genes linked to this term	% of total genes in the list (749)	P Value
1	4.24	GO:0048545~response to steroid hormone stimulus	26	3.47	2.5E-07
		GO:0031960~response to corticosteroid stimulus	12	1.60	6.2E-04
		GO:0051384~response to glucocorticoid stimulus	11	1.47	1.2E-03
2	4.09	GO:0009725~response to hormone stimulus	40	5.34	3.7E-08
		GO:0009719~response to endogenous stimulus	41	5.47	1.8E-07
		GO:0048545~response to steroid hormone stimulus	26	3.47	2.5E-07
		GO:0043627~response to estrogen stimulus	17	2.27	4.8E-06
		GO:0010033~response to organic substance	55	7.34	9.3E-06
		GO:0032355~response to estradiol stimulus	11	1.47	5.3E-05
		GO:0032870~cellular response to hormone stimulus	16	2.14	3.2E-04
		GO:0043434~response to peptide hormone stimulus	16	2.14	1.5E-03
		GO:0032868~response to insulin stimulus	9	1.20	5.0E-02
		GO:0032869~cellular response to insulin stimulus	7	0.93	5.7E-02
		GO:0008286~insulin receptor signaling pathway	4	0.53	1.9E-01
3	3.67	GO:0007517~muscle organ development	26	3.47	1.5E-06
		GO:0014706~striated muscle tissue development	15	2.00	3.3E-04
		GO:0060537~muscle tissue development	15	2.00	5.4E-04
		GO:0007519~skeletal muscle tissue development	10	1.34	1.3E-03
		GO:0060538~skeletal muscle organ development	10	1.34	1.3E-03
4	3.67	GO:0007517~muscle organ development	26	3.47	1.5E-06
		GO:0014706~striated muscle tissue development	15	2.00	3.3E-04
		GO:0060537~muscle tissue development	15	2.00	5.4E-04
		GO:0007519~skeletal muscle tissue development	10	1.34	1.3E-03

		GO:0060538~skeletal muscle organ development	10	1.34	1.3E-03
5	3.67	GO:0007010~cytoskeleton organization	37	4.94	4.2E-05
		GO:0030036~actin cytoskeleton organization	24	3.20	4.8E-05
		GO:0030029~actin filament-based process	24	3.20	1.3E-04
		GO:0007015~actin filament organization	9	1.20	8.5E-03

3.3.6.1.2 Network graph MCL clusters

MCL clusters (from Figure 3.6) and the main GO terms describing their functions are summarised in Table 3.7. One of the largest clusters (MCL2) is composed mostly of genes upregulated in labour and associated with inflammation. MCL14, MCL17, MCL21 and all further MCL clusters were too small to undergo functional enrichment.

Table 3.7. MCL clusters and the main GO terms describing their functions.

MCL Cluster	Number of probesets in cluster	Number of genes in cluster	Main functional category based on gene ontology (GO)	Expression appears higher in labouring samples	Expression appears higher in non-labouring samples	No differential expression
1	209	205	RNA processing	✔		
2	147	147	Inflammation	✔		
3	80	70	Transcription			✔
4	53	53	RNA processing	✔		
5	44	44	Cell death			✔
6	42	33	Transcription			✔
7	42	35	Transcription			✔
8	40	39	Protein transport		✔	
9	38	37	Transcription	✔		
10	37	36	Steroid biosynthesis	✔		
11	31	31	RNA processing			✔
12	31	31	Cell cycle	✔		
13	30	30	Cell cycle	✔		
15	25	21	Cell death			✔
16	24	24	Inflammation	✔		
18	23	23	Inflammation	✔		
19	21	21	Protein phosphorylation	✔		
20	20	20	Cell death	✔		
22	17	17	RNA processing	✔		

3.3.6.2 KEGG PATHWAYS

Pathways upregulated in labour include NOD-like receptor signalling (Figure 3.11), MAPK activation and cytokine-cytokine receptor signalling. These are all important inflammatory pathways. Pathways downregulated in labour include vascular smooth muscle contraction (Figure 3.12) and calcium signalling. This may represent a change in the composition (effectively dilution in quantity) of the myometrial tissue due to leukocyte infiltration at labour.

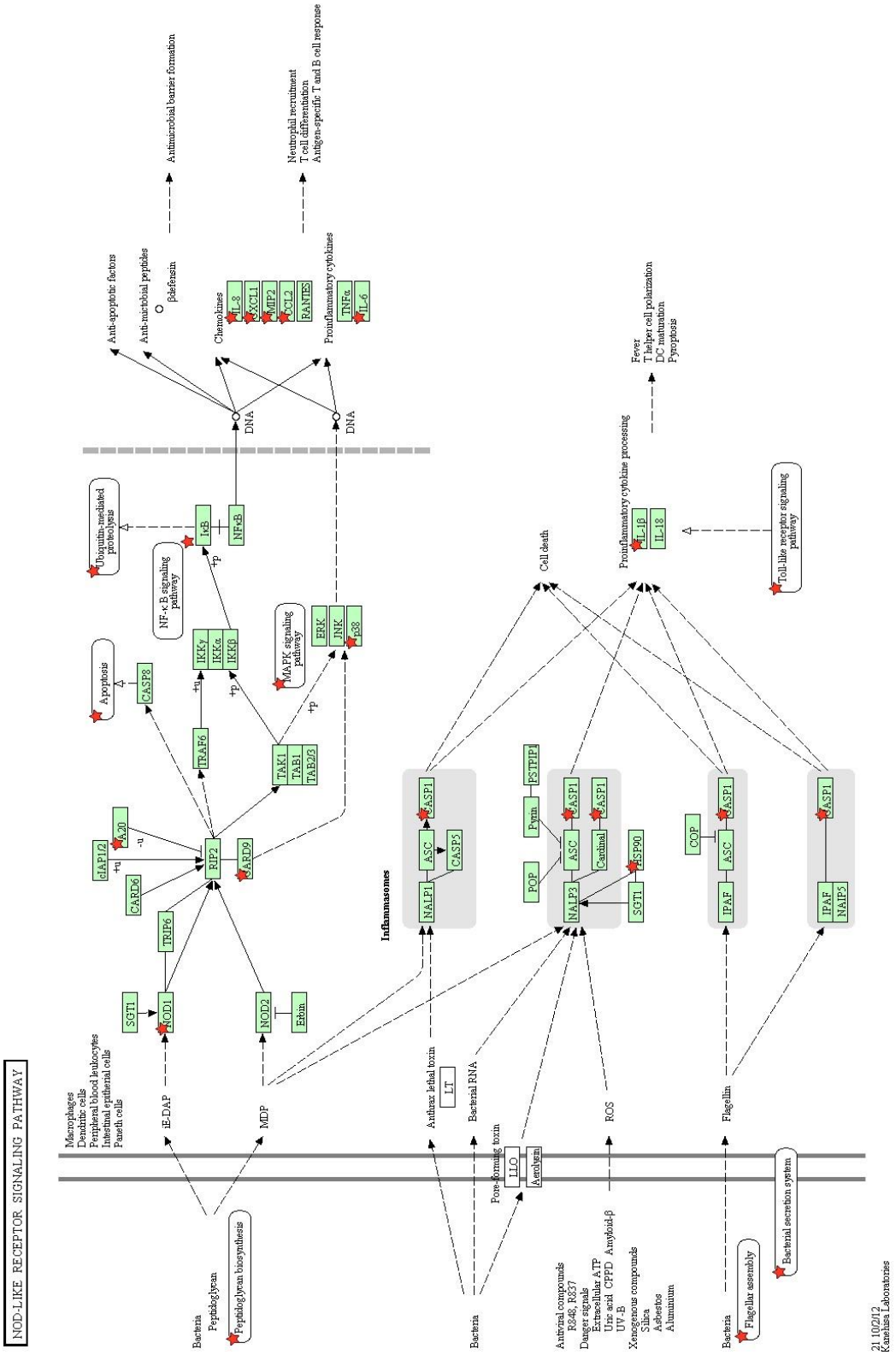


Figure 3.11. The KEGG diagram for the Nod-like Receptor Signalling Pathway. Red stars indicate genes found to be upregulated in labour.

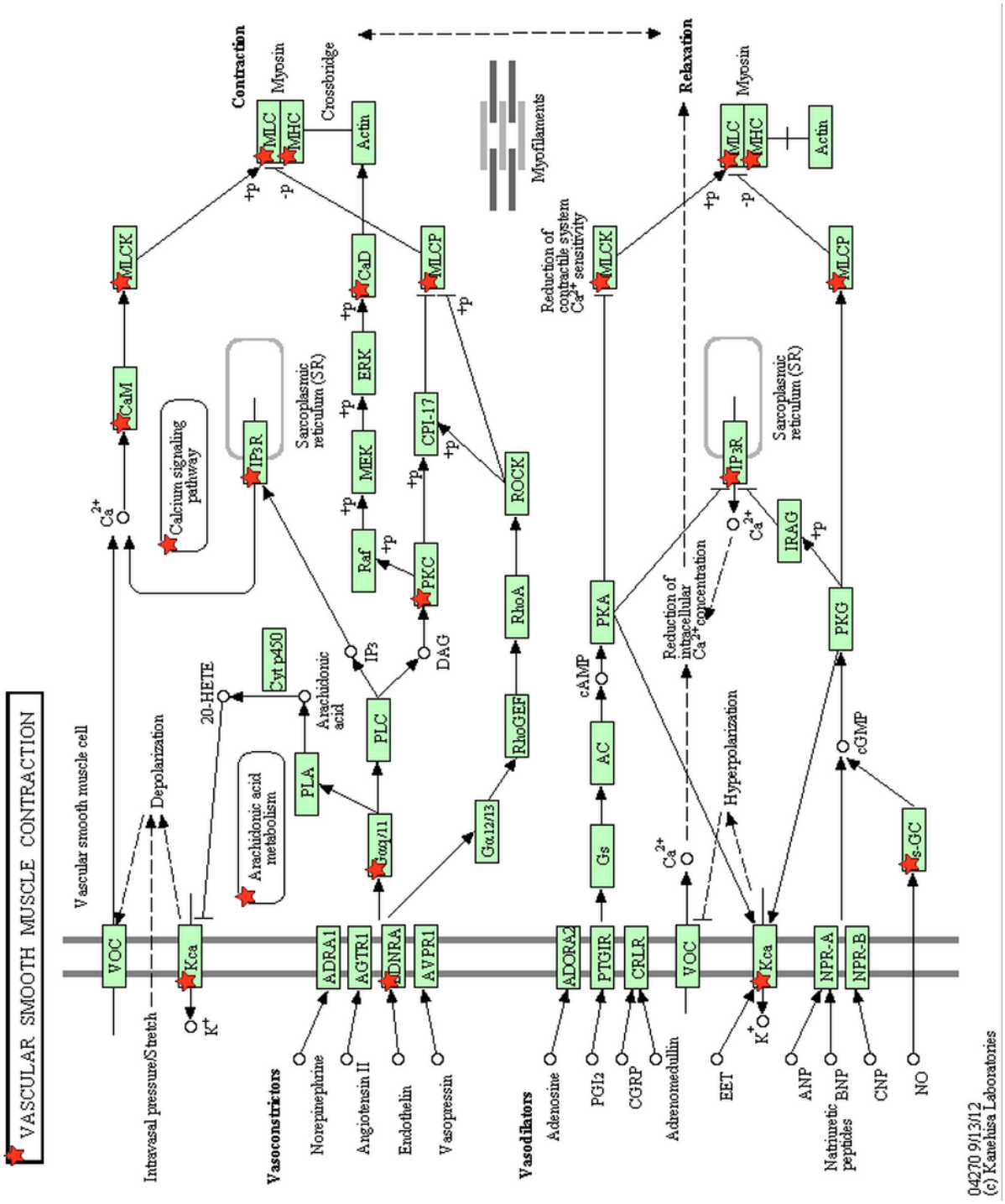


Figure 3.12. The KEGG pathway diagram for (vascular) smooth muscle contraction. Red stars indicate genes found to be downregulated in labour.

3.3.7 Assessing changes in tissue composition

Functional analysis showed that labour is associated with an upregulation of inflammatory genes/processes and a downregulation of genes/processes involved in smooth muscle contraction. It was hypothesised that both of these observations may be in part due to a change in the myometrial tissue; specifically that more leukocytes are present in labouring compared to non-labouring tissue and this "dilutes" the relative contribution of the smooth muscle cells. To illustrate this argument, the $\log^2$ mean ratio of labouring to non-labouring expression (including confidence intervals calculated using Fieller's theorem) was plotted for:

- **Smooth muscle markers** (Figure 3.13a) – Calponin 1 (CNN1), Desmin (DES), Transgelin (TAGLN), Tropomyosin 1 (TPM1) and Myosin heavy chain 11 (MYH11) are all considered classical markers of smooth muscle myofilament.(220) If there is a large influx of leukocytes into the myometrium at labour, these smooth muscle markers are expected to be downregulated.
- **Leukocyte markers** (Figure 3.13b)–CD14, CD68 and CD163 were the only Cluster of Differentiation genes expressed above background level. These are all macrophage markers.(221) The classic pan-leukocyte marker, CD45, was not expressed above background level (using 4/4 probes), although this does not suggest that leukocytes are absent. If there is a large increase in the number of leukocytes in the myometrium at labour, these leukocyte markers are expected to be upregulated.
- **Genes reported in the literature as associated with smooth muscle contraction** (Figure 3.13c) – Caldesmon 1 (CALD1), Calmodulin 1 (CALM1), and Myosin light chain kinase (MYLK) are key proteins in smooth muscle contraction and Protein phosphatase 1 regulatory subunit 12A (PPP1R12A) is a subunit of myosin phosphatase, which plays a key role in smooth muscle relaxation.(222) Expression of these contraction-associated proteins might be expected to increase in labour. Connexin43 (Cx43), is a component of gap junctions and important for synchronised contraction. Cx43 transcript increases dramatically in labour in the rat, but no such increase is seen in humans.(56)
- **Inflammatory cytokines** (Figure 3.13d)–Interleukin 1β (IL-1β), Interleukin 6 (IL-6), Interleukin 8 (IL-8), Interleukin 10 (IL-10). Expression of the pro-inflammatory

cytokines IL-1β , IL-6 and IL-8 have all been previously identified in the myometrium, largely restricted to leukocytes,(5) and upregulated in labour compared to non-labour.(6,55,115,142,223) The anti-inflammatory cytokine, IL-10, has previously been identified as upregulated in labour in the choriodecidua and amniotic fluid,(224,225) but not the myometrium.(115)

- **Genes reported in the literature as associated with myometrial activation** (Figure 3.13e) – Estrogen receptor alpha (ERα), Oxytocin receptor (OXTR), Prostaglandin E synthase 2 (PTGES2), and Cyclooxygenase 2 (COX2), have all been shown to be upregulated in the myometrium at labour.(26,33,226,227)

This analysis showed that leukocyte markers are upregulated and smooth muscle cell markers are downregulated at labour. Pro-inflammatory cytokines are highly upregulated, but genes classically associated with smooth muscle contraction are downregulated. Most genes classically associated with myometrial activation or stimulation are apparently not differentially expressed, which may be an effect of myocyte dilution (in other words, the myocytes do express greater amounts of activation stimulation markers, but their apparent density is diluted by the larger number of leukocytes in labouring samples, leading to an apparent lack of change in marker expression).

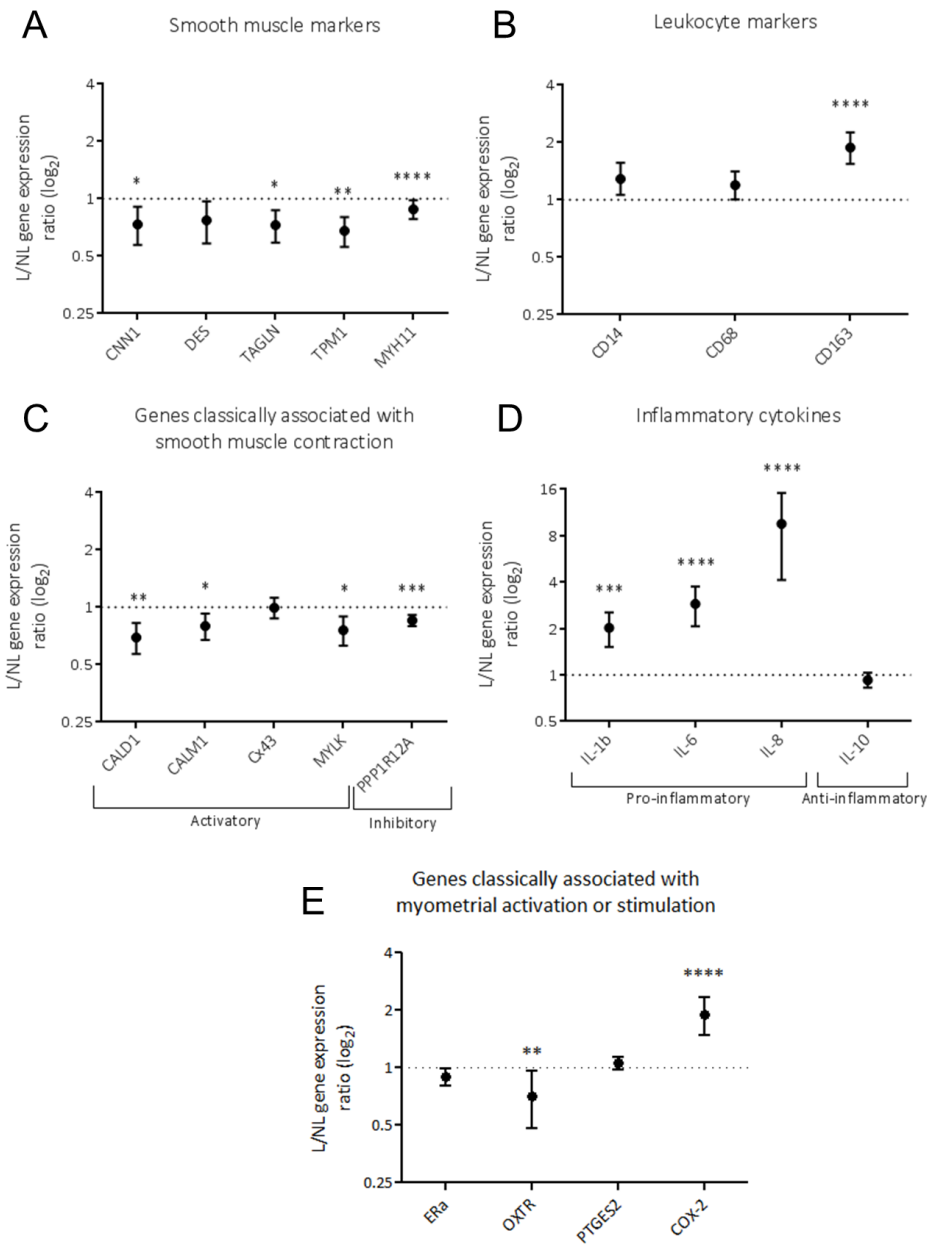


Figure 3.13. Microarray results to illustrate changes in gene expression between labouring and non-labouring groups. Error bars indicate 95% confidence intervals of the gene expression ratio between groups. Stars indicate the strength of statistical significance, calculated using t-tests on labouring and non-labouring expression values.

Higher numbers of leukocytes in labouring compared to non-labouring myometrium has been demonstrated previously (4). To confirm this in these samples, immunohistochemistry was carried out on 10 randomly selected myometrial samples (5 labouring, 5 non-labouring) from the same patients who were included in the microarray. Samples were stained for the leukocyte marker CD45 as described in Chapter 2. Although

there was a high level of variation between samples, there tended to be more leukocytes in labouring samples compared to non-labouring samples (Figure 3.14).

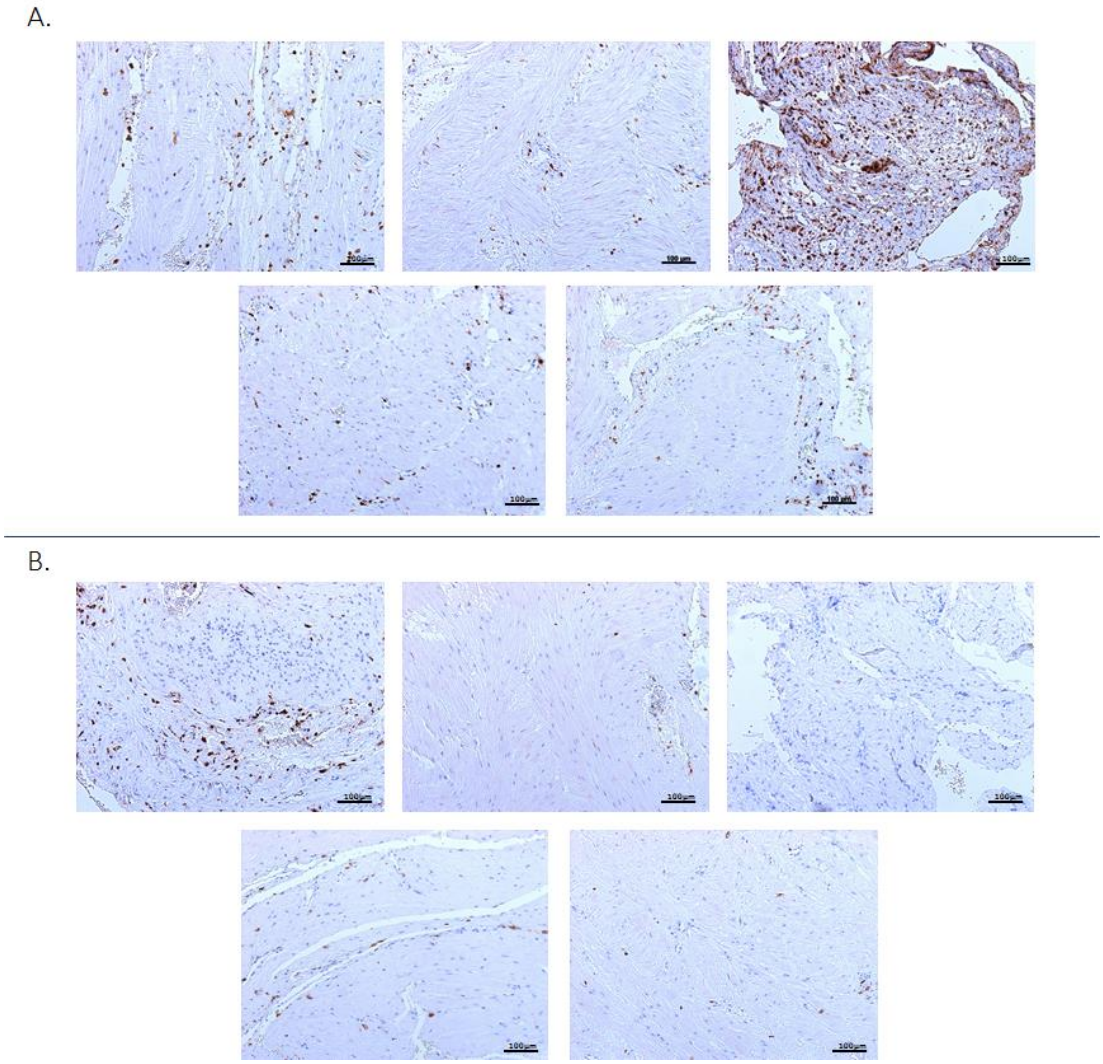


Figure 3.14. Staining for CD45 (brown) in myometrial tissue from (A) five labouring patients, and (B) five non-labouring patients. Pictures obtained using a 10x objective.

3.3.8 Assessing confounding by other patient characteristics

When observing the network graph in which nodes represent samples (previously shown in Figure 3.4), there are no significant associations between MCL cluster assignment and BMI category (P=0.2), gestation category (P=1, calculated with a Fishers exact test because cell numbers are low), and maternal age category (P=0.6, the 40+ group was combined with the 35-39 year old group to avoid invalidating the Chi-squared test due to low counts) (Figure 3.5). This suggests that these characteristics have little effect on

myometrial gene expression, however this technique is dependent on values assigned to the MCL clustering algorithm, so it does not provide strong evidence.

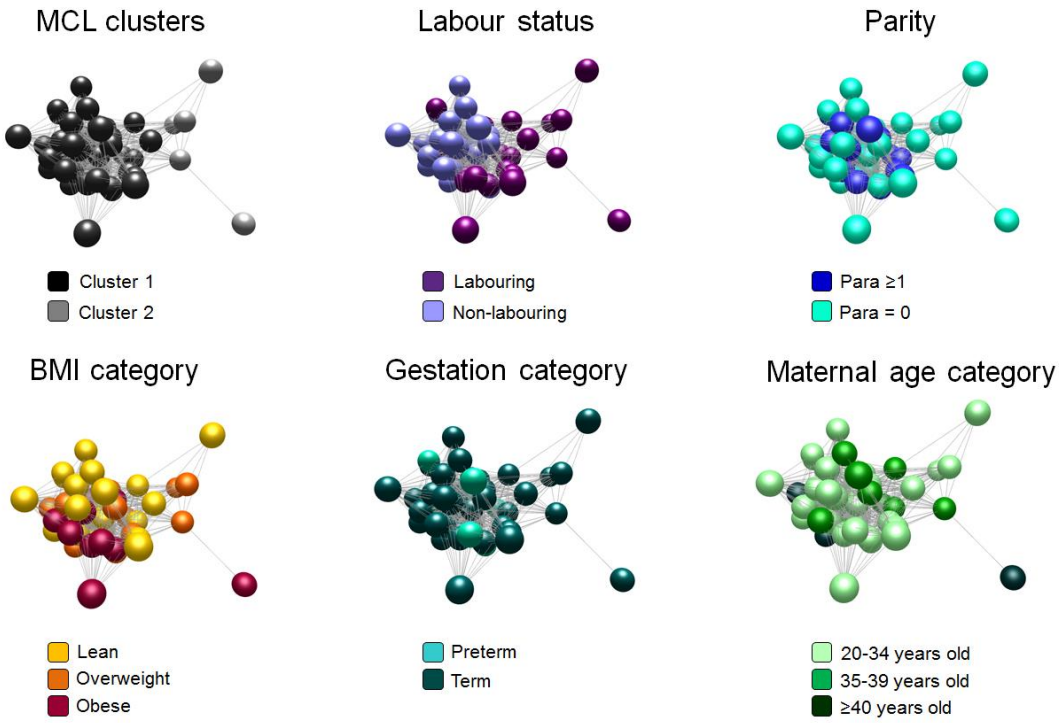


Figure 3.15. A network graph in which each node represents a different sample on the array. The same graph is coloured according to Markov chain (MCL) cluster number and several patient characteristics.

The association between MCL cluster assignment and parity category (para=0/para≥1) was not significant (P=0.1 calculated using a Fishers exact test), however all 15 para≥1 patients included on the graph appear in MCL cluster 1. This suggests that further work to confirm whether parity plays a small role in shaping the myometrial gene expression signature would be useful. In order to do this, lists of genes significantly associated only with first delivery (para=0) and only with subsequent deliveries (para≥1) were created. There were 246 genes associated with para=0 and 175 genes associated with para≥1 (the full list appears on the CD accompanying this thesis).

The number of genes that were significantly associated with parity and labour status was assessed (Figure 3.16). Although most genes are not associated with parity (i.e. they are not differentially expressed with a fold change of >1.2 or <-1.2, and an RP-PFP <0.05 between para=0 and para≥1 samples), 150 genes that are upregulated in labour are associated only with first deliveries (para=0) and 38 are associated only with subsequent (para≥1) deliveries. 24 genes that are downregulated in labour are associated with para=0 and 55 are associated with para≥1.

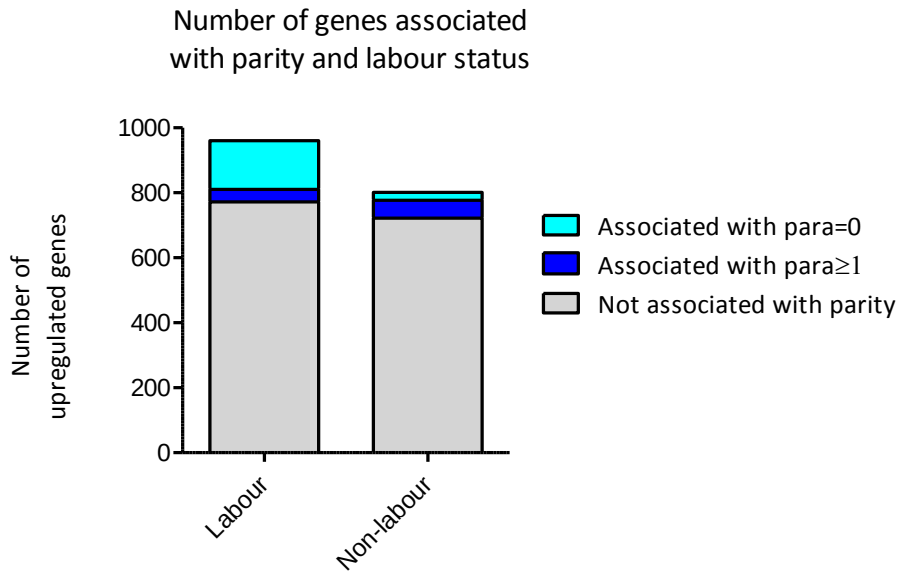


Figure 3.16. The number of genes significantly associated with parity and labour status.

Lists of genes upregulated in labour and para=0, and in labour and para≥1, were uploaded to DAVID to perform functional analyses. Table 3.8 shows the top five clusters of GO terms associated with the genes upregulated in labour and para=0. According to this, first labour is mainly associated with biological processes involved in inflammation, cell (probably leukocyte) movement/migration, responses to hormone stimulus, cytoskeleton organisation and the cell cycle. These are very similar to the biological processes associated with labour, regardless of parity (Table 3.5).

Table 3.9 shows the top 10 clusters of GO terms associated with the genes upregulated in labour and para≥1. Again, the biological processes are similar to those identified as associated with labour regardless of parity. There are also no clear differences in the

biological processes identified as associated with para=0 and those identified as associated with para≥1.

Table 3.8. The top five clusters of GO terms associated with genes found to be upregulated in labour and para=0.

DAVID cluster number	Cluster enrichment score	GO Term	Number of genes linked to this term	Percentage of total genes in the list (149)	P Value
1	14.25	GO:0006952~defense response	39	26.17	1.8E-20
		GO:0006954~inflammatory response	23	15.44	9.2E-13
		GO:0009611~response to wounding	27	18.12	1.0E-11
2	3.30	GO:0006909~phagocytosis	6	4.03	1.1E-04
		GO:0006897~endocytosis	10	6.71	3.4E-04
		GO:0010324~membrane invagination	10	6.71	3.4E-04
		GO:0016044~membrane organization	12	8.05	1.4E-03
		GO:0016192~vesicle-mediated transport	15	10.07	1.7E-03
3	3.10	GO:0010033~response to organic substance	23	15.44	2.2E-06
		GO:0048545~response to steroid hormone stimulus	10	6.71	1.2E-04
		GO:0009719~response to endogenous stimulus	14	9.40	1.9E-04
		GO:0009725~response to hormone stimulus	13	8.72	2.9E-04
		GO:0031960~response to corticosteroid stimulus	5	3.36	1.0E-02
		GO:0014070~response to organic cyclic substance	5	3.36	3.2E-02
		GO:0051384~response to glucocorticoid stimulus	4	2.68	4.2E-02
4	3.09	GO:0009617~response to bacterium	9	6.04	6.7E-04
		GO:0042742~defense response to bacterium	7	4.70	8.6E-04
		GO:0050830~defense response to Gram-positive bacterium	4	2.68	9.6E-04
5	2.68	GO:0007010~cytoskeleton organization	15	10.07	1.1E-04
		GO:0030036~actin cytoskeleton organization	9	6.04	1.8E-03
		GO:0030029~actin filament-based process	9	6.04	2.7E-03
		GO:0007015~actin filament organization	4	2.68	3.5E-02

Table 3.9. The top five clusters of GO terms associated with genes found to be upregulated in labour and para≥1.

DAVID cluster number	Cluster enrichment score	GO Term	Number of genes linked to this term	Percentage of total genes in the list (37)	P Value
1	1.85	GO:0030334~regulation of cell migration	4	10.81	6.7E-03
		GO:0040012~regulation of locomotion	4	10.81	9.4E-03
		GO:0051270~regulation of cell motion	4	10.81	9.6E-03
		GO:0001558~regulation of cell growth	4	10.81	9.7E-03
		GO:0042127~regulation of cell proliferation	6	16.22	3.2E-02
		GO:0040008~regulation of growth	4	10.81	4.2E-02
2	1.68	GO:0009617~response to bacterium	4	10.81	9.6E-03
		GO:0006955~immune response	6	16.22	1.9E-02
		GO:0006952~defense response	5	13.51	5.0E-02
3	1.54	GO:0007204~elevation of cytosolic calcium ion concentration	4	10.81	2.0E-03
		GO:0010627~regulation of protein kinase cascade	5	13.51	2.4E-03
		GO:0051480~cytosolic calcium ion homeostasis	4	10.81	2.4E-03
		GO:0006874~cellular calcium ion homeostasis	4	10.81	8.3E-03
		GO:0055074~calcium ion homeostasis	4	10.81	8.9E-03
		GO:0006875~cellular metal ion homeostasis	4	10.81	1.0E-02
		GO:0055065~metal ion homeostasis	4	10.81	1.1E-02
		GO:0008284~positive regulation of cell proliferation	5	13.51	1.4E-02
		GO:0030005~cellular di-, tri-valent inorganic cation homeostasis	4	10.81	1.5E-02
		GO:0055066~di-, tri-valent inorganic cation homeostasis	4	10.81	1.7E-02
		GO:0030003~cellular cation homeostasis	4	10.81	2.0E-02
		GO:0019725~cellular homeostasis	5	13.51	2.1E-02
		GO:0055080~cation homeostasis	4	10.81	2.7E-02
		GO:0009967~positive regulation of signal transduction	4	10.81	2.9E-02
		GO:0042127~regulation of cell proliferation	6	16.22	3.2E-02
		GO:0010647~positive regulation of cell communication	4	10.81	3.9E-02
		GO:0007166~cell surface receptor linked signal transduction	9	24.32	5.3E-02
		GO:0006873~cellular ion homeostasis	4	10.81	5.3E-02
		GO:0055082~cellular chemical homeostasis	4	10.81	5.5E-02
		GO:0050801~ion homeostasis	4	10.81	6.6E-02
		GO:0042592~homeostatic process	5	13.51	9.0E-02
		GO:0007186~G-protein coupled receptor	6	16.22	1.1E-01

		protein signaling pathway			
		GO:0048878~chemical homeostasis	4	10.81	1.1E-01
		GO:0043085~positive regulation of catalytic activity	4	10.81	1.2E-01
		GO:0044093~positive regulation of molecular function	4	10.81	1.5E-01
		GO:0030030~cell projection organization	3	8.11	2.1E-01
		GO:0007242~intracellular signaling cascade	4	10.81	5.6E-01
4	1.28	GO:0007610~behavior	5	13.51	2.1E-02
		GO:0042330~taxis	3	8.11	5.2E-02
		GO:0006935~chemotaxis	3	8.11	5.2E-02
		GO:0007626~locomotory behavior	3	8.11	1.3E-01
5	1.26	GO:0010627~regulation of protein kinase cascade	5	13.51	2.4E-03
		GO:0032101~regulation of response to external stimulus	4	10.81	5.6E-03
		GO:0032103~positive regulation of response to external stimulus	3	8.11	9.4E-03
		GO:0012501~programmed cell death	6	16.22	1.2E-02
		GO:0008284~positive regulation of cell proliferation	5	13.51	1.4E-02
		GO:0008219~cell death	6	16.22	2.3E-02
		GO:0016265~death	6	16.22	2.3E-02
		GO:0009967~positive regulation of signal transduction	4	10.81	2.9E-02
		GO:0009611~response to wounding	5	13.51	3.2E-02
		GO:0042127~regulation of cell proliferation	6	16.22	3.2E-02
		GO:0010647~positive regulation of cell communication	4	10.81	3.9E-02
		GO:0006915~apoptosis	5	13.51	4.7E-02
		GO:0010740~positive regulation of protein kinase cascade	3	8.11	5.6E-02
		GO:0051345~positive regulation of hydrolase activity	3	8.11	6.3E-02
		GO:0048584~positive regulation of response to stimulus	3	8.11	1.0E-01
		GO:0042981~regulation of apoptosis	5	13.51	1.1E-01
		GO:0043067~regulation of programmed cell death	5	13.51	1.1E-01
		GO:0010941~regulation of cell death	5	13.51	1.1E-01
		GO:0043085~positive regulation of catalytic activity	4	10.81	1.2E-01
		GO:0044093~positive regulation of molecular function	4	10.81	1.5E-01
		GO:0051336~regulation of hydrolase activity	3	8.11	1.8E-01
		GO:0043066~negative regulation of apoptosis	3	8.11	1.9E-01
		GO:0043069~negative regulation of programmed cell death	3	8.11	2.0E-01

	GO:0060548~negative regulation of cell death	3	8.11	2.0E-01
	GO:0043065~positive regulation of apoptosis	3	8.11	2.6E-01
	GO:0043068~positive regulation of programmed cell death	3	8.11	2.6E-01
	GO:0010942~positive regulation of cell death	3	8.11	2.6E-01

3.3.8.1 Assessing confounding within the labouring group

Another network graph was produced using only the labouring samples. This allows assessment of the impact on myometrial gene expression of patient characteristics that are only relevant to the labouring group. Figure 3.17 shows the network graph coloured according to MCL cluster number, whether the patient was administered prostaglandins/syntocinon, and rough stage of labour (assessed as 'early' if cervical dilation was ≤3cm, and 'late' if cervical dilation was ≥6cm). Fisher's exact or χ^2 tests showed no significant association between MCL cluster membership and either stage of labour (P=0.6) or induction/augmentation status (P=0.7).

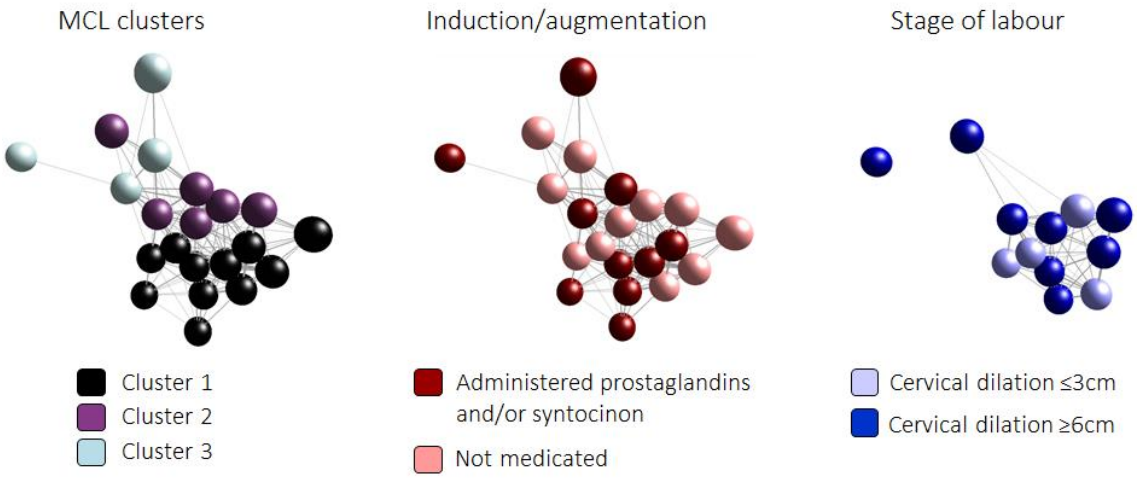


Figure 3.17. A network graph in which each node represents a different *labouring* sample on the array. The same graph is coloured according to Markov chain (MCL) cluster number and two patient characteristics: induction/augmentation of labour status and rough stage of labour. There appears to be fewer nodes in the 'Stage of labour' graph because nodes representing patients with cervical dilation between 3 and 6cm are hidden, as are nodes representing patients for whom cervical dilation was not recorded.

3.3.9 Microarray meta-analysis

3.3.9.1 Literature search

A literature search was conducted to identify microarrays comparing labouring and non-labouring myometrial gene expression.

Table 3.10 shows the number of results for keyword searches in online literature repositories.

Table 3.10. Number of results for keyword searches in online repositories of abstracts and microarray data.

	Abstract repositories				Microarray data repositories	
Search terms	**PubMed**	**Google Scholar**	**Web of Science**	**Scopus**	**ArrayExpress**	**Gene Expression Omnibus**
myometrium + microarray	76	3790	145	123		
"uterine smooth muscle" + microarray	6	659	15	11		
myometrium					31	37
myometrial					17	3
“uterine smooth muscle”					31	2

3.3.9.2 Selected studies

The 10 eligible studies are outlined in Table 3.11. Although ten studies were identified, raw data was only available for three: Bukowski 2006 (55), O’Brien 2008 (119), and Weiner 2010 (115). These studies were included in the meta-analysis, along with the study conducted as part of this PhD, referred to as “Sharp 2013”. This would equate to 27 extra (not including Sharp 2013) samples, however when the authors of the O’Brien study (119) supplied the data files, they confirmed that two samples (one term in labour and one term not in labour) had been removed from their analysis because they “deviated from the others”. These samples were also omitted from the meta-analysis. Therefore the meta-analysis used 29 samples (15 labouring, 14 non-labouring), plus the 48 used in Sharp 2013, giving a total of 77 samples (37 labouring, 40 non-labouring).

Table 3.11. Studies eligible for the meta-analysis.

First author	Microarray platform	Groups compared (number of LS myometrial samples)	Inclusion in meta-analysis	
Aguan (2000)(112)	Atlas Human cDNA array blots	TNIL (3) TIL (3)	×	No reply from authors
Bethin (2003)(59)	Affymetrix Human U95A	PTNIL (3) PTIL (3) TIL (3)	×	No reply from authors
Charpigny (2003)(114)	Atlas Human 1.2 and 1.2II nylon cDNA expression macroarrays	PTNIL (4) TNIL (4) TIL (4)	×	Little information available regarding unusual platform
Havelock (2005)(138)	Incyte Human UniGEM V .14	NIL (6) TIL (4)	×	No reply from authors
Bukowski (2006)(55)	Affymetrix Human U95A	TNIL (6) TIL (7)	✓	Raw data available in online repository
Esplin (2005)(141)	Unknown (cDNA hybridized to two glass chips and imaged using the Molecular Dynamics Generation III scanner)	TNIL (5) TIL (5)	×	Little information available regarding unusual platform
O'Brien (2008)(119)	Applied 130 Biosystems Genome Survey Microarray (version 2)	TNIL (3) TIL (3)	✓	Authors supplied preprocessed data on request
Bollapragada (2009)(223)	Affymetrix Human U133 + 2.0	TNIL (9) TIL (9)	×	Original raw data not available
Mittal (2010)(140)	Illumina Human HT-12 Expression Bead-chip	TNIL (20) TIL (19)	×	Authors declined to supply data
Weiner (2010)(115)	Affymetrix Human U133 + 2.0	PTNIL (3) PTIL (3) TNIL (3) TIL (3)	✓	Raw data available in online repository

TNIL, Term not in labour; TIL, Term in labour; PTNIL, Preterm not in labour; PTIL, Preterm in labour; NIL, Not in labour; LS, lower segment.

3.3.9.3 PREPROCESSING

The Weiner and Bukowski datasets were preprocessed in the same way as Sharp 2013 – by subjecting the data to RMA background correction before quantile normalisation to remove non-biological variation. Figure 3.18 shows boxplots to visualise the distribution of expression values across all probes and all samples before (A and C) and after (B and D) normalisation for both studies. Compared with Sharp 2013 (Figure 3.1) there are large variations in the raw data, indicating a large amount of inter-patient variation and perhaps

lower RNA quality. O'Brien 2008 did not undergo consistent preprocessing because the authors did not supply raw data.

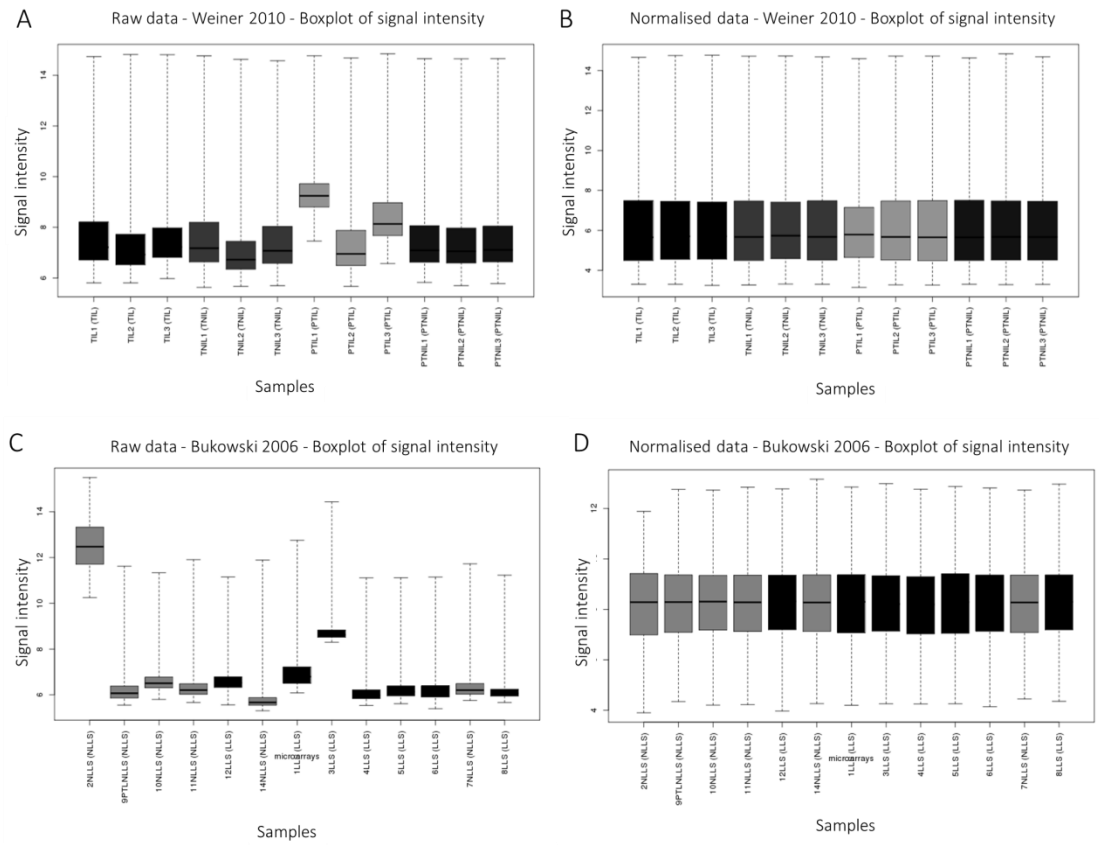


Figure 3.18. Boxplots to show the distribution of expression values across all samples in Weiner 2010 and Bukowski 2006 before and after quantile normalisation.

3.3.9.4 Assessing differential expression

Figure 3.19 shows the number of significantly (RP-PFP<0.05) differentially expressed (fold change of >1.2, <-1.2) genes in Sharp 2013, Weiner 2010 and Bukowski 2006, and the number of genes in agreement between studies. There are only 42 genes agreed to be differentially expressed (regardless of the direction of the change) in all three studies. These are listed along with their fold changes in Table 3.12.

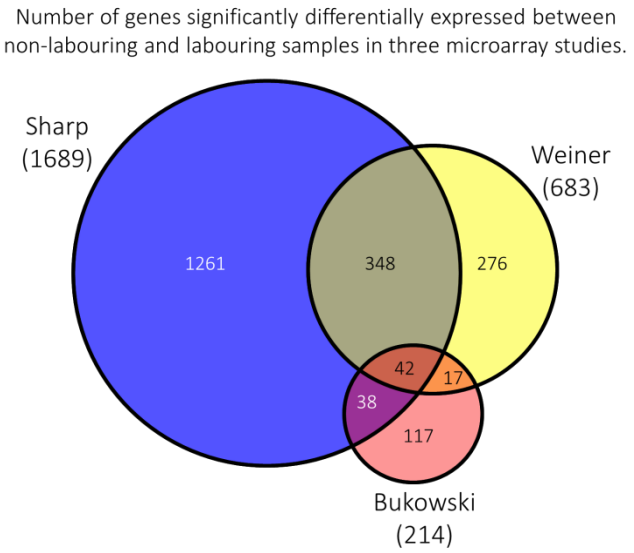


Figure 3.19. The number of genes significantly differentially expressed (regardless of the direction of the change) between non-labouring and labouring samples in three microarray studies.

Table 3.12. Fold changes of genes agreed to be significantly differentially expressed in all three studies considered in the microarray meta-analysis.

Gene ID	Gene name	Sharp 2013	Weiner 2010	Bukowski 2006	Agree on the direction of the change?
ABP1	amiloride binding protein 1	2.96	-1.47	2.92	No
AHNAK2	AHNAK nucleoprotein 2	-1.32	-1.55	-1.54	Yes
ALDH1A2	aldehyde dehydrogenase 1 family, member A2	-1.35	-1.88	-1.49	Yes
AP3S1	adaptor-related protein complex 3, sigma 1 subunit	-1.24	-1.26	-1.91	Yes
ATP2B4	ATPase, Ca++ transporting, plasma membrane 4	-1.24	-1.44	-1.70	Yes
CALD1	caldesmon 1	-1.44	-1.32	-1.50	Yes
CAV1	caveolin 1, caveolae protein, 22kDa	-1.29	-1.28	-1.45	Yes
CCL2	chemokine (C-C motif) ligand 2	4.01	3.41	1.32	Yes
CCT2	chaperonin containing TCP1, subunit 2 (beta)	1.32	1.25	-1.41	No
COL1A2	collagen, type I, alpha 2	-1.20	-1.48	-1.22	Yes
DAAM1	dishevelled associated activator of morphogenesis 1	-1.54	-1.86	-2.80	Yes
FHL1	four and a half LIM domains 1	-1.44	-1.46	-1.34	Yes
GAS1	growth arrest-specific 1	-1.26	-1.42	-1.39	Yes
HSP90AA1	heat shock protein 90kDa alpha (cytosolic), class A member 1	1.22	1.24	-1.49	No
ISCU	iron-sulfur cluster scaffold homolog (E. coli)	-1.26	-1.23	-1.67	Yes
LAMA2	laminin, alpha 2	-1.26	-1.65	-1.87	Yes

LDHA	lactate dehydrogenase A	1.38	1.38	-1.22	No
LITAF	lipopolysaccharide-induced TNF factor	1.31	1.52	-2.01	No
METTL7A	methyltransferase like 7A	-1.40	-1.57	-2.26	Yes
MFAP5	microfibrillar associated protein 5	-1.32	-1.44	-2.05	Yes
NR2F2	nuclear receptor subfamily 2, group F, member 2	-1.52	-1.46	-1.87	Yes
PARVA	parvin, alpha	-1.46	-1.31	-1.23	Yes
PGK1	phosphoglycerate kinase 1	1.35	1.35	-1.35	No
PLA2G2A	phospholipase A2, group IIA (platelets, synovial fluid)	1.64	2.71	1.49	Yes
PRG2	proteoglycan 2	2.14	1.62	4.24	Yes
PRKAR1A	protein kinase, cAMP-dependent, regulatory, type I, alpha	-1.28	-1.30	-2.27	Yes
PRSS23	protease, serine, 23	-1.21	-1.61	-1.36	Yes
RASSF2	Ras association (RalGDS/AF-6) domain family member 2	-1.41	-1.58	-2.00	Yes
RNASE1	ribonuclease, RNase A family, 1 (pancreatic)	-1.32	-1.25	-1.35	Yes
S100A8	S100 calcium binding protein A8	3.33	3.58	1.27	Yes
SPP1	secreted phosphoprotein 1	2.00	1.88	1.71	Yes
STAT3	signal transducer and activator of transcription 3	1.27	1.41	-1.66	No
SVIL	supervillin	-1.86	-1.38	-1.85	Yes
SYNM	synemin, intermediate filament protein	-1.52	-1.32	-1.89	Yes
TCEAL1	transcription elongation factor A (SII)-like 1	-1.51	-1.44	-1.63	Yes
TCEAL4	transcription elongation factor A (SII)-like 4	-1.63	-1.38	-1.28	Yes
TMEM123	transmembrane protein 123	-1.34	-1.22	-1.96	Yes
TMEM59	transmembrane protein 59	-1.21	-1.42	-1.82	Yes
TPM1	tropomyosin 1 (alpha)	-1.47	-1.22	-1.20	Yes
TUBB2A	tubulin, beta 2A	-1.44	-1.23	-1.60	Yes
VCL	vinculin	-1.44	-1.22	-1.52	Yes
ZFP36	zinc finger protein 36, C3H type, homolog (mouse)	1.71	1.24	-1.21	No

Figure 3.20 and Figure 3.21 show Forrest plots of the standardised mean difference in gene expression between labouring and non-labouring groups for each study for smooth muscle markers (markers of myofilament (220) Figure 3.20) and leukocyte markers (macrophage markers (221) Figure 3.21). In general, studies agree on the direction of the changes and heterogeneity (assessed using Tau^2 and significance of the Chi^2 statistics) is low.

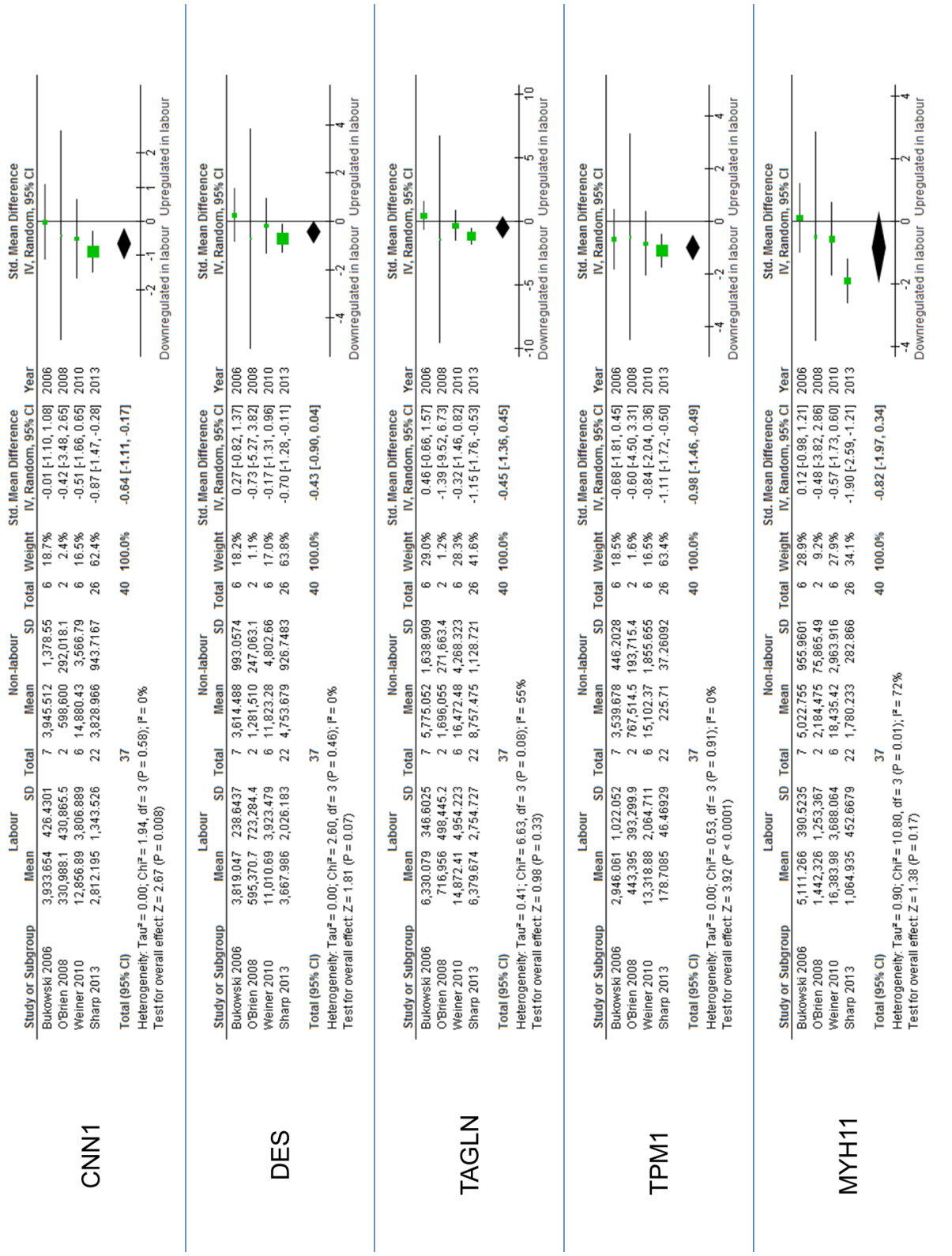


Figure 3.20. Forrest plots of the mean standardised difference in smooth muscle marker gene expression between labouring and non-labouring groups in each study.

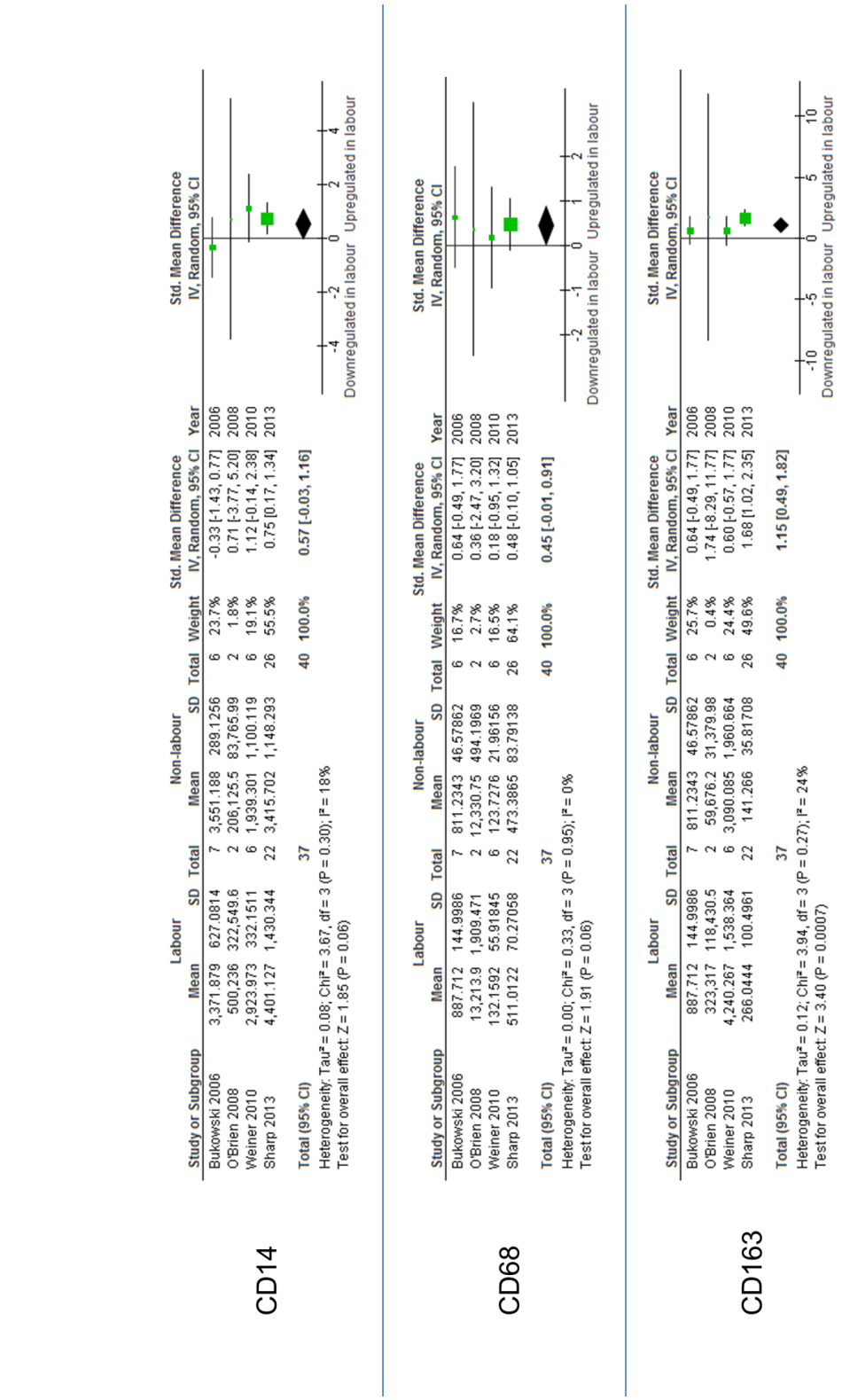


Figure 3.21. Forrest plots of the mean standardised difference in leukocyte marker gene expression between labouring and non-labouring groups in each study.

3.3.9.5 FUNCTIONAL ANALYSIS

As described above for Sharp 2013, functional analysis was performed on lists of official gene identifiers representing 1) down-regulated genes (fold change <-1.2 and RP-PFP ≤0.05) and 2) up-regulated genes (fold change >1.2 and RP-PFP ≤0.05) for each study individually (Again, O'Brien 2008 could not be used to perform these analyses). Figure 3.22 shows the number of GO terms up and down regulated in labour according to each study, and each combination of studies. 25 GO terms were agreed to be upregulated by all three studies and 52 were agreed to be downregulated. The main classes of these GO terms are summarised in Table 3.13.

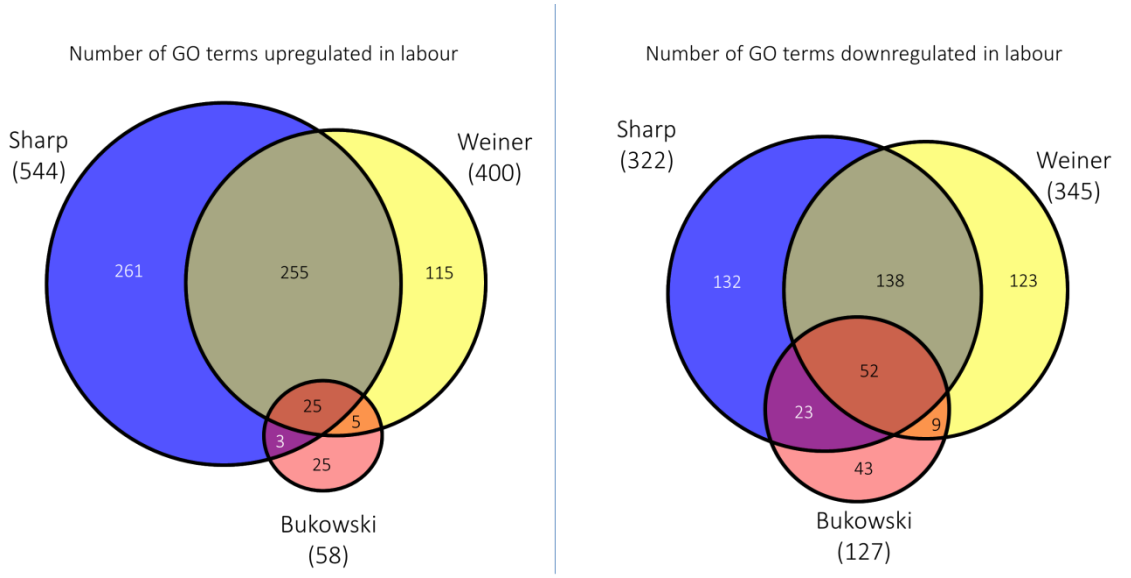


Figure 3.22. Venn diagrams to illustrate the number of GO terms upregulated and downregulated in labour according to each study.

Table 3.13. Summarised classes of the main GO terms agreed to be up or down regulated in labour by Sharp 2013, Weiner 2010 and Bukowski 2006.

Upregulated in labour	Downregulated in labour
Defense response	Response to hormone stimulus
Response to organic substance	Response to endogenous stimulus
Response to wounding	Response to steroid hormone stimulus
Inflammatory response	Response to estrogen stimulus
Regulation of cell proliferation	Muscle organ development

3.4 DISCUSSION

The myometrial transcriptome of labouring women differed considerably from that of non-labouring women. In terms of overall gene expression, non-labouring samples were

very similar to each other and different from labouring samples, but labouring samples were more varied. This suggests that we can be confident that the women supplying non-labouring samples were not in labour, and the women supplying the labouring samples were in labour but likely to be at different stages. 1761 genes were differentially expressed between the two groups, which is considerably more than identified by previous microarray experiments of this kind (115,140). Network graph and Gene Ontology analysis showed that labour was associated with an upregulation of biological processes such as inflammation and cell movement/migration, and a downregulation of muscle-specific processes and pathways such as muscle development, smooth muscle contraction and calcium signalling. The main protein pathways identified as associated with labour (Nod-like receptor signalling, MAPK activation and cytokine-cytokine receptor signalling) are all downstream or involved in crosstalk with pathways to NF-kappa B (NF-κB) activation. These observations are in agreement with the results of previous human microarray studies (115,130,140,155,228), might provide further evidence to support the hypothesis that the initiation of parturition is associated with inflammation. However, it was also hypothesised that these observations may be in part due to a change in the composition of the myometrial tissue at labour. It has previously been demonstrated that more leukocytes are present in labouring myometrium compared to non-labouring myometrium. This has been interpreted as evidence of leukocytes infiltrating the myometrium at labour (4), but an alternative explanation is that labouring samples contain more blood than non-labouring samples due to the trauma of emergency Caesarean delivery. Staining for the leukocyte marker CD45 confirmed the presence of more leukocytes in labouring compared to non-labour samples in this study. The microarray data showed that leukocyte markers were "upregulated" and smooth muscle cell markers were "downregulated" in labouring compared to non-labouring samples. Some genes that have been implicated in the initiation of labour by previous studies were not differentially expressed in this study, suggesting the contribution of myometrial smooth muscle cells may be diluted by the greater leukocyte composition. The feasibility of adjusting for this effect (for example, by normalising gene expression values to the expression of leukocyte markers) was considered. However, this was unlikely to be useful or reliable in noisy data.

The patient samples were selected to maximise the number of samples in the experiment, rather than minimise interpatient variation. Therefore, there was a high degree of patient

variation within (but not between) each group on BMI, maternal age, gestational age, parity and stage of labour (in the labouring group, assessed by cervical dilation). Network graph analysis suggested that these factors were unlikely to have a significant impact on the results; samples appeared to cluster by labour status rather than any other characteristic. However, this does not provide definitive evidence of no significant confounding. Previous studies have shown a significant difference in the myometrial gene expression signature of women in preterm and term labour (115,229). Our results did not support this finding, but the number of preterm samples in our analysis was small (4 labouring, 2 non-labouring). Parity appeared to have a small effect on myometrial gene expression, with 421 genes being significantly differentially expressed in first deliveries (para=0) compared to subsequent deliveries (para≥1). However, functional analysis did not show any clear differences in the biological processes identified as associated with para=0 and those identified as associated with para≥1.

The results were compared to the published results of similar microarrays using a meta-analysis. This is the first time microarray data to study parturition has been meta-analysed using raw, complete data. The meta-analysis showed that although different studies show limited agreement on the precise myometrial gene signature associated with labour, studies do agree that labour is associated with an "upregulation" of inflammatory processes and a "downregulation" of muscle-specific processes. No microarray study can suggest whether these results are due to leukocyte infiltration into the myometrium as a mechanism initiating labour, or due to larger numbers of leukocytes in labouring myometrium as a consequence of the trauma of emergency Caesarean section.

The main strengths of this study are its size and the comprehensive, novel methods used to analyse the data. In studies of parturition, gene expression is likely to show a particularly high degree of variation between individuals. This is not only due to genetic and environmental effects, but also because the time to the onset of spontaneous labour in non-labouring samples is unknown (55). The large sample size used in this study helps to reduce some of this bias by increasing the power of the study to detect real changes in gene expression. Additionally, network graph and meta-analytical approaches have been used for the first time to study the myometrial transcriptome at labour. The network graph approach offered new insights into the effect of patient variability on gene

expression signatures, and the meta-analysis further increased the statistical power, reliability and generalizability of results.

There are, however, some limitations to this study that must be considered when interpreting the results.

1) No "control" sample was included on all of the four microarray chips. If such a control had been included, a more accurate estimation of any batch-effect could have been made and adjusted for.

2) The qRT-PCR validation of the microarray results was conducted on the same samples used in the array. More reliable validation could have been achieved using samples from different patients, but this was impossible due to lack of availability of such samples. The large number of samples used in this study means this is less of a limitation than it is for smaller microarray experiments.

3) IHC confirmed the presence of the CD45 protein in non-labouring and labouring samples, so it is surprising that the mRNA transcript was not expressed above background. qRT-PCR for CD45 would have been useful for validating this result. Furthermore, the IHC results show a lot of variation between patients. More compelling evidence of a difference between non-labouring and labouring samples could have been achieved by examining more than one tissue section per patient, counting CD45-stained cells to give a quantitative measure of differential protein expression, and staining for other leukocyte markers.

4) The non-labouring samples were collected during elective Caesarean section that was sometimes performed because the woman had previously delivered via Caesarean section due to an underlying uterine pathology. Additionally, the labouring samples were collected during emergency Caesarean section, performed for reasons such as breech presentation, failure to progress in labour and fetal distress. Therefore, many of the patients in the study cannot be considered physiologically "normal", which raises the question of whether or not the results can be generalised to a normal population.

5) Many of the women had undergone treatment with syntocinon, prostaglandins or both to induce or augment labour. It is unclear how these treatments may effect global gene expression in the myometrium, but they are likely to contribute to the upregulation of

genes with an inflammatory role. Network graph analysis of labouring samples suggested this effect was small.

6) Parity has been categorised as para=0 for women with no previous children and para≥1 for women with one or more children, however, this information is insufficient to state that women in the para=0 group have never undergone labour before or that women in the para≥1 group have.

7) Due to time restrictions, no protein-level analyses were conducted, which would have validated the gene-level results. Some similar microarray studies have confirmed differential expression using ELISA and Western blots (119,120,140). However, when published, our data will be available in full online, providing the possibility for results to be validated at the protein-level by any other researcher.

8) The network graph approach is useful for offering insights into the effect of patient variability on gene expression and for providing useful visualisations of the data, however the MCL clustering algorithm is dependent on values assigned to its parameters. These values are defined through empirical testing to achieve a result the researcher believes to be most informative and useful. However, this introduces a level of bias and subjectivity, that should be considered when interpreting these graphs.

9) Tissue was collected from the upper flap of the lower transverse incision through the uterine wall, which is the most convenient location at the time of Caesarean. However, the upper segment is known to show increased contractility relative to the lower segment (55,230,231), making it arguably a more appropriate tissue to study. Furthermore, previous studies have shown significant differences in gene expression and molecular pathways associated with the initiation of labour between the upper and lower uterine segments (55,56), so the results of this study cannot be assumed to be relevant to the molecular events taking place throughout the whole organ (54).

10) Finally, although microarray experiments can highlight differences in gene expression between non-labouring and labouring samples, they are not able to prove that these differences cause a progression towards labour, and other hypothesised explanations (such as greater blood content of labouring samples due to the trauma of emergency

Caesarean section, or inflammation caused by vigorous myometrial contraction) remain possibilities.

In conclusion, labour appears to be characterised by an "upregulation" of inflammatory genes and a "downregulation" of muscle-specific processes in human lower segment myometrium. Both the large sample size and the meta-analysis with previous studies improve the reliability and generalisability of these results. Our results suggest that other patient characteristics such as gestational age, BMI, maternal age and parity do not to have a large effect on the myometrial transcriptome. Appropriate analysis of high throughput studies such as this are essential for improving our understanding of the molecular mechanisms underlying parturition, and will provide a basis for understanding differences between normal and dysfunctional or preterm labour.

4 A COMPUTATIONAL MODEL OF A KEY SIGNALLING PATHWAY IN INFECTION-INDUCED PRETERM LABOUR

4.1 INTRODUCTION

In addition to the high throughput data generation and analysis described in chapter two, the novel approach of systems biology also provides useful strategies to integrate complex interactions within biological systems through building computational models. Such models can be used to develop comprehensive *in silico* reproductions of "pregnant" tissues that demonstrate "emergent" properties, not obvious by conventional analysis (156). Computational models of the molecular mechanisms initiating parturition could (i) identify gaps in knowledge where additional wet lab experiments are required, and (ii) provide the basis for an *in silico* model of parturition for "testing" novel drugs to treat or prevent preterm labour. For example, a model could predict how a myometrial smooth muscle cell might respond to the inhibition of an intracellular pathway and used to inform the design of further *in vitro* experiments in the initial stages of drug design.

This chapter describes the use of published data to develop a comprehensive model of an important candidate signalling pathway in infection-induced preterm labour: that of lipopolysaccharide (LPS) -induced activation of nuclear factor kappa B (NF-κB). This is an important candidate signalling pathway in infection-induced preterm labour. The actions of LPS and NF-κB are well characterised: LPS is a gram negative bacterial endotoxin that triggers an inflammatory response in many cells including uterine smooth muscle cells (232). LPS is often used in animal and culture studies to mimic intrauterine infection which subsequently induces preterm labour (228,233–235), therefore the actions of LPS could be considered to replicate the actions of the initiator of some cases of infection-induced preterm labour. LPS binds to its receptor, Toll-Like Receptor 4 (TLR4) to trigger an innate immune response. TLR4 is expressed in the decidua, amnion and myometrium, and expression is higher at term compared to preterm (236,237). TLR4-neutralising monoclonal antibody and TLR4 antagonists have been shown to inhibit LPS- and inflammation-induced preterm labour in a mouse (238) and a primate model These studies suggest that TLR4 plays an important role in the induction of gram-negative bacteria infection-induced uterine contractility. Other toll-like receptors (such as TLR2) also appear to play a key role in parturition (236,237,239,240). TLR2 recognises

endotoxins associated with gram-positive bacteria and shows increased expression in chorion, amnion and myometrium at labour (237).

All toll-like receptor signalling pathways culminate in activation of NF-κB (241). NF-κB is a protein complex transcription factor with a particular role in the immune response to infection. It is activated in response to pro-inflammatory stimuli (242), but also regulates the transcription of inflammatory genes (243–245). NF-κB activity increases in human labour, particularly in the fetal membranes (246), but also in the myometrium, where labour is associated with an increase in the NF-κB p65-p50 herterodimer in pregnancy and labouring tissue compared to non-labouring tissue (247). In this way NF-κB may act as a feed-forward mechanism for the inflammatory events associated with labour (248). In fact, the microarray experiment in the previous chapter (in agreement with other studies (130,133,155)) identified several inflammatory pathways associated with labour, all of which are downstream of or involved in crosstalk with pathways that activate NF-κB. Therefore, in a uterine smooth muscle cell, the LPS/NF-κB signalling pathway is likely to be involved in triggering preterm labour in response to intrauterine infection.

4.2 Summary of methods

Detailed model development methods are discussed in Chapter 2. Briefly, a literature search was conducted to provide specific information on the structure and kinetic behaviour of the LPS/ NF-κB signalling pathway (described in section 2.2.3). The reasons for selecting this model target are explained in section 2.2.1). Graphical representations of the pathway that detail every reaction were drawn using two standardised graphical notations (SBGN and mEPN, described in section 2.2.4). The software tool Copasi was used to build a kinetic model of the pathway based on ordinary differential equations (ODEs, the reasons for selecting this modelling approach are discussed in section 2.2.2 and model development is discussed in section 2.2.5). The model was developed through successive rounds of running steady-state and time-course simulations, comparing outputs to published descriptions of the pathway's behaviour *in vitro*, and adjusting the model (described in section 2.2.6). An attempt was also made to test the ability of the model to recapitulate the *in vitro* behaviour of the system when either of two upstream proteins (MyD88 and TRIF) is knocked-out.

4.3 Results

4.3.1 The literature search

An extensive literature search (described in Chapter 2) provided information on the LPS-NF-κB pathway. Figure 4.1 shows a simplified overview of the canonical LPS-NF-κB p65-p50 pathway, derived from generally accepted interactions in the literature.

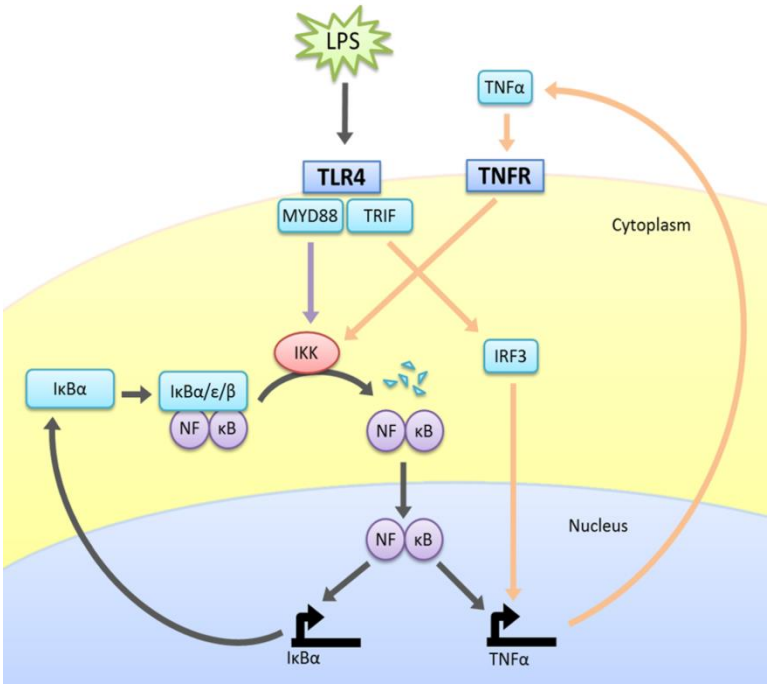


Figure 4.1. A simplified overview of the LPS-induced NF-κB signalling pathway. IκB, nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor (alpha, beta and epsilon isoforms are incorporated into the model); IKK, IκB kinase; LPS, lipopolysaccharide; MyD88, myeloid differentiation primary response gene 88; NF-κB, nuclear factor kappa B; TLR4, toll-like receptor 4; TNFα, tumour necrosis factor alpha; TNFR, tumour necrosis factor receptor; TRIF, Tir-Domain-Containing Adapter-Inducing Interferon-β.

The pathway can be split into three modules: 1) The MyD88-dependent pathway, 2) The TRIF-dependent pathway, 3) IκB-NF-κB signalling.

The reactions that describe LPS activation of its receptor Toll Like Receptor 4 (TLR4) are common to both modules 1 and 2: LPS binds to LPS-binding protein (LBP) (249) and this complex binds to and activates CD14 on the cell membrane.(250) CD14 then helps transport the LPS-LBP complex to TLR4 where LPS activates the receptor (251–254).

Further downstream signalling occurs via two pathways: one dependent on Myeloid Differentiation Primary Response Gene 88 (MyD88) and the other dependent on Tir-Domain-Containing Adapter-Inducing Interferon-β (TRIF). The MyD88-dependent pathway is largely responsible for the expression of proinflammatory cytokines (255–260), and the main role of the TRIF-dependent pathway is to induce expression of co-stimulatory molecules and IFN-inducible genes (261–264). NF-κB can be activated via either pathway, and *in vivo* the two pathways probably interact to maximise expression of inflammatory cytokines (265).

Module 1) The MyD88-dependent pathway

The MyD88-dependent pathway begins when TIRAP (toll-interleukin 1 receptor domain containing adaptor protein) is recruited to the cytoplasmic domain of TLR4 (266,267). TIRAP recruits MyD88,(266–268) which binds IRAK4 (interleukin-1 receptor-associated kinase 4) (269,270), IRAK1 (interleukin-1 receptor-associated kinase 1) (271) and TRAF6 (TNF receptor associated factor 6)(272,273). IRAK1 is phosphorylated by IRAK4 (269,274,275) and then the IRAK1/TRAF6 complex dissociates from the receptor (271,275). TRAF6 then interacts with a complex consisting of the MAPK kinase kinase transforming growth factor-beta-activated kinase 1 (TAK1) and its binding proteins TAB1 and TAB2 (276). The TRAF6/IRAK1/TAK1/TAB1/TAB2 complex is inactive until it translocates to the cytoplasm and forms a larger complex with the ubiquitin-conjugating enzymes Ubc13 and Uev1A. This ubiquitination activates TAK1 and marks IRAK1 for degradation by the proteasome (276,277). TAK1 phosphorylates an IκB kinase (IKK) complex consisting of two catalytic subunits (IKKα and IKKβ) and a regulatory subunit (IKKγ) (278–281). This phosphorylated IKK complex feeds into the IκB-NF-κB signalling module of the pathway.

Module 2) The TRIF-dependent pathway

Compared to the MyD88-dependent pathway, less is known about the TRIF-dependent module. Here we have modelled the pathway according to the suggestions put forward by Covert *et al*. (205). They observed delayed NF-κB activation in MyD88-deficient mouse embryo fibroblasts (MEFs) treated with LPS. They hypothesised that the delay occurs because the TRIF-dependent pathway requires protein synthesis before NF-κB

can be activated via a secondary signal in an autocrine manner. Microarray analysis of LPS stimulated MyD88-deficient MEFs showed significant upregulation of the tumour necrosis factor alpha (TNFα) transcript, which (along with the findings of additional validation experiments) led them to suggest that TNFα is the autocrine signal responsible for the late-phase activation of NF-κB in response to LPS.

The pathway they put forward begins when TRIF-related adapter molecule (TRAM) binds to active TLR4 (282–284) and acts as a bridging adaptor connecting TLR4 to TRIF (283,285). This complex activates two non-canonical IKKs, TANK-binding kinase 1(TBK1) and IKKε (286–288), which phosphorylate Interferon Regulatory Factor 3 (IRF3) (286,287,289). IRF3 then translocates to the nucleus and regulates transcription of TNFα. TNFα is secreted into the extracellular space where it binds to its membrane receptor TNFR1 (290). This leads to the recruitment of TRAF2, TRADD and RIP1. TRAF2 then recruits the IKK complex, which allows it to be phosphorylated by RIP1 (291–295).

Module 3) IκB-NF-κB signalling

In resting cells, NF-κB is sequestered in an inactive state in the cytoplasm by nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor (IκB). Active IKK can bind to and phosphorylate IκB, which causes it to degrade and therefore release NF-κB (288,296). Free NF-κB then translocates to the nucleus where it regulates transcription of target genes, including the IκB isoform IκBα (297–299). The production of IκBα results in a negative feedback loop that causes oscillations in NF-κB activity (204,206). Two other IκB isoforms (IκBβ and IκBε) are also included in the model. They act in the same way as IκBα, but their production is not modelled as NF-κB-dependent and their activity is in antiphase with IκBα activity. This antiphase activity acts to dampen NF-κB oscillations.

Much of the above information was derived from cell types other than uterine smooth muscle cells. There was insufficient data in the literature to confirm whether any deviations of this standard pathway occur in uterine smooth muscle cells and in the scenario of pregnancy.

4.3.2 Graphical depiction

The details of the individual molecular interactions of the pathway, which are well characterised in the experimental literature, were built into a single detailed graphical depiction of the canonical pathway. These present the topographical model according to two standardised graphical notations, Systems Biology Graphical Notation (SBGN) and Modified Edinburgh Pathway Notation (mEPN).. A key to these notations is provided in Appendix 4). These graphical representations detail every entity used in the final model, the reactions they are involved in, and in what way they react (binding, phosphorylation, etc.). They act as blueprints to the static structure of the model. Although SBGN and mEPN each have their own merits, arguably mEPN provides a clearer depiction of the model structure in this case. These are the first standardised graphical depictions of this pathway. High resolution images of these graphical depictions are included on the CD accompanying this thesis.

4.3.3 The kinetic model

4.3.3.1 Structure and kinetic parameters

A kinetic model was built using known parameters for each of the processes shown in the graphical depiction. The model reaction and kinetic equations used in the final model are listed in Table 4.1. Table 4.2 lists the initial concentrations of molecular species used in the model. The full model, encoded in SBML format, is available in the Biomodels database (http://www.ebi.ac.uk/biomodels-main/MODEL1303230000).

An extensive search of the literature retrieved no data on time course behaviour or kinetic parameter values specific to uterine smooth muscle cells, so for the IKK-NF-κB portion of the model, kinetic values were taken from the model of Hoffmann et. al which was derived from experiments on mouse embryo fibroblasts (MEFs)(204). Kinetic values for the novel reactions that we included upstream of NF-κB activation (for example LPS to IKK, and production and actions of TNFα) were not available in the literature and were therefore imputed to fit the time course NF-κB activity profile observed by Covert *et al*. (205) in LPS-treated MEFs. Data-fitting is a standard technique in computational modelling and can be achieved automatically through algorithms that find optimal values (203), however in this case, parameters were adjusted manually because data were sparse.

Table 4.1. Summary of reactions and reaction parameters used in the model.

	Reaction name	Reaction equation	Rate equation	Kinetic parameters
1	MYD881) Receptor Complex Formation	LPS + LBP + CD14 + TLR4 + MyD88 + IRAK4 + TIRAP = LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4	$v^1 = \frac{kf^1 \cdot [LPS] \cdot [LBP] \cdot [CD14] \cdot [TLR4] \cdot [MYD88] \cdot [IRAK4] \cdot [TIRAP] - kr^1 \cdot [LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4]}{}$	kf^1 0.0001 ml^6/(µmol^6*s) kr^1 0.0001 1/s
2	MYD882) IRAK1 and TRAF6 Phosphorylation	IRAK1 + TRAF6 = TRAF6:IRAK1[P]; LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4	$v^2 = \frac{kf^2 \cdot [LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4] \cdot [TRAF6] \cdot [IRAK1] - kr^2 \cdot [TRAF6:IRAK1[P]]}{}$	kf^2 0.001 ml^2/(µmol^2*s) kr^2 0.001 1/s
3	MYD883) TAK/TAB Binds to TRAF6	TRAF6:IRAK1[P] + TAK1:TAB1:TAB2 = TAK1:TAB1:TAB2:TRAF6	$v^3 = \frac{kf^3 \cdot [TRAF6:IRAK1[P]] \cdot [TAK1:TAB1:TAB2] - kr^3 \cdot [TAK1:TAB1:TAB2:TRAF6]}{}$	kf^3 0.003 ml/(µmol*s) kr^3 0.01 1/s
4	MYD884) IKK Phosphorylation by TAK1	IKK -> IKK[P]; TAK1:TAB1:TAB2:TRAF6	$v^4 = \frac{k^4 \cdot [TAK1:TAB1:TAB2:TRAF6] \cdot [IKK]}{km^4 + [IKK]}$	k^4 0.1 1/s km^4 0.1 µmol/ml
5	MYD885) IKK[P] Dephosphorylation	IKK[P] -> IKK	$v^5 = kf^5 \cdot [IKK[P]]$	kf^5 0.1 1/s
6	TRIF01) Receptor Complex Formation	LPS + LBP + CD14 + TLR4 + TRIF + TRAM + RIP1 + TBK1/IKKe = LPS:LBP:CD14:TLR4:RIP1:TRAM:TRIF:TBK1/IKKe	$v^6 = \frac{kf^6 \cdot [LPS] \cdot [LBP] \cdot [CD14] \cdot [TLR4] \cdot [TRIF] \cdot [TRAM] \cdot [RIP1] \cdot [TBK1/IKKe] - kr^6 \cdot [LPS:LBP:CD14:TLR4:TRAM:TRIF:TBK1/IKKe]}{}$	kf^6 0.0001 ml^7/(µmol^7*s) kr^6 0.0001 1/s
7	TRIF02) IRF3 Phosphorylation	IRF3 -> IRF3[P]; LPS:LBP:CD14:TLR4:RIP1:TRAM:TRIF:TBK/IKKe	$v^7 = \frac{k^7 \cdot [LPS:LBP:CD14:TLR4:RIP1:TRAM:TRIF:TBK/IKKe] \cdot [IRF3]}{km^7 + [IRF3]}$	k^7 0.1 1/s km^7 0.1 µmol/ml
8	TRIF03) IRF3 Dephosphorylation	IRF3[P] -> IRF3	$v^8 = kf^8 \cdot [IRF3[P]]$	kf^8 0.1 1/s
9	TRIF04) IRF3 Nuclear Import/Export	IRF3[P] = IRF3[P](nuc)	$v^9 = kf^9 \cdot [IRF3[P]] - kr^9 \cdot [IRF3[P](nuc)]$	kf^9 0.1 1/s kr^9 0.1 1/s
10	TRIF05) Inducible TNFa Synthesis	2 * IRF3[P](nuc) -> TNFa + 2 * IRF3[P](nuc)	$v^{10} = kf^{10} \cdot 2 \cdot [IRF3[P](nuc)]$	kf^{10} 0.02 ml/(µmol*s)
11	TRIF06) Constitutive TNFa Synthesis	source -> TNFa	$v^{11} = kf^{11} \cdot [source]$	kf^{11} 0.001 1/s
12	TRIF07)Inducible TNFa Synthesis by NFkB	2 * NFkB(nuc) -> TNFa + 2 * NFkB(nuc)	$v^{12} = kf^{12} \cdot 2 \cdot [NFkB(nuc)]$	kf^{12} 0.001 ml/(µmol*s)
13	TRIF08) TNFa Degradation	TNFa -> sink	$v^{13} = kf^{13} \cdot [TNFa]$	kf^{13} 0.1 1/s
14	TRIF09) TNFa Receptor Complex Formation	TNFa + TNFR1 + TRAF2 + TRADD + RIP1 = TNFa:TNFR1:TRAF2:TRADD:RIP1	$v^{14} = \frac{kf^{14} \cdot [TNFa] \cdot [TNFR1] \cdot [TRAF2] \cdot [TRADD] \cdot [RIP1] - kr^{14} \cdot [TNFa:TNFR1:TRAF2:TRADD:RIP1]}{}$	kf^{14} 0.1 ml^4/(µmol^4*s) kr^{14} 0.1 1/s
15	TRIF10) IKK Phosphorylation by RIP1	IKK -> IKK[P]; TNFa:TNFR1:TRAF2:TRADD:RIP1	$v^{15} = \frac{k^{15} \cdot [TNFa:TNFR1:TRAF2:TRADD:RIP1] \cdot [IKK]}{km^{15} + [IKK]}$	k^{15} 0.1 1/s km^{15} 0.1 µmol/ml
16	NFkB01) IkBa:NFkB Binding	IkBa + NFkB = IkBa:NFkB	$v^{16} = kf^{16} \cdot [IkBa] \cdot [NFkB] - kr^{16} \cdot [IkBa:NFkB]$	kf^{16} 0.5 ml/(µmol*s) kr^{16} 0.0005 1/s
17	NFkB02) IKK:IkBa:NFkB Binding (1)	IKK[P] + IkBa:NFkB = IKK[P]:IkBa:NFkB	$v^{17} = kf^{17} \cdot [IKK[P]] \cdot [IkBa:NFkB] - kr^{16} \cdot [IKK[P]:IkBa:NFkB]$	kf^{17} 0.185 ml/(µmol*s) kr^{17} 0.0125 1/s
18	NFkB03) IkBb:NFkB Binding	IkBb + NFkB = IkBb:NFkB	$v^{18} = kf^{18} \cdot [IkBb] \cdot [NFkB] - kr^{18} \cdot [IkBb:NFkB]$	kf^{18} 0.5 ml/(µmol*s) kr^{18} 0.0005 1/s
19	NFkB04) IKK:IkBb:NFkB Binding (1)	IKK[P] + IkBb:NFkB = IKK[P]:IkBb:NFkB	$v^{19} = kf^{19} \cdot [IKK[P]] \cdot [IkBb:NFkB] - kr^{19} \cdot [IKK[P]:IkBb:NFkB]$	kf^{19} 0.048 ml/(µmol*s) kr^{19} 0.00175 1/s
20	NFkB05) IkBe:NFkB Binding	IkBe + NFkB = IkBe:NFkB	$v^{20} = kf^{20} \cdot [IkBe] \cdot [NFkB] - kr^{20} \cdot [IkBe:NFkB]$	kf^{20} 0.5 ml/(µmol*s) kr^{20} 0.0005 1/s

Red italics = parameters derived during model fitting to experimental data from Covert et al. (205); *all other values* = parameters used in Hoffmann et al.'s model of NF-κB signalling (204). *v*= reaction rate, *kf*= rate of the forward reaction, *kr* = rate of the reverse reaction.

Table 4.1 continued.

	Reaction name	Reaction equation	Rate equation	Kinetic parameters
21	NFkB06) IKK:IkBe:NFkB Binding (1)	IKK[P] + IkBe:NFkB = IKK[P]:IkBe:NFkB	$v^{21} = kf^{21} \cdot [IKK[P]] \cdot [IkBe:NFkB] - kr^{21} \cdot [IKK[P]:IkBe:NFkB]$	kf^{21} 0.07 ml/(µmol*s) kr^{21} 0.00175 1/s
22	NFkB07) IKK:IkBa:NFkB Catalysis	IKK[P]:IkBa:NFkB -> IKK[P] + NFkB	$v^{22} = kf^{22} \cdot [IKK[P]:IkBa:NFkB]$	kf^{22} 0.0204 1/s
23	NFkB08) IKK:IkBb:NFkB Catalysis	IKK[P]:IkBb:NFkB -> IKK[P] + NFkB	$v^{23} = kf^{23} \cdot [IKK[P]:IkBb:NFkB]$	kf^{23} 0.0075 1/s
24	NFkB09) IKK:IkBe:NFkB Catalysis	IKK[P]:IkBe:NFkB -> IKK[P] + NFkB	$v^{24} = kf^{24} \cdot [IKK[P]:IkBe:NFkB]$	kf^{24} 0.011 1/s
25	NFkB10) IkBa:NFkB Constitutive Degradation	IkBa:NFkB -> NFkB	$v^{25} = kf^{25} \cdot [IkBa:NFkB]$	kf^{25} 2.25E-05 1/s
26	NFkB11) IkBb:NFkB Constitutive Degradation	IkBb:NFkB -> NFkB	$v^{26} = kf^{26} \cdot [IkBb:NFkB]$	kf^{26} 2.25E-05 1/s
27	NFkB12) IkBe:NFkB Constitutive Degradation	IkBe:NFkB -> NFkB	$v^{27} = kf^{27} \cdot [IkBe:NFkB]$	kf^{27} 2.25E-05 1/s
28	NFkB13) NFkB Nuclear Import/Export	NFkB = NFkB(nuc)	$v^{28} = kf^{28} \cdot [NFkB] - kr^{28} \cdot [NFkB(nuc)]$	kf^{28} 0.09 1/s kr^{28} 8.00E-05 1/s
29	NFkB14) Nuclear IkBa:NFkB Binding	IkBa(nuc) + NFkB(nuc) = IkBa:NFkB(nuc)	$v^{29} = kf^{29} \cdot [IkBa(nuc)] \cdot [NFkB(nuc)] - kr^{29} \cdot [IkBa:NFkB(nuc)]$	kf^{29} 0.5 ml/(µmol*s) kr^{29} 0.0005 1/s
30	NFkB15) Nuclear IkBb:NFkB Binding	IkBb(nuc) + NFkB(nuc) = IkBb:NFkB(nuc)	$v^{30} = kf^{30} \cdot [IkBb(nuc)] \cdot [NFkB(nuc)] - kr^{30} \cdot [IkBb:NFkB(nuc)]$	kf^{30} 0.5 ml/(µmol*s) kr^{30} 0.0005 1/s
31	NFkB16) Nuclear IkBe:NFkB Binding	IkBe(nuc) + NFkB(nuc) = IkBe:NFkB(nuc)	$v^{31} = kf^{31} \cdot [IkBe(nuc)] \cdot [NFkB(nuc)] - kr^{31} \cdot [IkBe:NFkB(nuc)]$	kf^{31} 0.5 ml/(µmol*s) kr^{31} 0.0005 1/s
32	NFkB17) Constitutive IkBa mRNA Synthesis	source -> IkBa_mRNA	$v^{32} = kf^{32} \cdot [source]$	kf^{32} 1.54E-06 1/s
33	NFkB18) Inducible IkBa mRNA Synthesis	2 * NFkB(nuc) -> IkBa_mRNA + 2 * NFkB(nuc)	$v^{33} = kf^{33} \cdot 2 \cdot [NFkB(nuc)]$	kf^{33} 0.0165 ml/(µmol*s)
34	NFkB19) IkBa mRNA degradation	IkBa_mRNA -> sink	$v^{34} = kf^{34} \cdot [IkBa\_mRNA]$	kf^{34} 0.00028 1/s
35	NFkB20) Constitutive IkBb mRNA Synthesis	source -> IkBb_mRNA	$v^{35} = kf^{35} \cdot [source]$	kf^{35} 1.78E-07 1/s
36	NFkB21) IkBb mRNA degradation	IkBb_mRNA -> sink	$v^{36} = kf^{36} \cdot [IkBb\_mRNA]$	kf^{36} 0.00028 1/s
37	NFkB22) Constitutive IkBe mRNA Synthesis	source -> IkBe_mRNA	$v^{37} = kf^{37} \cdot [source]$	kf^{37} 1.27E-07 1/s
38	NFkB23) IkBe mRNA degradation	IkBe_mRNA -> sink	$v^{38} = kf^{38} \cdot [IkBe\_mRNA]$	kf^{38} 0.00028 1/s
39	NFkB24) IKK:IkBa Binding	IKK[P] + IkBa = IKK[P]:IkBa	$v^{39} = kf^{39} \cdot [IKK[P]] \cdot [IKBa] - kr^{39} \cdot [IKK[P]:IkBa]$	kf^{39} 0.0225 ml/(µmol*s) kr^{39} 0.00125 1/s
40	NFkB25) IkBa Translation	IkBa_mRNA -> IkBa + IkBa_mRNA	$v^{40} = kf^{40} \cdot [IkBa\_mRNA]$	kf^{40} 0.00408 1/s
41	NFkB26) IkBa Degradation	IkBa -> sink	$v^{41} = kf^{41} \cdot [IkBa]$	kf^{41} 0.000113 1/s

Red italics = parameters derived during model fitting to experimental data from Covert et al. (205); *all other values* = parameters used in Hoffmann et al.'s model of NF-κB signalling (204). *v*= reaction rate, *kf*= rate of the forward reaction, *kr* = rate of the reverse reaction.

Table 4.1 continued.

	Reaction name	Reaction equation	Rate equation	Kinetic parameters
42	NFkB27) IkBa Nuclear Import/Export	IkBa = IkBa(nuc)	$v^{42} = kf^{42} \cdot [IkBa] - kr^{42} \cdot [IkBa(nuc)]$	kf^{42} 0.0003 1/s kr^{42} 0.0002 1/s
43	NFkB28) IKK:IkBb Binding	IKK[P] + IkBb = IKK[P]:IkBb	$v^{43} = kf^{43} \cdot [IKK[P]] \cdot [IKBb] - kr^{43} \cdot [IKK[P]:IkBb]$	kf^{43} 0.006 ml/(µmol*s) kr^{43} 0.00175 1/s
44	NFkB29) IkBb Translation	IkBb_mRNA -> IkBb + IkBb_mRNA	$v^{44} = kf^{44} \cdot [IkBb\_mRNA]$	kf^{44} 0.00408 1/s
45	NFkB30) IkBb Degradation	IkBb -> sink	$v^{45} = kf^{45} \cdot [IkBb]$	kf^{45} 0.000113 1/s
46	NFkB31) IkBb Nuclear Import/Export	IkBb = IkBb(nuc)	$v^{46} = kf^{46} \cdot [IkBb] - kr^{46} \cdot [IkBb(nuc)]$	kf^{46} 0.00015 1/s kr^{46} 0.0001 1/s
47	NFkB32) IKK:IkBe Binding	IKK[P] + IkBe = IKK[P]:IkBe	$v^{47} = kf^{47} \cdot [IKK[P]] \cdot [IKBe] - kr^{47} \cdot [IKK[P]:IkBe]$	kf^{47} 0.009 ml/(µmol*s) kr^{47} 0.00175 1/s
48	NFkB33) IkBe Translation	IkBe_mRNA -> IkBe + IkBe_mRNA	$v^{48} = kf^{48} \cdot [IkBe\_mRNA]$	kf^{48} 0.00408 1/s
49	NFkB34) IkBe Degradation	IkBe -> sink	$v^{49} = kf^{49} \cdot [IkBe]$	kf^{49} 0.000113 1/s
50	NFkB35) IkBe Nuclear Import/Export	IkBe = IkBe(nuc)	$v^{50} = kf^{50} \cdot [IkBe] - kr^{50} \cdot [IkBe(nuc)]$	kf^{50} 0.00015 1/s kr^{50} 0.0001 1/s
51	NFkB36) IKK:IkBa:NFkB Binding (2)	IKK[P]:IkBa + NFkB = IKK[P]:IkBa:NFkB	$v^{51} = kf^{51} \cdot [IKK[P]:IkBa] \cdot [NFkB] - kr^{51} \cdot [IKK[P]:IkBa:NFkB]$	kf^{51} 0.5 ml/(µmol*s) kr^{51} 0.0005 1/s
52	NFkB37) IkBa:NFkB Nuclear Export	IkBa:NFkB(nuc) -> IkBa:NFkB	$v^{52} = kf^{52} \cdot [IkBa:NFkB(nuc)]$	kf^{52} 0.0138 1/s
53	NFkB38) IKK:IkBb:NFkB Binding (2)	IKK[P]:IkBb + NFkB = IKK[P]:IkBb:NFkB	$v^{53} = kf^{53} \cdot [IKK[P]:IkBb] \cdot [NFkB] - kr^{53} \cdot [IKK[P]:IkBb:NFkB]$	kf^{53} 0.5 ml/(µmol*s) kr^{53} 0.0005 1/s
54	NFkB39) IkBb:NFkB Nuclear Export	IkBb:NFkB(nuc) -> IkBb:NFkB	$v^{54} = kf^{54} \cdot [IkBb:NFkB(nuc)]$	kf^{54} 0.0052 1/s
55	NFkB40) IKK:IkBe:NFkB Binding (2)	IKK[P]:IkBe + NFkB = IKK[P]:IkBe:NFkB	$v^{55} = kf^{55} \cdot [IKK[P]:IkBe] \cdot [NFkB] - kr^{55} \cdot [IKK[P]:IkBe:NFkB]$	kf^{55} 0.5 ml/(µmol*s) kr^{55} 0.0005 1/s
56	NFkB41) IkBe:NFkB Nuclear Export	IkBe:NFkB(nuc) -> IkBe:NFkB	$v^{56} = kf^{56} \cdot [IkBe:NFkB(nuc)]$	kf^{56} 0.0052 1/s
57	NFkB42) IKK:IkBa Catalysis	IKK[P]:IkBa -> IKK[P]	$v^{57} = kf^{57} \cdot [IKK[P]:IkBa]$	kf^{57} 0.00407 1/s
58	NFkB43) IKK:IkBb Catalysis	IKK[P]:IkBb -> IKK[P]	$v^{58} = kf^{58} \cdot [IKK[P]:IkBb]$	kf^{58} 0.0015 1/s
59	NFkB44) IKK:IkBe Catalysis	IKK[P]:IkBe -> IKK[P]	$v^{59} = kf^{59} \cdot [IKK[P]:IkBe]$	kf^{59} 0.0022 1/s

Red italics = parameters derived during model fitting to experimental data from Covert et al. (205); *all other values* = parameters used in Hoffmann et al.'s model of NF-κB signalling (204). *v*= reaction rate, *kf*= rate of the forward reaction, *kr* = rate of the reverse reaction.

Table 4.2. . Initial (pre-steady state) concentrations of species used in the model.

Species	Initial concentration (µmol/ml)
CD14	1*
IkBa	0
IkBa(nuc)	0
IkBa:NFkB	0
IkBa:NFkB(nuc)	0
IkBa_mRNA	0
IkBb	0
IkBb(nuc)	0
IkBb:NFkB	0
IkBb:NFkB(nuc)	0
IkBb_mRNA	0
IkBe	0
IkBe(nuc)	0
IkBe:NFkB	0
IkBe:NFkB(nuc)	0
IkBe_mRNA	0
IKK	0.1
IKK[P]	0
IKK[P]:IkBa	0
IKK[P]:IkBa:NFkB	0
IKK[P]:IkBb	0
IKK[P]:IkBb:NFkB	0
IKK[P]:IkBe	0
IKK[P]:IkBe:NFkB	0
IRAK1	1
IRAK4	1*
IRF3	1
IRF3[P]	0
IRF3[P](nuc)	0
LBP	1*
LPS	1*
LPS:LBP:CD14:TLR4:RIP1:TRAM:TRIF:TBK/IKKe	0
LPS:LBP:CD14:TLR4:TIRAP:MyD88:IRAK4	0
MyD88	1*
NFkB	0.1
NFkB(nuc)	0
RIP1	1*
sink	0*
source	1*
TAK1:TAB1:TAB2	1*
TAK1:TAB1:TAB2:TRAF6	0

TBK1/IKKe	1*
TIRAP	1*
TLR4	1*
TNFa	0
TNFa:TNFR1:TRAF2:TRADD:RIP1	0
TNFR1	1*
TRADD	1*
TRAF2	1*
TRAF6	1*
TRAF6:IRAK1[P]	0
TRAM	1*
TRIF	1*

All values were set to 1µmol/ml if they were assumed to be present at time=0s and 0µM if they were assumed to be absent. NF-κB was set to 0.1µmol/ml as in Hoffmann *et al*.(204). Concentrations marked with an asterisk were 'fixed' at their initial concentrations to avoid overcomplicating the model by modelling synthesis and degradation of these species.

4.3.3.2 Steady state behaviour

The model was run to steady state (i.e. when there were no further changes in concentrations over time) using different concentrations of LPS. This showed that both IKK (active and inactive) and free NF-κB (nuclear and cytoplasmic) show a dose response (Figure 4.2). At LPS doses over 0.4, there is a switch from the majority of IKK being inactive at steady state, to the majority being phosphorylated at steady state. Concentrations of nuclear and cytoplasmic NF-κB both increase with higher doses of LPS because LPS treatment leads to an increase in free cytoplasmic NF-κB via phosphorylated IKK, and this free cytoplasmic NF-κB then translocates to the nucleus.

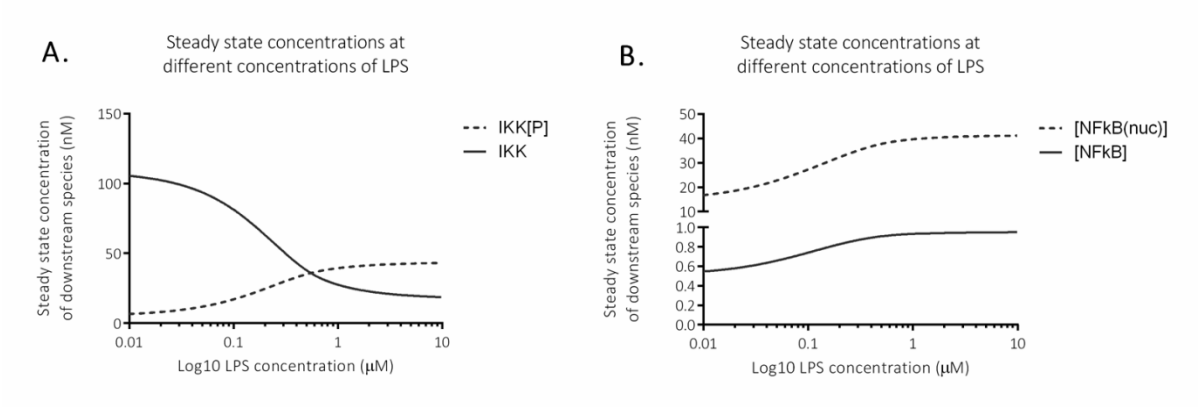


Figure 4.2. Steady state concentrations at different doses of LPS. A. *In silico* simulation of steady state concentrations of inactive and phosphorylated IKK, and B. *In silico* simulation of steady state concentrations of free nuclear and cytoplasmic NF-κB.

4.3.3.3 Time course behaviour

The model mimics *in silico* the activation of NF-κB over time that Covert *et al.*(205) described *in vitro* in time course experiments on LPS-treated wild-type MEFs, and *in silico* in their model of IKK-NF-κB signalling (Figure 4.3). Nuclear (active) concentration of NF-κB shows damped oscillatory behaviour. The model also mimics *in silico* the time-course concentration of phosphorylated IKK found by Covert *et al*. (205) *in vitro* (Figure 4.4). The *in silico* model presented here mimics the pattern of these behaviours, although not to the exact degree; the timing, number of oscillations and exact concentrations are not the same as found by Covert *et al*. (205) *in vitro*. The model predicts that the concentration of the phosphorylated (active) form of IKK should increase with a small amount of oscillation and reach a maximum around 4 hours after LPS treatment. Covert *et al.*'s (205) Western blot analysis suggests a similar pattern, but with a faster and steadier increase in phosphorylated IKK, reaching a maximum at around two hours after treatment.

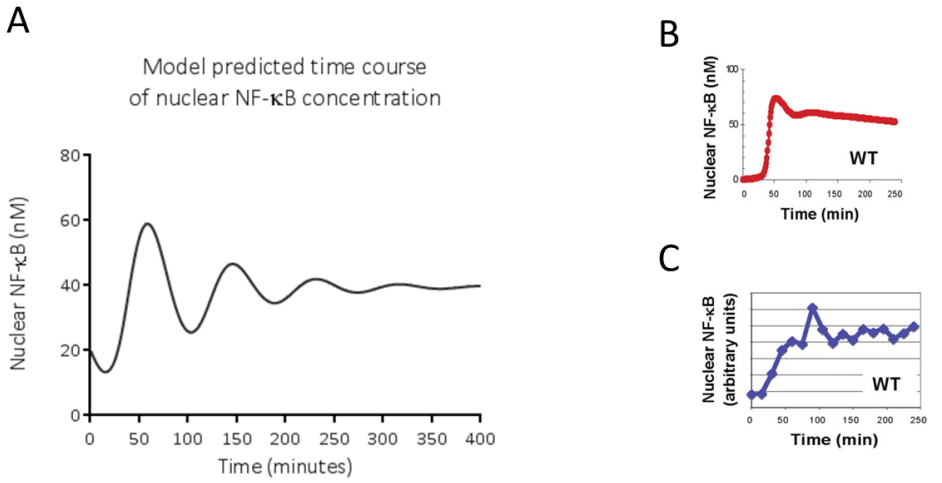


Figure 4.3. NF-κB time course behaviour of A. Concentrations of nuclear NF-κB over time, simulated using the *in silico* model described here, B. Concentrations of nuclear NF-κB over time, simulated using the *in silico* model described by Covert *et al.*, C. Experimental data from LPS-treated mouse embryo fibroblasts as described by Covert *et al.* (205)

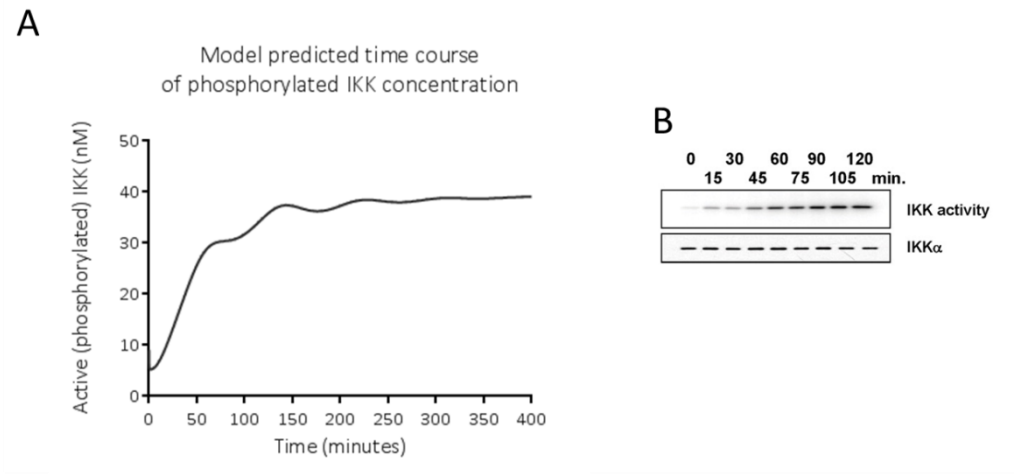


Figure 4.4. IKK time course behaviour. A. Concentrations of nuclear IKK over time, simulated using the *in silico* model described here, B. Experimental data from LPS-treated mouse embryo fibroblasts as described by Covert *et al.* (205)

The model mimics *in silico* the activation of NF-κB over time that Covert *et al.* (205) found in TRIF (Tir-Domain-Containing Adapter-Inducing Interferon-β) and MyD88 (Myeloid Differentiation Primary Response Gene 88) knock-out cells *in vitro* (Figure 4.5). After LPS treatment of both knock-out cells, the time-course experiments of Covert *et al.* (205) showed increased oscillatory NF-κB activation compared to wild-type cells, and the initiation of NF-κB activation was delayed by around 30 minutes in MyD88 knock-out

compared to wild-type cells. Covert *et al*. (205) argued that this 30 minute delay occurs because the TRIF-dependent pathway relies on the synthesis and actions of TNFα. Whereas Covert *et al*. (205) could only mimic this behaviour *in silico* by introducing an artificial delay to mimic MyD88 knock-out, the extended model presented here allows MyD88 or TRIF to be 'knocked-out' directly, by fixing their concentrations at 0 before running steady state and time course simulations as described above. The model described here also includes the reactions involved in the synthesis of TNFα and its autocrine actions on the cell. Therefore it can be used to test the downstream effects of MyD88 and TRIF knock-out more naturally. The model successfully predicts increased oscillations in NF-κB activation when TRIF or MyD88 is 'knocked-out', and there is a simulated delay in NF-κB activation when MyD88 is 'knocked-out'. Again, although the model captures the general pattern of the *in vitro* data, it is not quantitatively accurate.

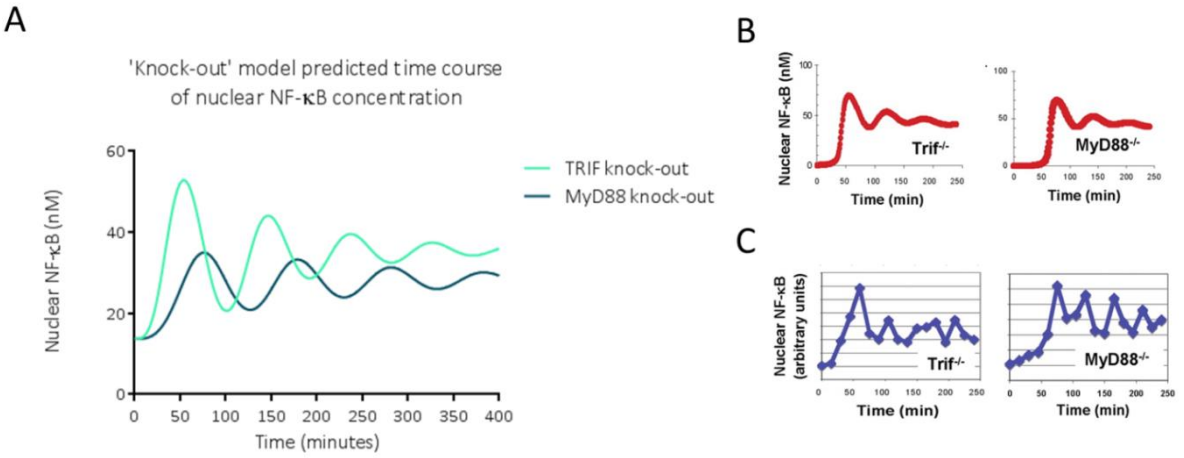


Figure 4.5. NF-κB time course behaviour in TRIF and MyD88 knock-out conditions. A. Concentrations of nuclear NF-κB over time, simulated using TRIF and MyD88 knock-out versions of the *in silico* model described here, B. Concentrations of nuclear NF-κB over time, simulated using TRIF and MyD88 knock-out versions of the *in silico* model described by Covert *et al*., C. Experimental *in vitro* data from LPS-treated TRIF or MyD88 knock-out mouse embryo fibroblasts as described by Covert *et al*. (205)

4.4 Discussion

Using only published data, a graphical depiction and kinetic model of LPS-induced NF-κB activation has been produced. The model successfully recapitulates *in silico* the published behaviour of the system *in vitro* in MEFs.

Computational modelling is a major growth area in biomedical research, but has only rarely been applied to pregnancy physiology or pathology. The model described here is the first kinetic model of a signalling pathway involved in parturition. The LPS-NF-kB pathway is an important candidate signalling pathway in infection-induced preterm labour. There is only one report of another computational model to study the molecular events initiating labour (121). This model by Equils *et al*. uses published data to model the immune-endocrine interactions in a uterine smooth muscle cell with an increase in the ratio of progesterone receptor A (PR-A) to progesterone receptor B (PR-B) as an endpoint. It showed that NF-κB increased the PR-A:PR-B ratio, and that higher doses of NF-κB shortened the time to reach the PR-A:PR-B ratio observed in labour. The model assumes that NF-κB is a marker of infection so these results reflect the known association between infection and preterm birth. This is an encouraging and useful first step towards modelling preterm labour, however the model does not include the molecular interactions upstream of NF-κB activation that initiate the whole pathway, and so does not allow *in silico* exploration of the importance of these interactions to the system. Additionally, the model does not include the complex interactions between molecules at an intracellular level and therefore risks oversimplifying the system. The model described here extends this work by comprehensively including interactions upstream of NF- κB activation.

Models of NF-κB activity have previously been published in scenarios outwith pregnancy. These generally use IKK as an input to allow the model to be adapted to simulate NF-κB activity following any treatment (204–206). However, in a pregnancy scenario, LPS is a more appropriate input than IKK because LPS could be considered the initiator of some cases of infection-induced preterm labour. Therefore, the model described here extends previous models by using LPS as an input and explicitly modelling molecular interactions upstream of IKK activation, including LPS to IKK and the production and action of TNFα. This allows closer analysis of the interactions that activate IKK and therefore affect downstream NF-kB activity. After validation using cells from human uterine smooth muscle cells, this will allow *in silico* testing of drugs targeting these upstream interactions.

This is the first attempt to explicitly model these upstream events appropriately. One previous attempt by Selvarajoo (300) was flawed because although the kinetic rate equations were based on mass action kinetics, they did not describe physical interactions between individual entities. For example, the first reaction in the Selvarajoo model (300),

"TLR4 <-> MyD88" (rate equation: Kf [TLR4] – Kr[MyD88]), describes a reaction where TLR4 is reversibly converted to MyD88, which does not represent the true physical interaction between these two molecules.

Explicitly modelling the upstream events produces a more complete model that is able to reproduce *in silico* the published behaviour of the system *in vitro* in wild-type, MyD88 knock-out and TRIF knock-out MEFs. However, it should be noted that, to an extent, the model was built for this purpose. At the time of model development, we were aware of the behaviour of NF-κB in LPS-treated MyD88 and TRIF knock-outs that Covert et al. (205) reported, but we did not develop the model under knock-out conditions. Instead, we built the "wild-type" model and used MyD88- and TRIF- "knock-out" conditions to test the model in silico. Several "tests" and subsequent changes to the wild-type model were made during model development. A true test of the model would be to "knock-out" or inhibit another component of the pathway and compare the in silico behaviour to that achieved in vitro. We are unaware of any published data that would allow us to make this comparison.

Although the model described here can simulate the pattern of the *in vitro* behaviour, the exact timing, number of oscillations and exact concentrations were different. However, this is unlikely to invalidate the model because the kinetics of the pathway are also likely to alter in different experimental conditions and in different cell types. There are undoubtedly more LPS targets that could be incorporated into extended versions of the model to make it more comprehensive and improve its potential to make predictions about the relative importance of different parts of the pathway. In particular, it would be interesting to expand the model to include more inflammatory cytokines aside from TNFα. Inflammatory cytokines (such as IL-8 and IL-6) have been shown to induce preterm labour and uterine contractions in animal models (301,302), and their expression is also regulated by NF-κB (243–245), therefore a model of inflammatory cytokine-induced NF-κB activation would be useful in the context of parturition. In the early model described here, a decision was made to use LPS as a model input, because it is exogenous to the cell and could therefore be modelled with a fixed concentration. Future, more complex, versions of the model could use other inputs that are also produced by the cell. A model of NF-κB activation via TLR2 would also be useful, because TLR2 is activated by naturally occurring endotoxins in vivo.

No published data was found to describe the structure or kinetics of the LPS-NF-κB pathway in uterine smooth muscle cells and therefore it cannot be confirmed that the pathway does not deviate from that described in MEFs. Although this is a major limitation of the model, it is unlikely that there would be any large deviations because the pathway appears to be well conserved (303) and the main components of the pathway (NF-κB, TLR4, TNFα and IκBα) have been shown to be expressed in uterine smooth muscle (232,304). However, wet lab experiments using uterine smooth muscle cells should be conducted to validate the model in uterine smooth muscle cells. For example, quantitative western blots and/or ELISAs could be used to quantitate the levels of active and inactive proteins in the pathway at short time intervals up to around two hours. The most important parameters to define would be those describing the behaviour of TLR4, IKK, NF-κB, IκBα and TNFα in response to LPS stimulation. NF-κB activation may be best measured using electromobility shift assay (EMSA). PCR to measure mRNA transcripts of IκBα and TNFα would also be useful. The lack of available published data also highlights the need for the publication of detailed data from time course experiments to aid with model building. More data in the literature would have been useful for validating the model. For example, detailed results of in vitro experiments on IKK knock-out cells would have allowed us to test if our model could predict changes to NF-κB activation under these circumstances. The SBML version of this model has been uploaded to PloSONE (as supplementary material to the paper) and the Biomodels database (305) to encourage other researchers to use it. This is an important next step in validating the model and confirming its usefulness.

This work provides proof of concept that it is possible to build computational models of signalling pathways relevant to labour. When validated using wet lab experiments on cells derived from human gestational tissue (for example, uterine smooth muscle cells), such models could be used for drug testing *in silico*, providing a rapid, safe, economical and ethical strategy to identify candidate effective therapies for further testing. Thus, these models have the potential to improve our understanding of parturition and translate into improved pregnancy outcomes.

5 A NETWORK GRAPH APPROACH TO EXPLORE EPIDEMIOLOGICAL DATA

5.1 INTRODUCTION

Chapters 3 and 4 describe applications of data-driven, systems biology approaches to study the molecular biology of parturition. This chapter explores the use of a systems approach to study the epidemiology of parturition-related events.

Conventional data analysis in epidemiology involves testing specific hypotheses generated by the researcher. Although undoubtedly useful, this approach risks overlooking interesting and confounding associations the researcher does not expect to see. A similar limitation has been described of the reductionist approach in molecular biology. In that field, systems biology has been proposed as a complementary approach that can help overcome this limitation through hypothesis-free data-driven analyses. For the work described in this thesis, it was proposed that a similar, systems approach could be applied in epidemiology, to complement traditional methods. Specifically, this chapter describes the use of 3D, interactive, unbiased network graphs to visually explore the Walker Cohort – a database of nearly 50,000 birth records collected from 1952 to 1966 in Dundee. To our knowledge, this is the first time network graphs have been used to explore epidemiological data. BioLayout *Express*^{3D} (which is most commonly used to analyse gene expression data (150,151,153,306–308), and was used for this purpose in Chapter 3) was used to build three dimensional network graphs in which each 'node' represents either a clinical characteristic or a delivery episode, and these are connected by weighted 'edges' based on Pearson's correlation coefficients.

5.2 SUMMARY OF METHODS

The detailed methods for this chapter are described in Chapter 2. Briefly, the Walker Cohort of 48985 births (described in section 2.3.1) was prepared by removing erroneous values and converting continuous variables to discrete, binary variables (described in section 2.3.2). The data was then used to build an "obstetric characteristic graph", in which nodes represent obstetric characteristics such as small placental weight (described in section 2.3.3), and "delivery episode graphs", in which nodes represent delivery episodes

(described in section 2.3.4). An outline of the final protocol is provided in Figure 5.1. A description of all the variables available in the Walker dataset is provided in Appendix 5.

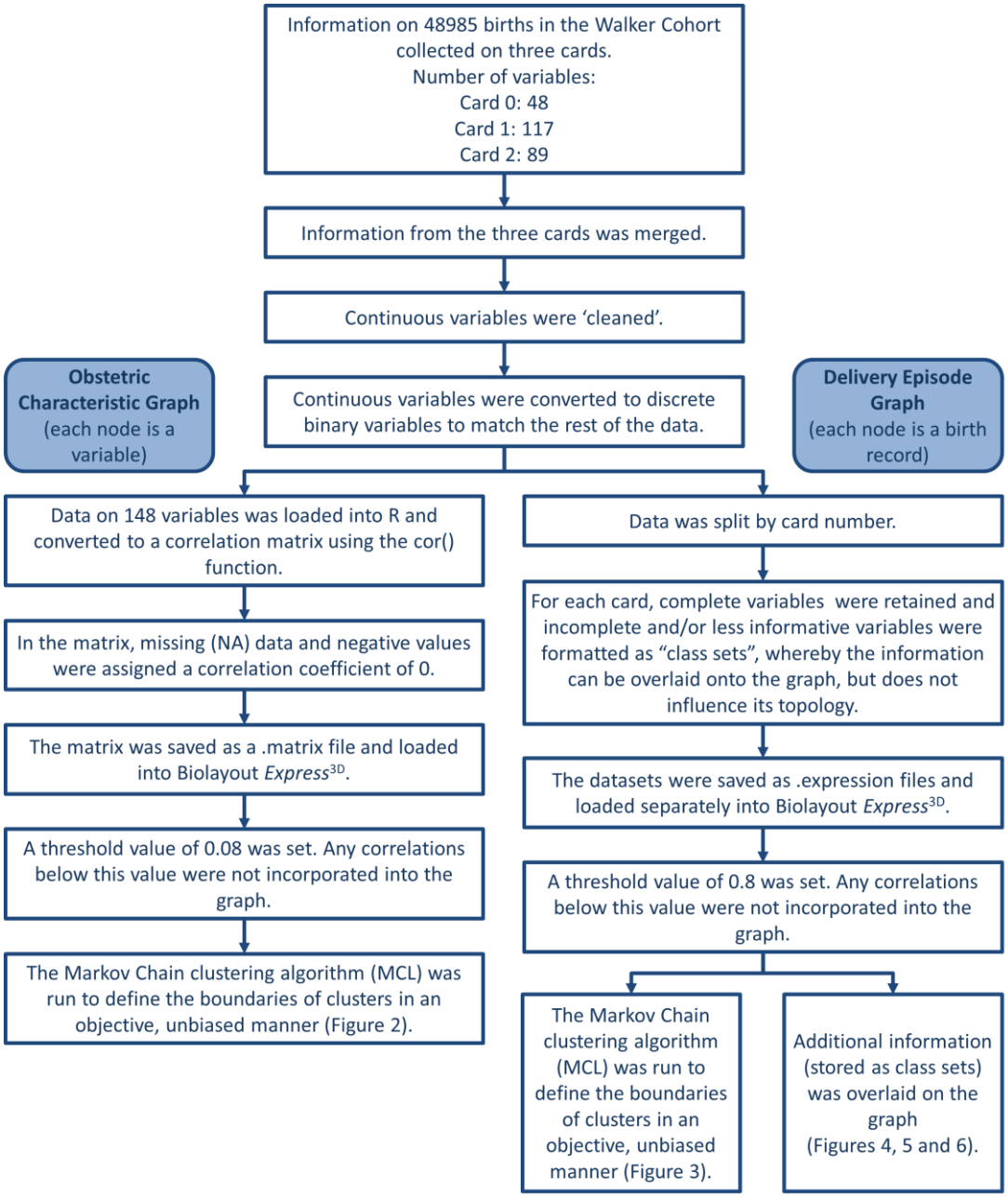


Figure 5.1. An outline of the final protocol developed to analyse the Walker data.

5.3 Results

5.3.1 Effect of data cleaning

Data stored as continuous variables were 'cleaned' to remove potentially inaccurate data. This affected the distribution of the data, in most cases making it more consistent with what would be expected (Figure 5.2 and Figure 5.3). For example, birth weight was more normally distributed after cleaning and maternal age was more clearly skewed to the left (younger mothers).

The two sub-populations that appear to exist in variables such as maternal/paternal height and head circumference are likely to have arisen because some values were recorded in centimetres and some were converted from imperial measurements. If the recorded value was metric, it is more likely to be an integer to the nearest centimetre, whereas converted values are more likely to be fractions. Therefore, there are unlikely to be two clear sub-populations in reality.

The bimodal distribution of crown-heel and crown-rump length suggests that these measurements were more likely to have been recorded if they deviated from a "normal" range of values. This bias is important to consider when interpreting data on any variable.

Cleaning removed many erroneous 0 values that were possibly used when data was missing. It also removed extreme low or high values that might have arisen through recording or typing errors. Table 5.1 shows key statistics describing the distribution of continuous data before and after cleaning. The procedure successfully removes the relatively small number of values that might introduce bias, without substantially altering the median and 25th and 75th centiles (except in variables such as paternal age and height, which contained many erroneous 0 values).

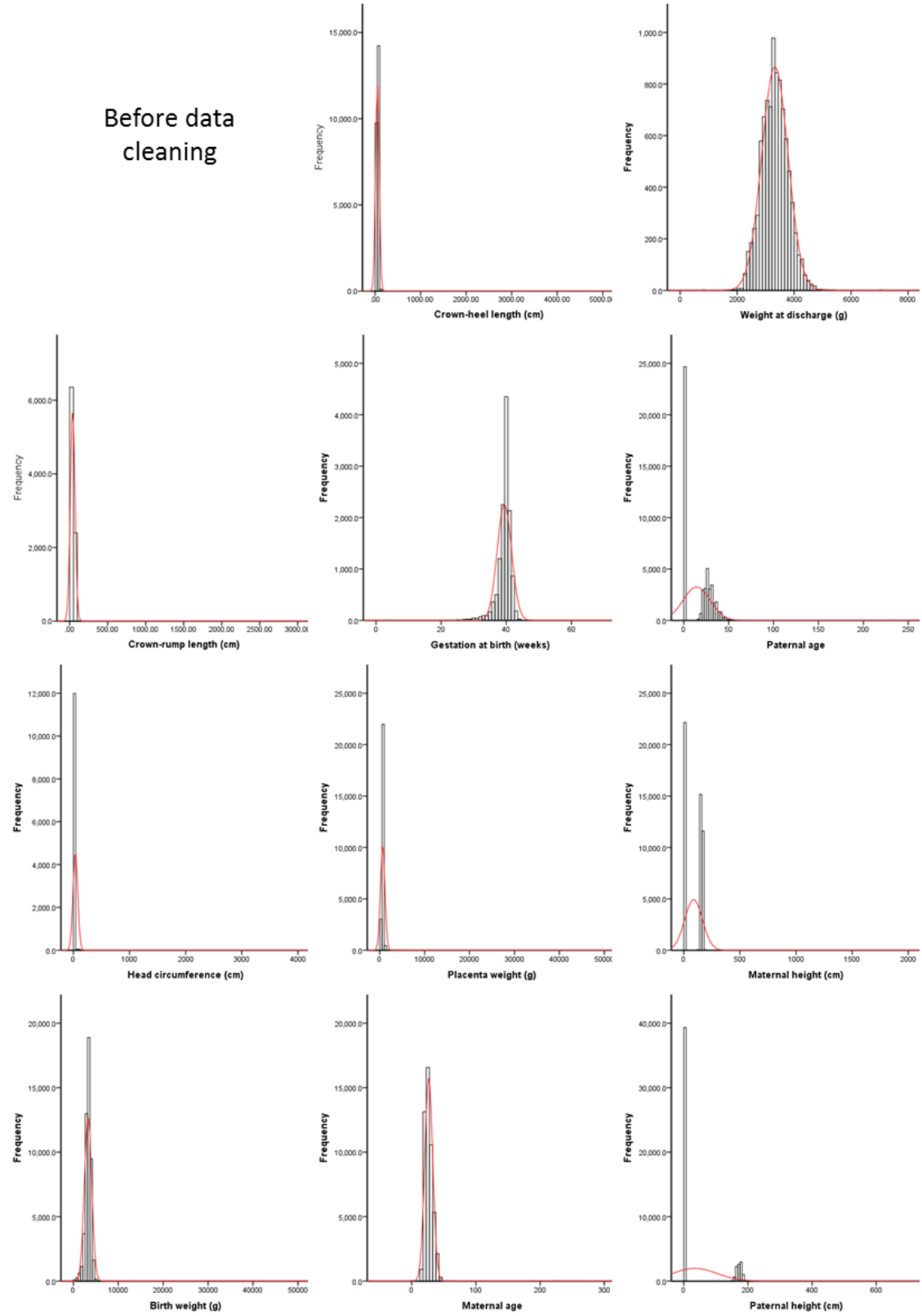


Figure 5.2. Histograms to show the distribution of data for continuous variables before data cleaning.

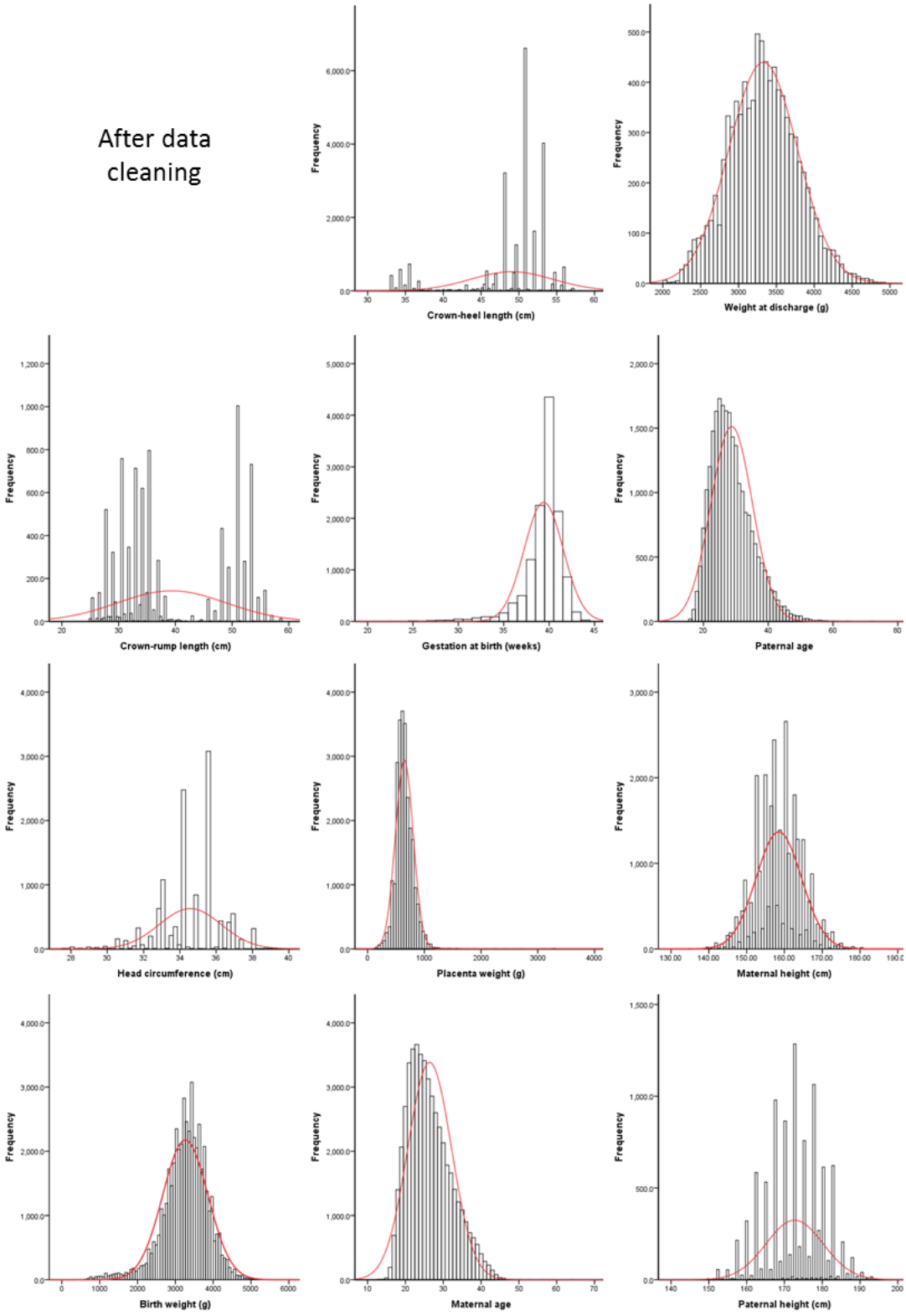


Figure 5.3. Histograms to show the distribution of data for continuous variables after data cleaning.

	Before data cleaning							After data cleaning						
	n	Mean	Median	Min	Max	25th %ile	75th %ile	n	Mean	Median	Min	Max	25th %ile	75th %ile
Crown heel length (cm)	24096	49.7	50.8	2.5	4914.9	48.3	52.1	23658	49.1	50.8	33.0	58.1	48.3	52.1
Crown rump length (cm)	8755	39.8	35.6	0.00	2709.0	31.1	50.8	8584	39.4	35.6	24.9	58.4	31.7	50.8
Gestation at birth (weeks)	12466	39.4	40	3	66	39	41	12456	39.4	40	23	44	39	41
Head circumference (cm)	12083	35.3	34.3	3.6	3638.0	33.6	35.6	11850	34.57	34.29	27.50	39.30	33.60	35.56
Birth weight (g)	48985	3267.3	3315.0	0.00	42060.0	2948.4	3630.0	48856	3261.8	3315.0	411.1	5273.1	2948.4	3630.0
Weight at discharge (g)	9028	3326.4	3316.9	793.8	7030.7	3005.1	3628.7	9009	3325.8	3316.9	2012.8	4904.5	3005.1	3628.7
Paternal age	48985	14.3	0.00	0.00	232	0	28	24299	28.7	28	16	72	24	32
Paternal height (cm)	48985	34.1	0.00	0.00	670.6	0	0.00	9646	172.7	172.7	149.9	194.3	167.6	177.8
Maternal age	48985	26.4	25	0	253	22	30	48943	26.4	25	13	64	22	30
Maternal height (cm)	48985	86.88	150.50	0.00	1651.6	0	158.8	26392	158.7	158.8	145.4	179.7	154.9	162.6
Placenta weight (g)	25507	656.8	630.00	9.00	46720.0	560.0	730.0	25462	647.5	630.0	141.8	3316.9	560.0	730.0

Table 5.1. Key statistics to illustrate the distribution of continuous data before and after data cleaning.

5.3.2 Network graphs

BioLayout *Express*^{3D} worked well with binary data and produced informative network graphs that could be used to dynamically explore relationships between patient health characteristics, obstetric complications and fetal and maternal outcomes following delivery.

5.3.2.1 Obstetric characteristic graph

A network graph in which each node represents a variable was produced. The edges denote the strengths of the correlations between variables according to size (thicker edges show stronger correlations) and colour (from blue (weak correlation) to red (strong correlation)). Pearson's correlation coefficients are shown for each edge. The graph consisted of 54 nodes and 51 edges making up 14 small, unconnected graphs. Markov Chain (MCL) clustering revealed the four clusters shown in Figure 5.4. These clusters confirmed known associations between certain obstetric characteristics, and could therefore be categorised according to the main condition/characteristic likely to be linking them.

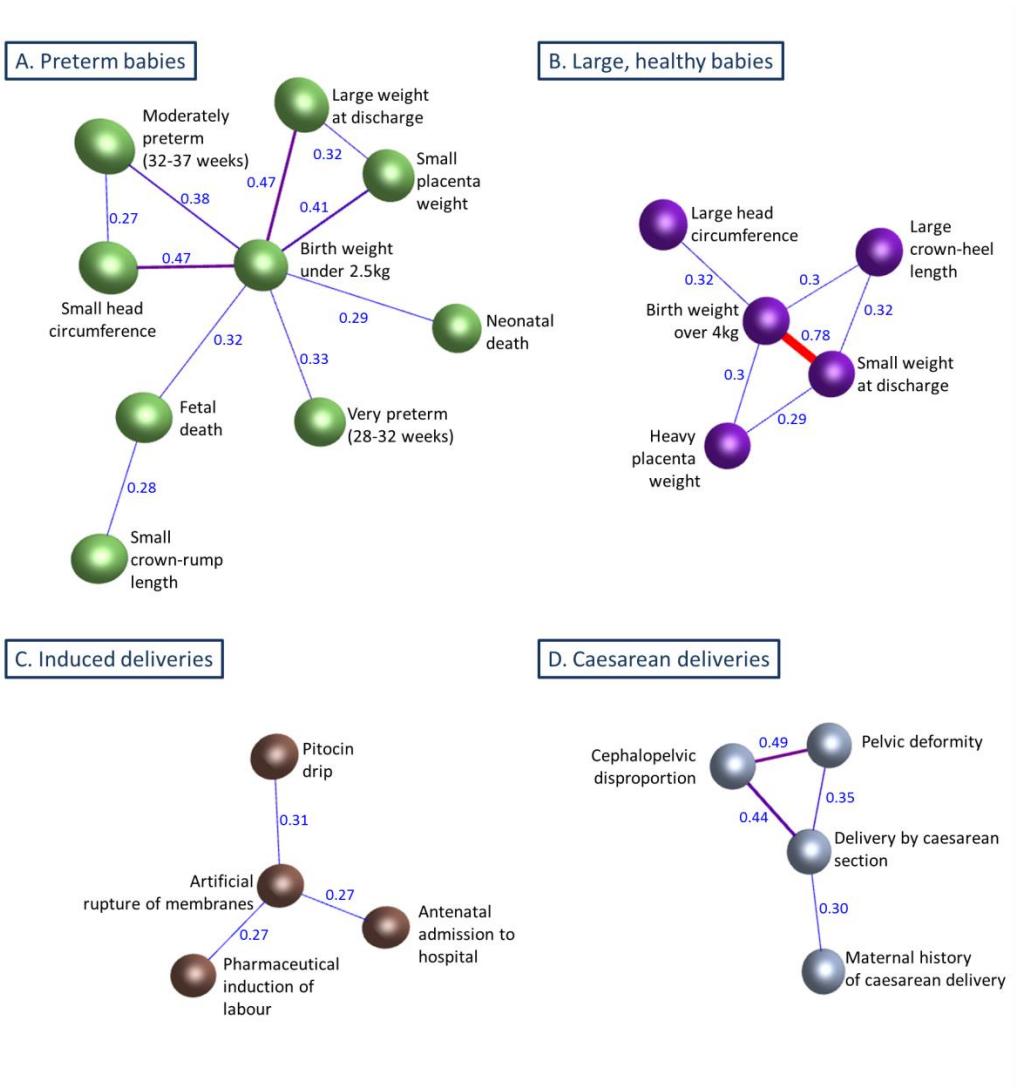


Figure 5.4. The four MCL clusters identified in the obstetric characteristic graph. Each cluster confirms known associations between certain obstetric characteristics. Large weight at discharge: ≥3912.2cm; Small weight at discharge: ≤2749.9cm; Small placenta weight: ≤481.94g; Heavy placenta weight: ≥830g; Small crown-rump length: ≤28.6cm; Small head circumference: ≤32.9cm; Large head circumference: ≥36.8cm; Large crown-heel length: ≥53.34cm.

5.3.2.2 Delivery episode graph

A larger network graph in which each node represents a different delivery episode was also produced. Over the study period, Walker data was collected on three different cards, which collected data on slightly different variables. Therefore, it was necessary to split the dataset by card number and create a separate graph using data from each card. After preparation, card 0 contained 6478 complete records over 39 variables, card 1 contained 10046 complete records over 36 variables, and card 2 contained 7982 complete records

over 24 variables. Here, only the graph built using data from card 0 is shown, because this is representative of the graphs built using data from the other cards. Card 0 also contained the most data (6478 delivery episodes * 39 variables = 252642 values).

The graphs consisted of a large main graph including most patients, and several smaller graphs including patients with very similar, but rarer types of delivery. Figure 5.5 shows the main graph built using data from card 0. Similar deliveries form MCL clusters that can be explored in more detail by viewing a chart of average values for every obstetric characteristic on the card (Figure 5.6). Although each individual delivery has a value of 1 (present) or 0 (absent) for each characteristic, the average is calculated across all deliveries, giving a value between 0 and 1. The clusters confirm known associations between risk factors and outcomes. For example, cluster 1 confirms known associations between preeclampsia and surgical induction of labour and cluster 8 confirms known associations between pelvic deformity, cephalopelvic disproportion and delivery by Caesarean section,

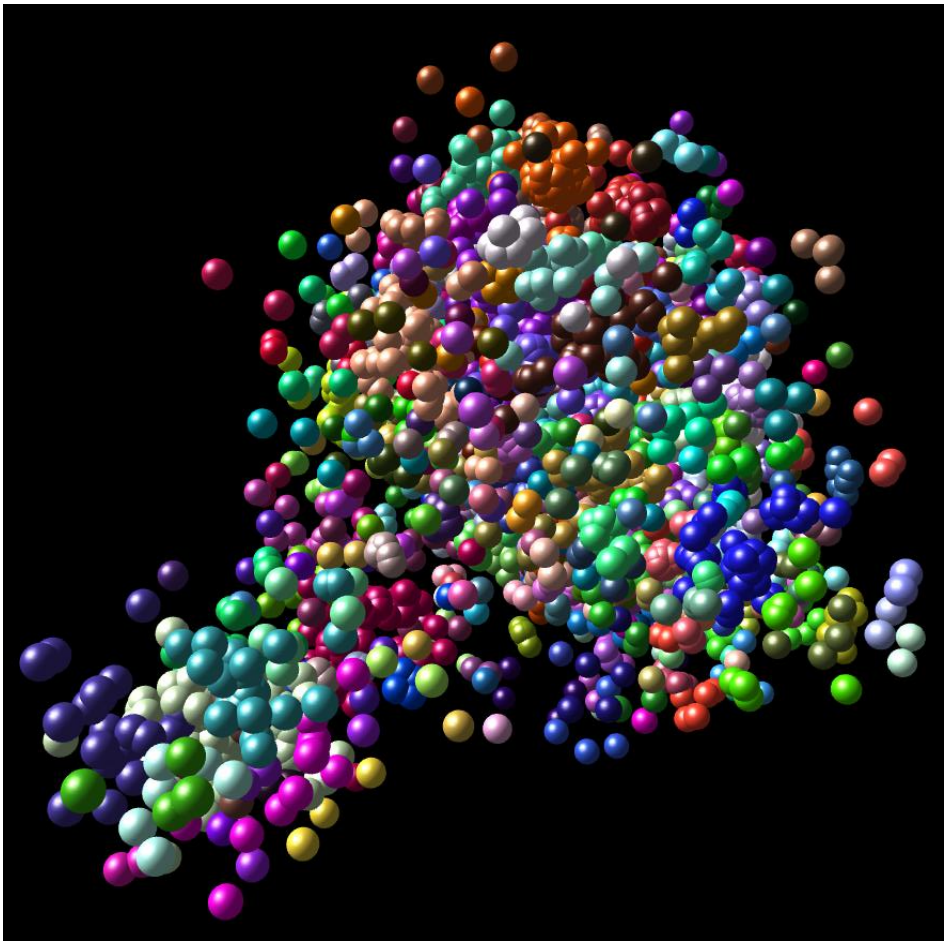


Figure 5.5. The main delivery episode network graph built using data from card 0.Edges have been hidden to aid visualisation.

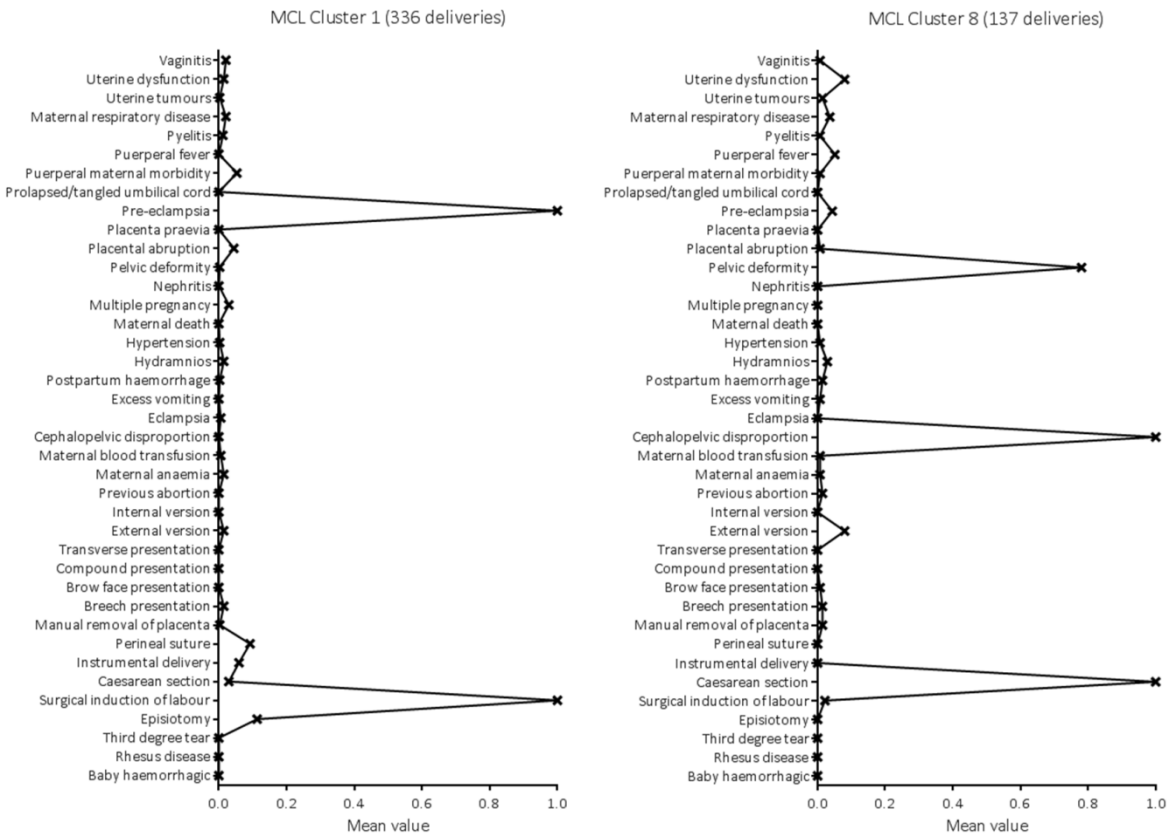


Figure 5.6. Examples of charts of average values for each obstetric characteristic in two MCL clusters. The charts show all obstetric characteristics used to create the network graph of data on Walker card 0.

The network graphs were overlain with additional information about each birth to explore how certain obstetric characteristics influence the data. This was particularly useful for highlighting key associations between risk factors and outcomes. For example, Figure 5.7 shows that nodes clearly cluster according to mode of delivery, and a higher incidence of pelvic deformity, uterine dysfunction and cephalopelvic disproportion are associated with assisted (surgical or instrumental) delivery compared to vaginal delivery. Figure 5.8 shows that nodes also cluster according to the duration of labour, with longer durations being associated with uterine dysfunction and assisted delivery. Figure 5.9 shows that fetal position is associated with different types of delivery, with right/left occiput posterior and lateral positions being associated with a higher risk of Caesarean section compared to the normal position (right/left occiput anterior).

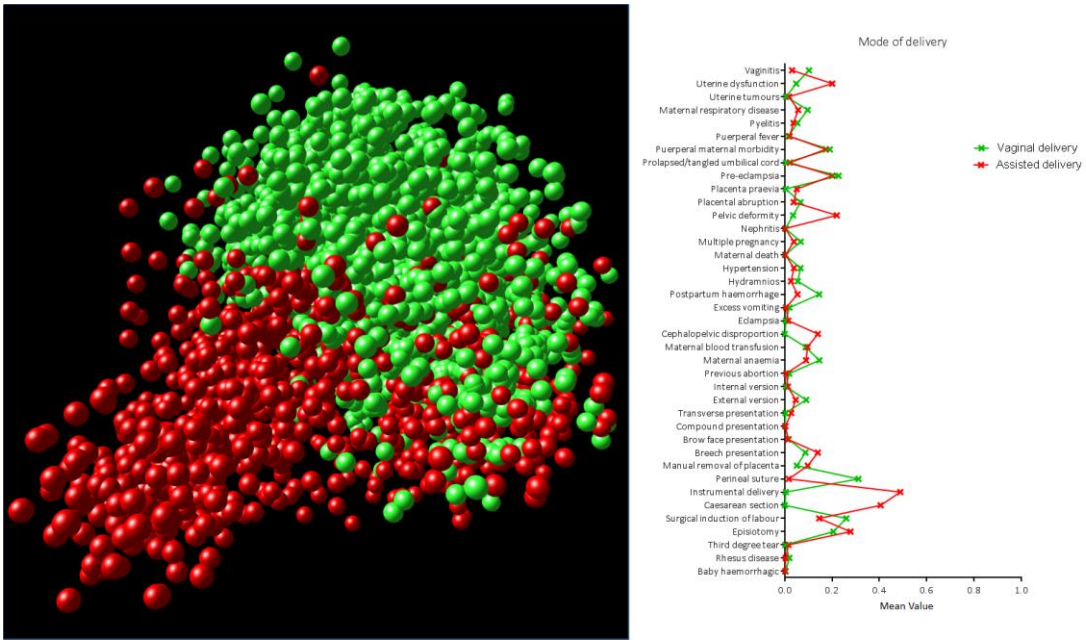


Figure 5.7. The main delivery episode graph for card 0 coloured according to mode of delivery.

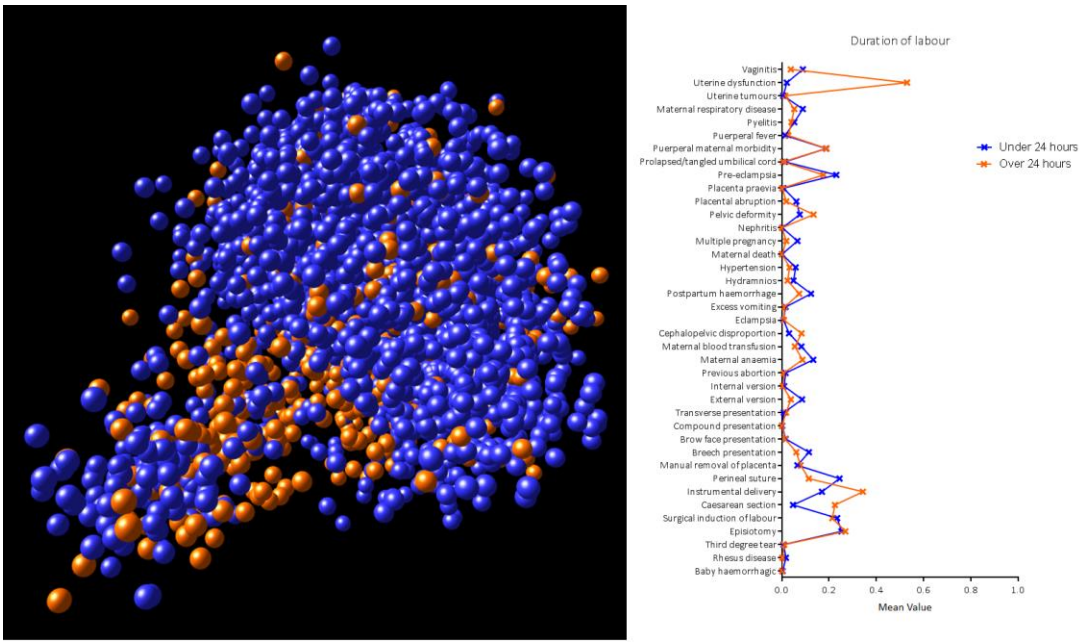


Figure 5.8. The main delivery episode graph for card 0 coloured according to duration of labour.

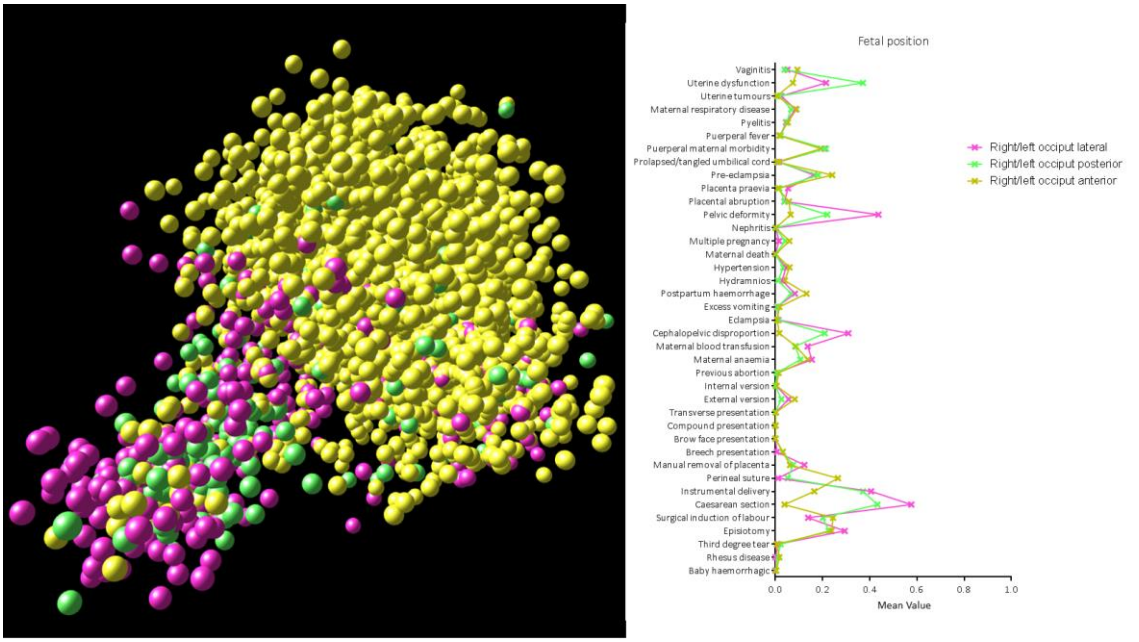


Figure 5.9. The main delivery episode graph for card 0 coloured according to fetal position.

5.4 Discussion

Biolayout Express^{3D} was used to produce network graphs to explore epidemiological birth record data. The network graphs confirmed known associations between risk factors and pregnancy outcomes.

Networks have previously been used to explore the transmission of communicable disease throughout populations (154,309–311), however, this is the first time a 3D network graph has been used to explore epidemiological data on non-communicable health outcomes.

The large size of the Walker project provided an ideal dataset in which to develop this technique. It contains over 48000 records, representing around 75% of births of all babies in Dundee during the 15-year period it was conducted (212). The relatively recent study period means the deliveries are more comparable to the present generation's deliveries than some other historical cohorts. The main limitation of the Walker data is the lack of information on maternal smoking, parental diet and BMI, which are likely to be important factors influencing pregnancy outcome.

When using nodes to represent each variable (as in the obstetric characteristic graph), relationships can be visualised and traced to a wider network to generate hypotheses

about covariates. The graph formed four MCL clusters that confirmed known associations between obstetric characteristics. The largest cluster (Figure 5.4a) contained nodes representing characteristics of small, compromised premature babies (fetal death, low birth weight, small crown-rump, small head circumference, light placenta weight), preterm birth (at 28-32 weeks and 32-37 weeks) and consequences of preterm birth (neonatal death) (66–68). Large weight at discharge also appears in this cluster, which initially seemed counterintuitive. However, we have hypothesised that the relatively high correlation between large discharge weight and low birthweight might be explained by neonatal growth occurring during a prolonged postnatal hospital stay. It is interesting to note that not all known associations cluster together, for example extreme prematurity (under 28 weeks) did not occur in this cluster. In fact, this characteristic did not appear in any cluster because it only correlated with one other characteristic (small crown-heel length) above the 0.08 Pearson's correlation value cut-off. This is likely to be due to the low prevalence of extreme prematurity in the Walker data (34 cases out of 12457 deliveries with information on this variable, 0.27%). The second largest cluster (Figure 5.4b) contained nodes characteristic of larger babies (large birthweight, large crown heel length and large head circumference). Small weight at discharge also appeared in this cluster, which we have hypothesised to be a result of a shorter postnatal hospital stay. Interestingly, these nodes do not connect to nodes representing obstetric complications classically related to macrosomia such as cephalopelvic disproportion or delivery by Caesarean section (312). Instead, these characteristics formed a different cluster (Figure 5.4d), along with two other common indications for Caesarean section – pelvic deformity (which was historically more common due to the higher prevalence of rickets(313)) and maternal history of Caesarean section (314). The final cluster (Figure 5.4c) contained nodes representing methods of inducing or augmenting labour (artificial rupture of membranes, pharmaceutical induction of labour, Pitocin (oxytocin) drip), along with antenatal admission to hospital, which is necessary to receive such treatment. This is as expected, because women often receive two or more interventions to induce/augment labour if initial attempts fail (315).

The "obstetric characteristic" graph approach copes well with missing data and allows discovery of 'emergent' relationships that could not be identified using a purely statistical approach. For example, the largest MCL cluster (Figure 5.4a) confirms associations between preterm birth and fetal and neonatal death. However, if we were to have used a

more traditional logistic modelling approach, it would have been impossible to infer the relationship between these variables because no Walker births have information on both gestation and fetal or neonatal death.

When using nodes to represent cases (as in the "delivery episode" graph), larger, more informative graphs can be explored to reveal factors that often occur together in different deliveries. This is useful for highlighting pregnancy risk factors for parturition complications. For example, in the delivery episode network graph for card 0, MCL cluster 1 comprises mostly deliveries in which women had pre-eclampsia and were subsequently surgically induced, this is in line with NICE guidelines on management of pre-eclampsia (316). Many of the women in this cluster also underwent episiotomy and experienced morbidity in the puerperal period. MCL cluster 8 comprises deliveries in which women tended to have pelvic deformities, experience cephalopelvic disproportion and, probably consequently, deliver via Caesarean section (313,317). Many of these women also experienced uterine dysfunction and/or underwent external version suggesting that the fetus was in an abnormal position. These are also both indications for Caesarean section (99,318–321).

Information that does not influence the structure of the delivery episode graph can be used to colour nodes to highlight more relationships in the data. The main (largest) graph created using data from Walker card 0 was overlaid with information on the mode of delivery, which showed that deliveries cluster according to whether they were assisted (Caesarean or instrumental deliveries) or unassisted (normal, vaginal deliveries). The graph was then overlaid with information on the duration of labour, which showed that labours lasting over 24 hours tend to cluster. Observing the chart of average values for each obstetric characteristic shows that long labours are more often associated with uterine dysfunction, which is a well-known association (38). By comparing with the graph overlaid with information on mode of delivery, it is also clear that long labour is more often associated with assisted rather than unassisted deliveries, which reflects the common clinical practice to resort to instrumental or surgical delivery in cases of difficult prolonged labour (99,322,323). Finally, the graph was overlaid with information on the position of the fetus, which showed that deliveries in which the fetus was in the most common "normal" left/right occiput anterior position clustered separately from deliveries in which the fetus was in a less common occiput lateral or posterior position. Charting

average values for each obstetric characteristic shows that abnormal positions are more commonly associated with pelvic deformity and cephalopelvic disproportion. By comparing with the other graphs, it is also possible to visually observe that the occiput lateral or posterior position is more frequently associated with labour over 24 hours and assisted delivery, as shown previously (319,320).

Although the network graph approach did not highlight any novel associations, it provided a way to quickly view relationships within the data and was particularly useful in identifying non-informative variables. For example, data on maternal and paternal blood group was initially included in the analyses, but was removed after observation of the resulting 'obstetric characteristic' network graphs suggested the data contained errors. The graphs showed that all maternal and paternal blood groups were forming two clusters, separate from any other characteristic. This would suggest that the blood groups were correlated (i.e. that it is possible for one patient to have more than one blood group). Further observation of the data showed that many mothers/fathers had been erroneously recorded as having two or more blood groups. Therefore, information on blood group was judged to be meaningless and was removed. The network graph approach may be more successful at highlighting novel associations if applied to a dataset in which less is known about the associations between its variables. In pregnancy research, such a dataset might contain variables on pregnancy exposures such as over/under the recommended gestational weight gain, gestational diabetes and preeclampsia, and the occurrence of each of these in their own daughters' pregnancies. In this way, the network graph approach could be used to highlight trends in intergenerational transmission of several obstetric complications simultaneously. Alternatively, epigenetic effects of pregnancy exposures could be investigated using this approach and a dataset of methylation of specific CpG sites in neonatal cord blood samples and pregnancy exposures such as those previously mentioned.

The network graph approach aims to create graphs that are informative and comprehensive, therefore a number of considerations should be made to optimise the size and amount of useful information included in the graph. Clearly, the graphs are dependent on the data used to build them, so care should be taken to prepare the input files appropriately. Omission of a key variable, or inclusion of a less useful/meaningless variable (such as parental blood groups in this example) can potentially mask interesting,

real associations. Similarly, selecting a Pearson's correlation cut-off value that is too stringent or too relaxed can lead to interesting associations not appearing in the graphs, or being hidden by weaker associations. Additionally, once the graph has been loaded, MCL clustering is highly dependent on the "inflation value" selected to run the algorithm. Lower inflation values result in fewer, larger, noisier clusters and higher inflation values result in many, smaller, cleaner clusters. As was the case in the development of the graphs shown here, a range of inflation and Pearson's correlation cut-off values should be explored when building network graphs. The size of the dataset is a key factor in producing useful network graphs. The obstetric complication graph was built using data on 48985 patients for 108 variables, but low correlation between most of the variables meant the final graph contained only 54 nodes (variables), and only 22 of these were assigned to MCL clusters. Therefore, we do not think this approach would be useful in smaller datasets, unless variables within the datasets correlate strongly.

Within Biolayout Express3D, network graphs are interactive; they can be viewed from any angle or zoom level and clicking on a node will reveal more information about how it relates to the rest of the data. Since network graphs are best viewed in this way, there are clear advantages to publishing the graphs online. Currently, the obstetric characteristic graph could potentially be made available to download to run in Biolayout Express3D (which is free and available to anyone). This would allow other researchers to experiment with other Pearson value thresholds and inflation values and therefore facilitate validation of the approach. The .matrix file (a correlation matrix) used to build the obstetric characteristic graph contains no patient-level data, so publishing it online would not breach the data protection rules set out by HIC. However, the delivery episode graph is built using a .expression file that can be opened in a program such as MS Excel to reveal the raw data. Therefore, HIC (who own the data, but also have a duty to protect the privacy of the cohort participants) would not currently permit this.

In conclusion, network graphs provide a quick and effective way to preliminarily explore relationships between risk factors, outcomes and confounders in an unbiased data-driven manner, but require large datasets of over or around 50,000 cases

6 Intergenerational transmission of postpartum haemorrhage risk – analysis of two Scottish birth cohorts

6.1 Introduction

The previous chapter describes analyses of birth record data from the Walker cohort. The historical nature of this dataset allows these records to be linked with a large number of current health-outcome databases, and allows linkage of records over generations. This chapter describes the results of a record-linkage study in which the Walker cohort was linked to a current database of maternity information, SMR02, to investigate intergenerational transmission of postpartum haemorrhage (PPH).

PPH is defined as ≥500ml blood loss from the genital tract in the first 24 hours after childbirth. It is the leading cause of maternal death worldwide, occurring in around 7-26% of all deliveries (63) and contributing to the deaths of an estimated 125,000 women each year (324). The annual incidence of PPH appear to be rising steadily, even in high resource countries (100). Known risk factors, causes and consequences of PPH are summarised in Figure 6.1, however the aetiology is often unclear and PPH may occur in women with no identifiable risk factors. PPH can be associated with a failure of the uterus to contract adequately after birth (atonic PPH; 90% of cases), trauma to the genital tract (traumatic PPH; 7% of cases), or bleeding due to retention of placental tissue or failure in the coagulation system (3% of cases) (101). Given the large percentage of PPH cases attributed to uterine atony and the focus on myometrial function in previous chapters, we also considered investigating intergenerational transmission of dystocia. However, dystocia was only defined as "labour over 24 hours" in the Walker cohort, and too few cases could be linked to SMR02 to provide adequate power to conduct this analysis. Therefore, we focussed our analyses on PPH.

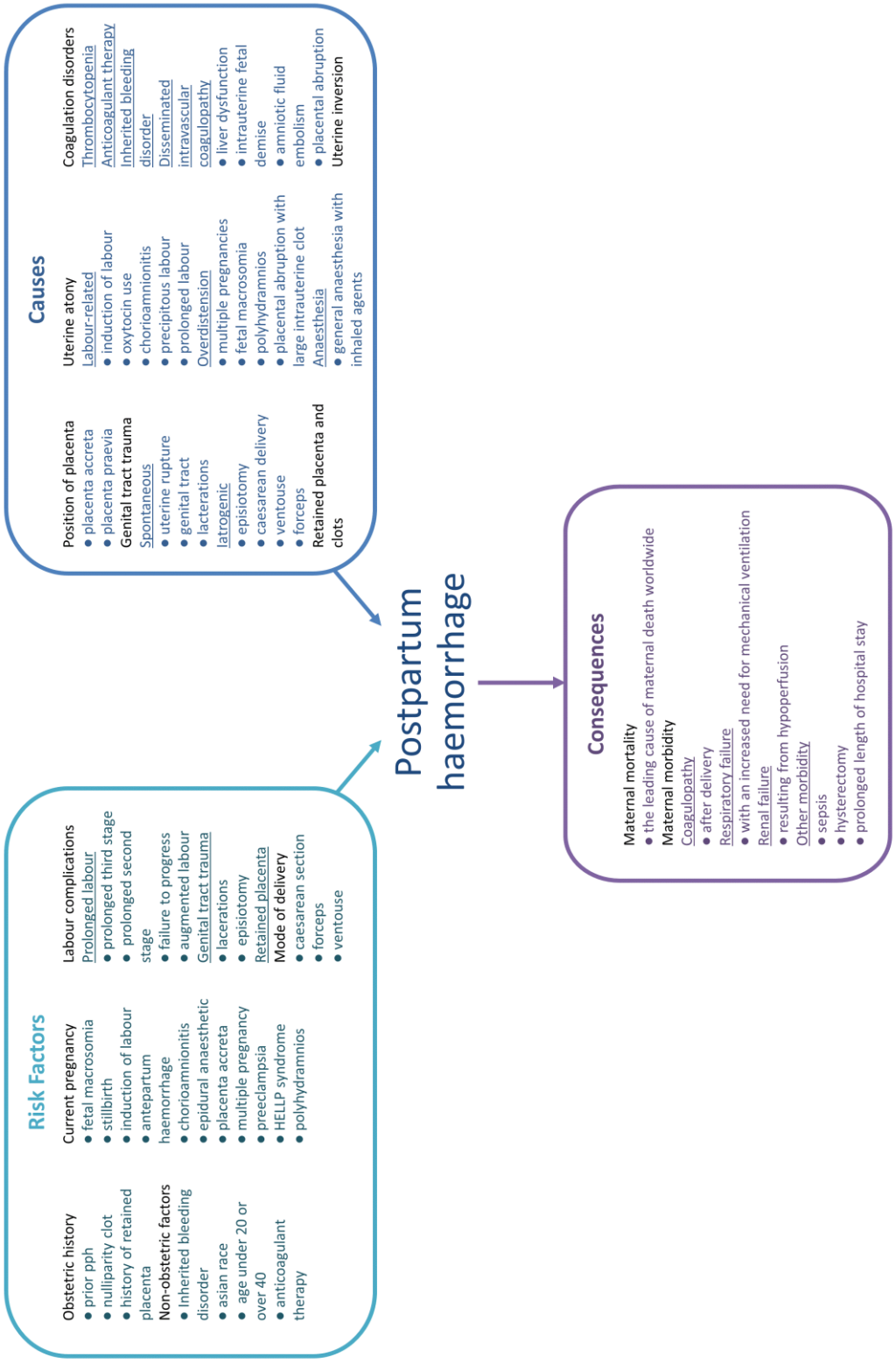


Figure 6.1. Causes, risk factors and consequences of postpartum haemorrhage, as identified previously.

Previous PPH is a significant risk factor for subsequent PPH, with several studies finding women two to three times more likely to have PPH in their second pregnancy if they had PPH in their first (325–328). If individual women are at increased risk, it is possible that this predisposition could be heritable, but no studies have previously addressed this. Understanding the biological and potentially heritable basis to PPH could be useful in understanding the aetiology of this important obstetric complication and developing better predictive and preventive tools. Additionally, it would help in the counselling of pregnant women, who are often aware of their family history of pregnancy related adverse events, including PPH.

Here, Scottish population data was used in which quality and consistency has previously been confirmed, and where database linkage is possible. This allowed patient-based analysis and analysis of intergenerational transmission over three generations of women.

6.2 Summary of methods

The methods for this chapter are described in detail in Chapter 2. Briefly, the Walker dataset of birth records collected between 1952-1966 was linked via the Community Health Identifier (CHI number) to the Scottish Morbidity Records Maternity Admissions dataset (SMR02) collected from 1975 to present (described in sections 2.4.1 to 2.4.4). Data was cleaned and continuous variables were used to create categorical variables (described in sections 2.4.7 and 2.4.8). Obstetric and demographic risk factors for PPH were assessed in the Walker and SMR02 datasets separately and as a pooled dataset by calculating unadjusted and adjusted odds ratios (described in section 2.4.9.2). To assess intergenerational transmission of PPH, generations were linked as follows:

- Generation 1 (Walker Mothers – Women who appear in the Walker cohort as mothers) was linked to Generation 2 (SMR02 Mothers – Women who appear in the Walker cohort as babies, and the SMR02 cohort as mothers).
- Generation 2 was linked to Generation 3 (SMR02 Daughters – Women who appear in the SMR02 cohort as babies, and also as mothers if they have had children themselves)
- Generation 1 was linked to Generation 3.
- Pooled mother-daughter analysis: mothers from Generations 1 and 2 were linked to daughters in Generations 2 and 3.

For each of these comparisons, logistic regression models were used to calculate unadjusted ORs to describe the relationship between PPH in the younger generation and PPH in the older generation. Generalised linear mixed models (GLMMs) were used to adjust for the non-independence arising from the appearance of the same woman in different (grand)mother-(grand) daughter pairs, as well as known risk factors for PPH identified through the unlinked analyses (described in section 2.4.9.3). In SMR02 only, ICD codes made it possible to determine which type of PPH a woman experienced, however there were too few cases to allow us to conduct subgroup analyses with adequate power. Furthermore, the Walker data did not contain information on the type of PPH. Therefore, in all analyses PPH was defined as "any type of PPH".

6.3 Results

6.3.1 Prevalence of PPH

Figure 6.2 outlines how records were linked in this study and the number of records used in the final analysis.

The overall prevalence of PPH (1089/25322, 4.3%) was similar in both the Walker (176/3847, 4.6%) and SMR02 (913/21475, 4.3%) cohorts. 82.3% (751) of cases of PPH in SMR02 deliveries were caused by uterine atony. PPH was diagnosed as delayed or secondary in 8.9% (81) of SMR02 cases and associated with retained placenta (third stage) in 8.4% (77). Coagulation defect was the least common recorded cause of PPH (0.4%, 4 cases).

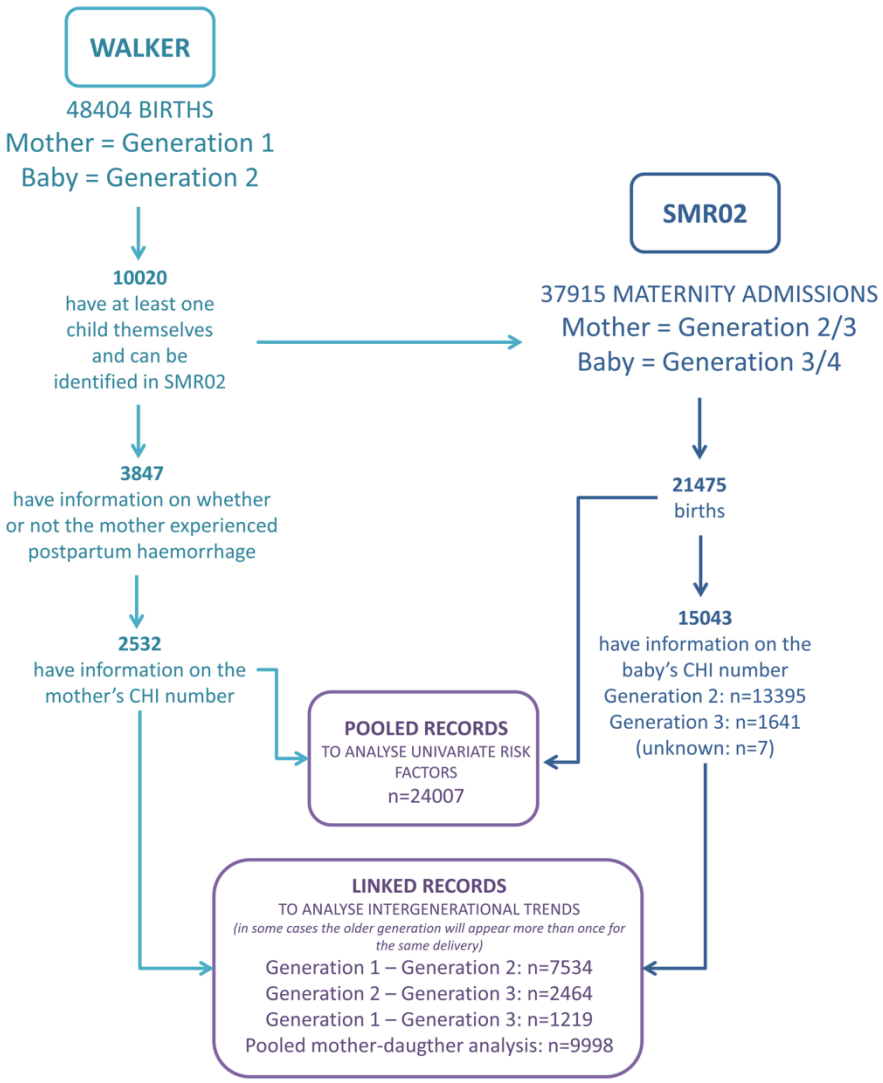


Figure 6.2. Method of record linkage and number of records analysed.

6.3.2 Effect of data cleaning

Data were cleaned as described in Chapter 2, section 2.4.7. Table 6.1 and Table 6.2 show the effects of data cleaning on the distribution of three key continuous variables in the SMR02 and the Walker dataset, respectively. This shows that data cleaning gave the data more 'reasonable' distributions, based on clinical knowledge.

Table 6.1. The effects of data cleaning on the distribution of SMR02 variables.

	Before cleaning				After cleaning			
	Mean	Median	Min.	Max.	Mean	Median	Min.	Max.
Birth weight (g)	2244.4	3010	0	9999	3309.4	3340	250	5710
Maternal age	27	25	14	48	26	25	14	48
Gestational age (weeks)	31.7	39	0	99	39.6	40	12	44

Table 6.2. The effects of data cleaning on the distribution of Walker variables.

	Before cleaning				After cleaning			
	Mean	Median	Min.	Max.	Mean	Median	Min.	Max.
Birth weight (g)	3275.7	3280	907.2	39973.5	3245.5	32319	1332	39974
Maternal age	26	25	2	228	27	26	16	228
Gestational age (weeks)	39.7	40	26	48	39.7	40	26	48

6.3.3 Cohort effect

Univariate analysis of the data was performed to quantify the effect of known and unknown risk factors for PPH. Univariate analysis of the Walker and SMR02 datasets separately showed different risk factors for PPH (Table 6.3), with high birth weight being the only common risk factor. To maximise the size of the dataset, and therefore statistical power, all further analysis was conducted using a 'pooled' dataset, combining SMR02 and Walker.

Table 6.3. Analysis of risk factors associated with PPH in separate Walker and SMR02 dataset analyses.

Risk factor	WALKER - unadjusted odds ratio* (95% confidence interval)	SMR02 – unadjusted odds ratio* (95% confidence interval)
Multiple pregnancy	0.422 (0.06 to 3.08) *ns*	2.27 (1.72 to 3.01)
Low birth weight (≤2.5kg)	0.68 (0.34 to 1.34) *ns*	1.20 (0.94 to 1.53) *ns*
High birth weight (≥4kg)	1.73 (1.01 to 2.94)	2.48 (2.09 to 2.94)
Maternal age ≤20-years-old	1.42 (0.81 to 2.49) *ns*	13.2 (11.9 to 14.7)
Maternal age ≥40-years-old	0.82 (0.299 to 2.26) *ns*	3.62 (2.56 to 5.11)
Nulliparity	1.03 (0.76 to 1.40) *ns*	1.20 (1.05 to 1.37)
Preterm birth (≤37 weeks' gestation)	1.03 (0.52 to 2.07) *ns*	1.36 (1.12 to 1.66)
Postterm birth (≥42 weeks' gestation)	1.63 (0.95 to 2.8) *ns*	0.70 (0.54 to 0.91)
Delivery by Caesarean section†	1.37 (0.53 to 3.56) *ns*	7.35 (5.96 to 9.08)
Instrumental delivery (forceps or ventouse)	0.47 (0.18 to 1.21) *ns*	0.71 (0.40 to 1.23) *ns*
Genital trauma or episiotomy	1.48 (0.91 to 2.39) *ns*	3.91 (1.50 to 10.20)
Mother smoked during pregnancy	n/a	0.07 (0.06 to 0.09)

ns: non significant.

* ratio of the odds of a birth being affected by PPH when the risk factor is present to the odds of a birth being affected by PPH when the risk factor is absent.

6.3.4 Univariate analysis of pooled data.

In univariate analyses of data pooled from Walker and SMR02 (Table 6.4), multiple pregnancy, baby birth weight over 4kg, maternal age over 40-years, preterm gestation ≤37 weeks, Caesarean delivery, nulliparity, genital trauma/episiotomy and smoking during pregnancy were significant risk factors for PPH (ORs ranging from 1.17 to 6.02). Delivery by forceps or ventouse was associated with a small but significant lower risk of PPH. There was insufficient data on PPH in a previous pregnancy to determine if this was a risk factor for PPH in a subsequent pregnancy.

Table 6.4. Analysis of risk factors associated with PPH in pooled (Walker and SMR02) analyses.

Risk factor	(A) Number of births with information on PPH and risk factor	(B) Number of births with PPH (% of column A)	(C) Number of births with PPH where risk factor is present (% of column B)	Unadjusted odds ratio* (95% confidence interval)	Adjusted† odds ratio* (95% confidence interval)
Multiple pregnancy	25322	1089 (4.3%)	58 (5.3%)	2.09 (1.59 to 2.76)	1.87 (0.77 to 4.53) *ns*
Low birth weight (≤2.5kg)	24935	1059 (4.2%)	81 (7.6%)	1.11 (0.88 to 1.39) *ns*	n/a
High birth weight (≥4kg)	24935	1059 (4.2%)	199 (18.8%)	2.37 (2.02 to 2.78)	1.84 (1.27 to 2.67)
Maternal age ≤20-years-old	20207	1015 (5.0%)	136 (13.4%)	1.34 (1.12 to 1.62)	1.13 (0.77 to 1.65) *ns*
Maternal age ≥40-years-old	20207	1015 (5.0%)	44 (4.3%)	2.73 (1.98 to 3.77)	1.32 (0.73 to 2.38) *ns*
Nulliparity	25293	1085 (4.3%)	541 (49.9%)	1.17 (1.03 to 1.32)	1.47 (1.10 to 1.98)
Preterm birth (≤37 weeks' gestation)	23741	1003 (4.2%)	127 (12.7%)	1.33 (1.1 to 1.61)	0.63 (0.41 to 0.97) *ns*
Postterm birth (≥42 weeks' gestation)	23741	1003 (4.2%)	81 (8.1%)	0.79 (0.63 to 1.00) *ns*	n/a
Delivery by Caesarean section†	3872	471 (12.2%)	283 (60.1%)	6.02 (4.92 to 7.38)	8.20 (6.19 to 10.86)
Instrumental delivery (forceps or ventouse)	3872	471 (12.2%)	21 (4.5%)	0.42 (0.27 to 0.66)	n/a
Genital trauma or episiotomy	18890	885 (4.7%)	25 (2.8%)	1.61 (1.07 to 2.44)	9.61 (2.15 to 43.02)
Mother smoked during pregnancy	8833	573 (6.5%)	140 (24.4%)	1.40 (1.15 to 1.71)	0.79 (0.57 to 1.08) *ns*

ns: non significant.

* ratio of the odds of a birth being affected by PPH when the risk factor is present to the odds of a birth being affected by PPH when the risk factor is absent.

†adjusted for all factors identified as significant in the univariate (unadjusted) analysis.

6.3.5 Multivariate analysis of pooled data

A logistic regression model incorporating significant risk factors from the univariate analysis allowed us to adjust for confounding and revealed that large birth weight, Caesarean delivery, nulliparity and genital trauma/episiotomy were significant independent risk factors for PPH (Table 6.4). After these adjustments (particularly for multiple pregnancy and Caesarean section, which are significant confounders), delivery at

≤37 weeks was associated with a significant decreased risk of PPH (OR 0.63 95% CI 0.55-0.97).

6.3.6 Intergenerational linked analyses

Table 6.5 shows that there is a small increased risk of PPH in women whose mothers and/or grandmothers had PPH across generations 1-2 and 1-3, but this trend did not reach statistical significance. Comparisons of generations 2 and 3 and pooling of mother and daughter comparisons showed a reverse trend, i.e a trend to a protective effect of maternal PPH on the risk of PPH in the daughter. This highlights the cohort effect observed in the data; generations 2 and 3 are comprised solely of women from SMR02. GLMMs were used to adjust for non-independence between related mother-daughter pairs and most risk factors identified as significant by the multivariate analysis (analyses could not be completely adjusted for delivery by Caesarean section due to incomplete data). These analyses again confirmed no statistically significant effect of maternal PPH on the risk of PPH in the daughter. These intergenerational analyses had 80% power at an alpha of 0.05 to detect the following odds ratios: 2.1 for generation 1 linked to generation 2; 1.3 for generation 2 linked to generation 3; 2.1 for generation 1 linked to generation 3; and 1.3 for the pooled mother to daughter analysis. Thus we can be reasonably confident that any intergenerational effect of maternal PPH, should it exist, is not greater than an odds ratio of PPH in the daughter of 1.3.

Table 6.5. Analysis of intergenerational trends in PPH.

Risk factor	(A) Number of linked births with information on PPH	(B) Number of linked births with PPH in younger generation (% of column A)	(C) Number of linked births with PPH in both generations (% of column B)	Unadjusted odds ratio (95% confidence interval)	Adjusted* odds ratio (95% confidence interval)
PPH in generation 1 as a risk factor for PPH in generation 2	2543	49 (1.9%)	3 (6.1%)	1.32 (0.41 to 4.32)	1.20 ($4.83x10^{-3}$ to 296.0)
PPH in generation 2 as a risk factor for PPH in generation 3	2464	290 (11.8%)	4 (1.4%)	0.68 (0.24 to 1.90)	0.58 ($8.0x10^{-3}$ to 41.61)
PPH in generation 1 as a risk factor for PPH in generation 3	519	65 (12.5%)	6 (9.2%)	2.21 (0.85 to 5.72)	1.33 ($9.43x10^{-5}$ to 1.88e+4)

PPH in mothers as a risk factor for PPH in daughters (pooled analysis)	5007	339 (6.8%)	7 (2.1%)	0.59 (0.27 to 1.27)	0.69 (0.06 to 7.62)

*calculated using a generalised linear mixed model to adjust for the non-independence between linked births and risk factors identified as significant in the multivariate analysis (excluding delivery by Caesarean section because of incomplete data).

6.4 Discussion

To our knowledge, this is the first attempt to investigate the intergenerational transmission of PPH. The analyses do not support a large increased risk of PPH for women whose mothers/grandmothers had PPH. Caesarean delivery, genital trauma or episiotomy, high birth weight and nulliparity were identified as risk factors for PPH in univariate and multivariate analysis, thus confirming the results of previous studies (64,327,329–331). In particular, the odds of PPH in nulliparous women (odds ratio 1.47) was very similar to that reported by Combs *et al*. (odds ratio 1.45) (327).

Univariate analysis of the Walker and SMR02 datasets separately showed different risk factors for PPH. This may be due to the use of different measure and record data in the different cohorts. To maximise the size of the dataset and therefore the statistical power, analyses were conducted on a 'pooled' dataset, which means this 'cohort effect' was largely ignored. However, it is worth noting that due to its larger size, SMR02 has a greater impact on the results of these analyses compared to Walker.

In both the SMR02 and Walker cohorts, the prevalence of PPH (4.3% and 4.6%, respectively) was lower than that reported for other populations. In a meta-analysis of 104 datasets, Calvert *et al*. (63) showed that PPH prevalence shows high regional variation, ranging from 7.2% in Oceania to 25.7% in Africa. In Europe, they found a prevalence of 12.7%, which is similar to the 13.2% incidence reported in the NHS maternity records for England and Wales in 2011-12 (60). However, the authors also found that the prevalence depends strongly on the method of diagnosis of PPH, with a subjective measurement of blood loss resulting in a lower prevalence compared to an objective measurement. A subjective measure is likely to have been used for the SMR02 and Walker cohorts, which may explain the relatively low prevalence of PPH in these datasets.

In line with previous studies (64,332,333), the most frequent cause of PPH was uterine atony (82.3% of cases in SMR02), which prevents constriction of blood vessels during placental separation. Unfortunately there was insufficient data to analyse risk factors for different types of PPH individually.

No previous reports have investigated family history of PPH as a risk factor for PPH. The historical Walker data linked to the more recent SMR02 data presented a unique opportunity to do this. SMR02 data collection began in 1975, so the same comparison could be made within this dataset. Special consideration was given to the appearance of the same women in more than one mother-daughter/grandmother-granddaughter pair. This is further complicated by the tendency for women to experience PPH in repeat pregnancies (325–328). This non-independence invalidates the assumptions of many statistical tests and can lead to spurious conclusions. One option for dealing with this "clustering" is to restrict analysis to one pregnancy per woman (for example, the first pregnancy). However, this reduces statistical power and ignores a lot of potentially important information. It also changes the definition of the study population from "all births within the dataset" to "all first births within the dataset", so it may not be possible to generalise the results to "all births"(334). Another possible tactic is to include data on all pregnancies and ignore the non-independence. This approach was used to calculate the "unadjusted odds ratios" for intergenerational transmission of PPH. However, this will lead to incorrect standard errors and potentially incorrect conclusions. Therefore, a mixed model was used in the final, multivariate analysis to adjust for both covariates (fixed effects) and within-woman clustering (random effects). This protects against bias and allows estimation of the size of the effect introduced by this clustering (334).

No significant association was shown between PPH in the mother and the odds of PPH in daughters. This study had 80% power to detect an OR of 1.29 for maternal influence on PPH in the daughter. This is a lower OR than conferred by birthweight > 4.0kg and nulliparity (1.87 and 1.47 respectively) and very much lower than conferred by maternal Caesarean section (8.2) and genital tract trauma (2.6). Thus any effect of the pregnant woman's maternal history of PPH is (if it exists) much less significant than those of the index pregnancy. These data contrast with the known intergenerational transmission of pre-eclampsia and of preterm delivery (335,336). Again, the 'cohort effect' was apparent– in the analyses that involved data from the Walker cohort (generation 1) there was a slight

increase in odds of PPH in women whose mothers had experienced PPH. However, these observations were not significant so they did not change the overall conclusions of the analyses.

Pregnant women whose mothers had PPH can be reassured that they are unlikely to be at any significantly increased risk, compared to those whose mothers did not have PPH.

7 FETAL ASSESSMENT METHODS FOR IMPROVING NEONATAL AND MATERNAL OUTCOMES FOLLOWING PRETERM PRELABOUR RUPTURE OF MEMBRANES – A SYSTEMATIC REVIEW

7.1 INTRODUCTION

The previous two chapters describe how epidemiology can make use of patient health records to explore complex relationships between risk factors and outcomes of obstetric complications. Epidemiology is also concerned with how clinical interventions affect the health of a population. As discussed in Chapter One, the efficacy of medical interventions is assessed through randomised controlled trials (RCTs), but the reliability of the evidence is greatly increased through meta-analysis of all available data from comparable RCTs. This chapter describes a Cochrane systematic review and meta-analysis of RCT data to assess the efficacy of different methods of fetal assessment in improving neonatal and maternal outcomes following preterm prelabour rupture of membranes (PPROM).

7.1.1 DESCRIPTION OF THE CONDITION

Preterm prelabour rupture of membranes (PPROM) is defined as rupture of the chorioamniotic membranes before 37 weeks' gestation, where there is at least one hour between rupture of membranes and the onset of contractions. It occurs in around 1% to 2% of pregnancies and is associated with around 30% to 40% of preterm births (61,337,338).

PPROM is often idiopathic, but it has been associated with a number of factors including a history of previous preterm delivery or PPROM, vaginal bleeding during pregnancy, uterine overdistension (76), black race (74), smoking, cervical cerclage (75) amniocentesis, infection, and low socioeconomic status (77).

The mechanisms responsible for PPROM are largely unknown, but might include excessive stretching of the membranes, decreased collagen content, placental abruption or programmed amniotic cell death (78). Between one-quarter and one-half of cases are associated with intrauterine infection and inflammation (79).

PPROM is associated with fetal and maternal morbidity and mortality. Cord prolapse, cord compression, placental abruption and maternal/neonatal infection are potential complications (79), and many women require interventions to expedite delivery, with induction of labour or Caesarean section. The major cause of perinatal morbidity following PPROM arises from preterm birth, which is a significant problem because most women will go into spontaneous labour within several days of PPROM (66). The neonatal morbidity associated with preterm birth includes respiratory distress syndrome, intraventricular haemorrhage, and infection (78).

Several Cochrane reviews have assessed interventions for improving outcomes following PPROM: Kenyon *et al*. (339) found some evidence that antibiotics may increase time to labour and decrease the risk of infection, but appear to have no effect on mortality. Mackeen *et al*. (83) found tocolytics increase time to labour but may also increase risk of chorioamnionitis, low Apgar scores and need for ventilation. Hofmeyr *et al*. (82) found transabdominal amnioinfusion decreased risk of neonatal sepsis, infection and death, but the evidence was insufficient to recommend routine use. Transcervical amnioinfusion improved fetal heart rate patterns and umbilical cord blood pH results but did not significantly improve substantive clinical outcomes. Abou El Senoun *et al*. (340) found home care is associated with a lower rate of Caesarean section compared with hospital care, although there was only a small effect on outcomes such as perinatal mortality, neonatal and maternal infection, latency period and neonatal admission to intensive care.

7.1.2 Description of the fetal assessment as an intervention

This review focuses on fetal assessment methods following PPROM to improve pregnancy outcomes. Although fetal assessment does not aim to "treat" the mother or the fetus, it is a way for medical professionals to monitor fetal wellbeing and aid their decisions about how to manage cases of PPROM. Fetal assessment may result in earlier delivery due to earlier detection of fetal compromise, which could improve neonatal and maternal outcomes. However, if the fetal assessment tool is inaccurate, it may cause inappropriate early delivery and worsen outcomes.

There are several methods for assessing fetal wellbeing, including:

- Fetal movement counting - mothers count fetal movements over a specified period or measure the time it takes for the fetus to make a specified number of movements. Reduced frequency of movements can be a sign of fetal compromise.
- Fetal cardiotocography - allows monitoring of the fetal heart rate over time. Changes in heart rate parameters outside of the accepted 'normal' limits can identify fetuses at risk of acute or chronic fetal hypoxia.
- Biophysical profile (BPP) - an assessment of overall fetal wellbeing through the combined assessment of fetal cardiotocography along with ultrasonic measurement of fetal movements, fetal tone, fetal breathing and estimation of amniotic fluid volume.
- Modified biophysical profile (MBPP) - less time-consuming than the BPP, the MBPP is an assessment of overall fetal wellbeing based only on fetal cardiotocography and estimation of amniotic fluid volume.
- Fetal and umbilical artery or venous Doppler ultrasound - ultrasonic measurement of blood flow through blood vessels of interest can indicate high vascular impedance and possible feto-placental compromise.
- Fetal MRI lung volumetrics - fetal lung volume is measured by MRI and used as an indication of fetal lung development and pulmonary hypoplasia.
- Amniocentesis for fetal lung maturity - levels of fetal lung surfactant in the amniotic fluid are used to gauge fetal lung maturity.

Several Cochrane reviews have examined the use of some of these tests in improving maternal and fetal outcomes in normal and compromised pregnancies (341–345). The scope of this review is to examine these assessment methods in improving outcomes in pregnancies affected by PPROM. Although some of these methods may be useful in identifying intrauterine infection (346–350), their effectiveness in improving outcomes following PPROM is uncertain.

7.1.3 Why it is important to do this review

Following PPROM, accurate methods for assessing fetal wellbeing and prognosis are needed to aid obstetrician's decisions in planning the time and mode of delivery. However, the value of the various methods of fetal assessment in ultimately improving neonatal and maternal outcomes has yet to be established.

7.1.4 Objectives

This review aims to compare methods of fetal assessment in improving outcomes following preterm prelabour rupture of membranes.

7.2 Summary of methods

Detailed methods of this review are described in Chapter 2. Briefly, the review was carried out according to Cochrane guidelines (described in section 2.5.1). Firstly, a detailed protocol was produced outlining the methods to be used (described in section 2.5.2). The Cochrane Trials Search Co-ordinator searched the Cochrane Pregnancy and Childbirth Group's Trials Register on 28th November 2012 (described in section 2.5.4). Criteria for considering studies for inclusion are described in section 2.5.3. All relevant randomised controlled trials involving women with PPROM before 37+0 weeks' gestation undergoing any method of fetal assessment (including no intervention) were included in the review. Primary outcomes of interest were fetal death (antenatal), neonatal death (in the first 28 days of life), maternal death, and serious maternal morbidity (septicaemia, need for intensive care, organ failure/need for ventilation, or need for hysterectomy). A complete list of secondary outcomes is provided in Table 2.6. Data was extracted (described in section 2.5.5) from the trial reports using a custom data extraction form (Appendix 7). For each RCT, the risk of bias was assessed (described in section 2.5.5.3). Statistical analysis was carried out using the Review Manager software to provide summary risk ratios (RRs) with 95% confidence intervals for dichotomous data, and mean differences for continuous data (described in section 2.5.5.4). Heterogeneity of treatment effects was also measured between trials (described in section 2.5.5.6).

This work was conducted by a team of three authors, although my contribution was the largest (precise author contributions are described in section 2.5.1.1).

7.3 Results

7.3.1 Results of the search

The search of the Cochrane Pregnancy and Childbirth Group's Trials Register identified four studies in six reports. Three studies (Carlan (1997) (351) Cotton (1984) (352) and Lewis (1999) (353)) reporting data on 271 women were included in the review. One study (Li *et al*., 2010 (354)) was excluded from the review because all the patients were over

37+0 weeks' gestation. Figure 7.1 shows a summary of information flow in the selection of papers for the review.

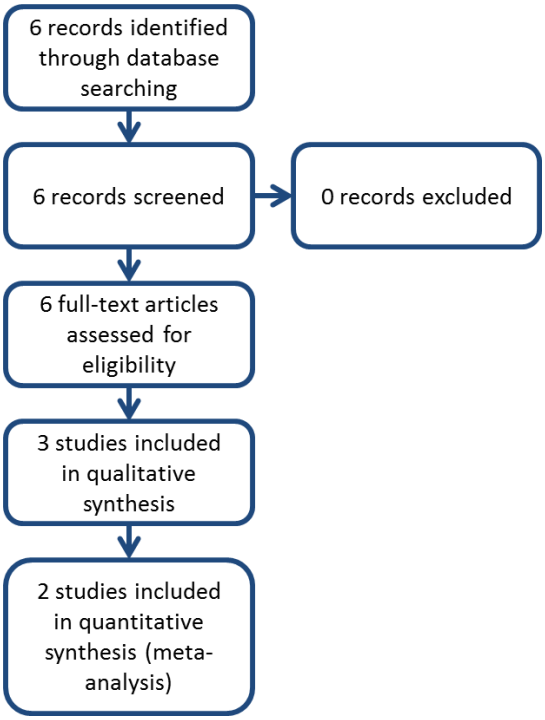


Figure 7.1. The summary of information flow in the selection of papers for the review.

7.3.1.1 Included studies

Details of the three included studies are described in Table 7.1, Table 7.2 and Table 7.3. Lewis (1999)(353) was associated with three publications: one abstract (355), one answer from Lewis *et al*. in response to a letter regarding the abstract (355), and one full article describing the results of the trial (353).

Table 7.1. A description of the key features of the Carlan (1997)(351) study.

	Carlan *et al*. (1997)(351)
Methods	A single centre unmasked randomised controlled trial conducted from May 1993 until June 1996 (the study was interrupted from March 1994 until December 1994 due to technical difficulties with the sonography machine).
Participants	93 women (92 available for analysis) with a diagnosis of PPROM between 24+0 and 34+0 weeks' gestation at Arnold Palmer Hospital for Children and Women (Orlando, Florida, USA).
Interventions	Weekly endovaginal ultrasound scans versus no endovaginal ultrasound scans.
Outcomes	• Neonatal death.

	• Maternal chorioamnionitis (diagnosed when maternal temperature reached 100.4°F, the patient exhibited uterine tenderness, and there was fetal or maternal tachycardia). • Maternal endometritis (diagnosed when maternal temperature reached 100.4F on two or more occasions during a 6-hour period after the first 24 hours postpartum, in association with uterine tenderness and no other source of infection). • Caesarean delivery. • Respiratory distress syndrome (determined by clinical presentation (oxygen therapy for more than 24 hours) and radiographic findings of a grapund-glass appearance with the presence of an air bronchogram). • Days of postnatal hospitalisation. • Days from randomisation to birth. • Birthweight. • Days from birth to discharge home. • Spontaneous labour before 34+0 weeks. • Maternal antibiotics after delivery. • Neonatal antibiotics. • Neonatal infection (diagnosed if there was fever, leukocytosis, or a positive culture of blood, urine, or gastric aspirate). • Corticosteroid use. • Number with Apgar over 5 at 5 minutes. • Intraventricular haemorrhage (defined as a haemorrhage within the germinal matrix, ventricular system of the brain, or intracerebral tissue. Diagnosed using cranial ultrasound). • Time from rupture to admission (days). • Time of maternal antibiotic treatment before labour (hours). • Cord blood pH.
Notes	The trial addressed a further outcome (the relation of cervical length to the time of labour in women who went into spontaneous labour), but this was not included in the review because it did not compare the effect of the intervention.
Risk of bias associated with sequence generation	Unclear risk: sequences were "randomly generated", but the method is not described.
Risk of bias associated with allocation concealment	Low risk: allocations were concealed in sealed envelopes.
Risk of bias associated with blinding method	High risk: the nature of the intervention prohibited participant and clinician blinding. It is presumed that the outcome assessor was also not blinded, because the outcomes include cervical

	length measurement, which was only available for the treatment group.
Risk of bias associated with incomplete outcomes	Low risk: loss of participants to follow-up and exclusion of participants after randomisation are described adequately. Only one participant was excluded after randomisation, representing a low risk of bias.
Risk of bias associated with selective reporting	Low risk: there is no suggestion of selective reporting of outcomes. All main outcomes described in the methods are reported in the results.
Risk of other bias	Low risk: no reason to suspect other bias.

Table 7.2. A description of the key features of the Cotton (1984) (352) study.

Cotton *et al*. (1984)	
Methods	A single centre unmasked randomised controlled trial.
Participants	47 women (44 available for analysis) with a diagnosis of PPROM between 26+0 and 34+0 weeks' gestation at Hermann Hospital (Houston, Texas, USA).
Interventions	Amniocentesis versus no amniocentesis.
Outcomes	• Latency period (days from rupture of membranes to birth; reported as a count of patients in different latency period categories). • Gestation at birth. • Birthweight. • Apgar scores. • Fetal death. • Neonatal death. • Neonatal sepsis. • Fetal sepsis. • Respiratory distress syndrome. • Transient tachypnea. • Intraventricular haemorrhage. • Pneumothorax. • Persistent ductus arteriosus. • Seizures. • Hyperbilirubinaemia. • Metabolic problems. • Apnea. • Feeding difficulties. • Days of postnatal hospitalisation (after discharge of mother). • Cardiotocographic abnormality in labour (fetal distress). • Maternal antepartum hospitalisation. • Maternal postpartum hospitalisation. • Maternal postpartum endometritis.

Notes	Days of postnatal hospitalisation could not be included in the review because data was presented as the median and range, which was insufficient to calculate the mean difference between the groups.
Risk of bias associated with sequence generation	Unclear risk: not described.
Risk of bias associated with allocation concealment	Unclear risk: not described.
Risk of bias associated with blinding method	High risk: the nature of the intervention prohibited participant and clinician blinding. Outcome assessor blinding is not described, therefore the risk of bias arising from this is unclear.
Risk of bias associated with incomplete outcomes	Low risk: loss of participants to follow-up and exclusion of participants after randomisation are described adequately.
Risk of bias associated with selective reporting	High risk: several outcomes listed in the methods section are not described in the results section. Some outcomes are reported only as showing no difference between the two groups, without any quantitative data.
Risk of other bias	Unclear risk: the number of patients with multiple gestations is not reported .These patients were not excluded from the study, but a subgroup analysis of their effect has not been described. Therefore the risk of bias arising from the inclusion of patients with multiple gestations in the analyses is unclear.

Table 7.3. A description of the key features of the Lewis (1999)(353) study.

Lewis *et al*. (1999)	
Methods	A single centre unmasked randomised controlled trial conducted during a 36 month period.
Participants	135 women with a diagnosis of PPROM ≤34+0 weeks' gestation at Louisiana State University School of Medicine, Shreveport, Louisiana, USA) and who were stable after 24 hours' monitoring.
Interventions	Daily nonstress test versus daily modified biophysical profile.
Outcomes	• Latency period (days from rupture of membranes to birth). • Days from birth to discharge home. • Birthweight • Neonatal antibiotic use (days) • Respiratory distress syndrome • Intraventricular haemorrhage • Necrotising enterocolitis • Bronchopulmonary dysplasia (interpreted in the review as oxygen treatment after 36+0 weeks) • Retinopathy of prematurity

	• Intra-amniotic infection (maternal chorioamnionitis) (diagnosed by analysis of amniotic fluid obtained from an amniocentesis or by a maternal temperature of 100.4°F, foul smelling fluid, and uterine tenderness). • Maternal endometritis • Neonatal sepsis • Presumed neonatal sepsis (diagnosed when there were clinical signs of infection with negative culture results and an abnormal leukocyte count). • Neonatal pneumonia (diagnosed by a positive radiographic finding plus evidence of sepsis). • Corticosteroid use. • Delivery for maturity. • Caesarean delivery. • Apgar score of 7 or less at 5 minutes.
Risk of bias associated with sequence generation	Low risk: random number tables with a 1:1 match were used to generate allocation sequences.
Risk of bias associated with allocation concealment	Low risk: allocation cards were kept in opaque envelopes and not opened until after informed consent was obtained.
Risk of bias associated with blinding method	High risk: the nature of the intervention prohibited participant or clinician blinding. Outcome assessor blinding is not described, therefore the risk of bias arising from this is unclear.
Risk of bias associated with incomplete outcomes	Low risk: no loss of participants to follow-up at each data collection point. No exclusion of participants after randomisation.
Risk of bias associated with selective reporting	Low risk: there is no suggestion of selective reporting of outcomes. All main outcomes described in the methods are reported in the results.
Risk of other bias	Unclear risk: the number of patients with multiple gestations is not reported. These patients were not excluded from the study, but a subgroup analysis of their effect has not been described. Therefore the risk of bias arising from the inclusion of patients with multiple gestations in the analyses is unclear.

7.3.1.1.1 Study design

All included studies were unmasked single-centre randomised controlled trials.

7.3.1.1.2 Sample sizes

Carlan (1997)(351) recruited 93 women (92 available for analysis), 46 were randomised to the intervention group and 47 to the control group. Cotton (1984)(352) recruited 47 women (44 available for analysis), 25 were randomised to the intervention group and 22

to the control group. Lewis (1999)(353) recruited 135 women (all 135 available for analysis), 66 were randomised to the group receiving daily modified biophysical profiling and 69 were randomised to the group receiving a daily nonstress test.

7.3.1.1.3 Study location

All included studies were conducted in the USA.

7.3.1.1.4 Participants

All included studies recruited women with a diagnosis of PPROM occurring ≤34+0 weeks. Lewis (1999) (353) only recruited women with this diagnosis who were stable after 24 hours' monitoring.

Carlan (1997) (351) excluded women with multiple pregnancies, cerclage in place, cervical dilatation on admission of more than 4cm, active regular contractions, or any indication for delivery. One woman was found to have a funic presentation (the appearance of the umbilical cord before the main presenting part of the fetus) and cervical dilation at the initial scan and was delivered by Caesarean immediately. This left 92 patients available for analysis (45 in the intervention group and 47 in the control group).

Cotton (1984) (352) excluded women with evidence of amnionitis, an indication for immediate delivery, sonographic evidence of gross fetal congenital anomalies, active labour with a cervical dilation of more than 4cm, or an inadequate amount of amniotic fluid for amniocentesis (assessed by real time ultrasound). In the treatment group, amniocentesis was unsuccessful in two patients and one patient withrdrew after the initial amniocentesis visit. This left 44 patients available for analysis (22 in each group).

Lewis (1999)(353) excluded women with signs of obvious clinical infection or indications for immediate delivery. No patients were excluded from analysis or lost to follow up, so 135 were available for analysis.

All three studies reported no significant difference between groups in maternal age (Carlan (1997)(351) data not reported; Cotton (1984) (352) data not reported; Lewis (1999) (353) mean 24.4 years in the nonstress test group and 25.7 years in the biophysical profile group) and parity (Carlan (1997)(351) data not reported; Cotton (1984) (352) data not reported; Lewis (1999) (353) 30.4% primigravid in the nonstress

test group and 19.7% primigravid in the biophysical profile group). Lewis (1999) (353) also reported no significant difference in ethnicity (27.5% White, 72.5% African American in the nonstress test group, 21.2% white, 78.8% African American in the modified biophysical profile group). Patient ethnicity was not reported for Carlan (1997)(351) or Cotton (1984) (352).

7.3.1.1.5 Types of intervention

Carlan (1997)(351) compared weekly endovaginal ultrasound scans versus no endovaginal ultrasound scans. Both also groups received intravenous antibiotics until genital culture results were available. If the culture results were positive, patients in both groups received intravenous antibiotics and were offered amniocentesis. Patients in both groups were offered steroids to accelerate pulmonary maturation until 31+0 weeks' gestation.

Cotton (1984) (352) compared amniocentesis versus no amniocentesis. Amniocentesis was performed at admission, 48 hours later, and then at weekly intervals until pulmonary maturity or bacteria in the amniotic fluid was identified, at which point the patient was delivered). Both groups also received bi-weekly electronic fetal heart rate monitoring. In the abasence of fetal pulmonary maturity, infection or fetal compromise, both groups received intravenous tocolysis. Steroids were also administered to both groups at the discretion of the managing physician.

Lewis (1999) (353) compared a daily nonstress test (a measurement of fetal cadiotocography) versus daily modified biophysical profile (as described by Chamberlain *et al.*(356)): total score of 10, with points each for tone, amniotic fluid, movement, breathing, and a nonstress test). Both groups also received intravenous antibiotics until delivery or for a total of seven days.

7.3.1.1.6 Outcome measures

None of the primary outcomes of this review (fetal/neonatal death, maternal death, and serious maternal morbidity) were addressed by all trials. Neonatal deaths (in the first 28 days of life) were reported for Carlan (1997)(351) and Cotton (1984) (352). From this, it was inferred that there were no fetal deaths (before birth) in either study. Lewis (1999) (353) did not report data for any of the primary outcomes.

Of the secondary outcomes, only maternal chorioamnionitis and respiratory distress syndrome were addressed by all three trials. Additionally, data on the frequency of intraventricular haemorrhage was also reported for all three trials. Accordingly, the review protocol was updated to include this as an outcome of interest, along with any other relevant outcomes described by each study but not identified in the original protocol. Interestingly, no study reported mean gestational age at delivery.

All addressed outcomes are listed in Table 7.1, Table 7.2 and Table 7.3.

7.3.2 Risk of bias in included studies

Figure 7.2 summarises judgements about overall risk of bias.

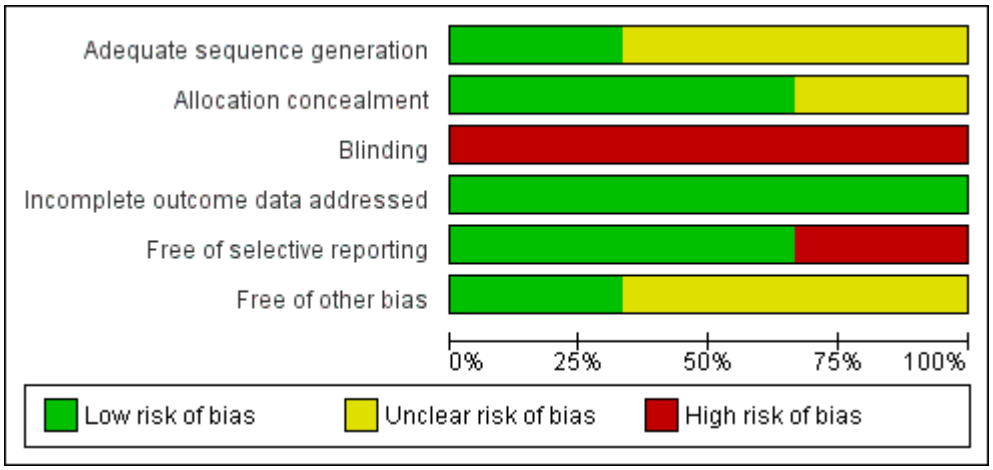


Figure 7.2. A summary of judgements about each risk of bias item presented as percentages across all included studies.

7.3.2.1 Allocation (selection bias)

Allocation sequence generation and concealment was not described for Cotton (1984) (352), so the risk of selection bias is unclear. For Carlan (1997)(351), the authors state that sequences were "randomly generated" but do not describe the method. Therefore the risk of bias can only be judged as unclear. However, the allocations were concealed in sealed envelopes, which is associated with a low risk of bias. For Lewis (1999) (353), there is an overall low risk of selection bias because random number tables were used to generate allocation sequences and allocations were concealed in opaque envelopes that were not opened until after informed consent was obtained.

7.3.2.2 Blinding (performance bias and detection bias)

The nature of the interventions meant that blinding of patients and clinicians was not possible. For Carlan (1997)(351) it is assumed that outcome assessors were also not blinded because one of the outcomes (cervical dilation) was only recorded for the treatment group. For Cotton (1984) (352) and Lewis (1999) (353), outcome assessor blinding is not discussed, so the risk of bias arising from this is unclear.

7.3.2.3 Incomplete outcome data (attrition bias)

The risk of attrition bias was judged to be low, because all three studies adequately described any loss of participants to follow up and any exclusion of participants after randomisation. Additionally, the level of attrition in all studies was low.

7.3.2.4 Selective reporting (reporting bias)

For Carlan (1997)(351) and Lewis (1999) (353) the risk of reporting bias is low because there is no suggestion of selective outcome reporting in either study. For Cotton (1984) (352), the risk of reporting bias is high because several outcomes listed in the methods section (gestation at delivery, birthweight and Apgar scores) are not described in the results section. Other outcomes (maternal antepartum and postpartum hospitalisation days, postpartum endometritis and fetal sepsis) are reported only as showing no difference between groups, without any quantitative data.

7.3.2.5 Other potential sources of bias

For Carlan (1997)(351), there is have no reason to suspect another potential source of bias. A potential source of bias in Cotton (1984) (352) and Lewis (1999) (353) arises from their treatment of patients with multiple pregnancies. Neither study excludes patients with multiple pregnancies but the reports do not mention how many (if any) patients had multiple pregnancies. If patients with multiple pregnancies are included in the analyses (without a subgroup analysis assessing their effect), then this is a potential source of bias.

7.3.3 Effects of interventions

Each study considered different comparisons of interventions, so these comparisons were considered separately in the review. It was also possible to consider a comparison of any assessment method versus no assessment method. This comparison included two studies (Carlan (1997)(351) and Cotton (1984) (352)). Lewis (1999) (353) could not be included in this comparison because the study did not consider the effect of no intervention.

Data on one outcome (days of postnatal hospitalisation) in one (Cotton (1984) (352)) study was not reported in a form that could be included in the meta-analysis (only the median and range were reported), this data was therefore omitted.

7.3.3.1 Weekly endovaginal ultrasound scan versus no endovaginal ultrasound scan.

Compared with the group receiving no fetal assessment, Carlan (1997)(351) found a significantly decreased relative risk (RR) of maternal endometritis in the group assessed by endovaginal ultrasound (RR 0.32, 95% CI 0.11-0.91). There was a non-significant decreased RR of maternal chorioamnionitis (RR 0.72, 95% CI 0.34-1.52) and neonatal respiratory distress syndrome (RR 0.71, 95% CI 0.4-1.27) in this group. However, there was also a non-significant increased RR of neonatal death (RR 7.30, 95% CI 0.39-137.54), Caesarean delivery (RR 1.39, 95% CI 0.33-5.88), spontaneous labour before 34+0 weeks (RR 1.04, 95% CI 0.75-1.45), corticosteroid use (RR 1.15, 95% CI 0.87-1.52), neonatal infection (RR 1.18, 95% CI 0.5-2.78), use of neonatal antibiotics (RR 1.02, 95% CI 0.89-1.17), neonatal intraventricular haemorrhage (RR 1.31, 95% CI 0.37-4.56), and maternal use of antibiotics after delivery (RR 1.34, 95% CI 0.55-3.3) associated with assessment by endovaginal ultrasound. There was no significant mean difference (MD) between groups in days from randomisation to birth (MD 1.9, 95% CI -11.61-15.41), days of postnatal maternal hospitalisation (MD -1, 95% CI -22.03-20.03), duration of maternal antibiotic treatment before labour (MD 21, 95% CI -67.6-109.6), birthweight (MD 31, 95% CI -636.35-698.35), days from birth to discharge home of the neonate (MD 4, 95% CI -45.82-53.82) or cord blood pH (MD -0.01, 95% CI -0.13-0.11). There was no difference in risk of Apgar scores over 5 at 5 minutes between the groups (RR 1, 95% CI 0.91-1.09). It was also inferred that there were no fetal deaths in either the treatment (endovaginal scan) group or the control (no scan) group (RR not estimable).

7.3.3.2 Amniocentesis versus no amniocentesis.

Compared with the group who were not assessed, Cotton (1984) (352) found a significantly decreased RR of neonatal apnea in the group assessed by amniocentesis (RR 0.13, 95% CI 0.02-0.92). Amniocentesis was associated with a non-significant decreased RR of maternal chorioamnionitis (RR 0.67, 95% CI 0.12-3.61), respiratory distress syndrome (RR 0.67, 95% CI 0.22-2.04), intraventricular haemorrhage (RR 0.75, 95% CI 0.31-1.8), neonatal sepsis (RR 0.33, 95% CI 0.04-2.96), cardiotocographic abnormality in

labour (RR 0.14, 95% CI 0.02-1.07), neonatal transient tachypnea (RR 0.75, 95% CI 0.31-1.8), hyperbilirubinaemia (RR 0.65, 95% CI 0.4-1.04) and metabolic problems (RR 0.38, 95% CI 0.11-1.23). However, amniocentesis was also associated with a non-significant increased RR of birth within 48 hours after rupture of membranes (RR 1.33, 95% CI 0.71-2.51), feeding difficulties (RR 2.33, 95% CI 0.69-7.88), seizures (RR 3, 95% CI 0.13-69.87) and persistent ductus arteriosus (RR 2, 95% CI 0.2-20.49). There was no difference in risk of neonatal death (RR 1, 95% CI 0.07-15) or pneumothorax (RR 1, 95% CI 0.07-15) between the groups. It was also inferred that there were no fetal deaths in either group (RR not estimable).

7.3.3.3 Daily nonstress test versus daily modified biophysical profile.

Compared with the group receiving daily full modified biophysical profiles, Lewis (1999) (353) found a significant decreased RR of maternal endometritis in the group receiving the daily nonstress test only (RR 0.25, 95% CI 0.1-0.64). There was also a non-significant decreased RR of Caesarean delivery (RR 0.64, 95% CI 0.35-1.15), maternal chorioamnionitis (RR 0.55, 95% CI 0.29-1.07), neonatal respiratory distress syndrome (RR 0.58, 95% CI 0.33-1.03), oxygen therapy after 36+0 weeks (RR 0.77, 95% CI 0.21-2.73), intraventricular haemorrhage (RR 0.72, 95% CI 0.17-3.08), necrotising entercolitis (RR 0.64, 95% CI 0.11-3.7), apgar scores less than 7 at 5 minutes (RR 0.78, 95% CI 0.35-1.77) and corticosteroid use (RR 0.88, 95% CI 0.64-1.21) in this comparison. However, there was a non-significant increased RR of retinopathy of prematurity (RR 1.34, 95% CI 0.45-4.01), delivery for maturity (RR 1.57, 95% CI 0.89-2.79), sepsis (RR 2.87, 95% CI 0.31-26.9) and presumed sepsis (RR 1.05, 95% CI 0.48-2.31). There was no significant mean difference (MD) between groups in days of neonatal antibiotic use (MD 0.36, 95% CI -0.26-0.98), days from birth to discharge home (MD -5.6, 95% CI -16.88-5.68), birthweight (MD 126.2, 95% CI -56.32-308.72), or days from randomisation to birth (MD 1, 95% CI -2.83-4.83).

7.3.3.4 Any assessment method versus no assessment method.

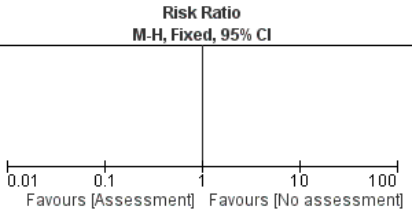
Where Carlan (1997)(351) and Cotton (1984) (352) reported the same outcomes, it was possible to conduct meta-analyses to assess the effect of any assessment method (fetal assessment using amniocentesis or endovaginal ultrasound) versus no assessment method. These meta-analyses showed no significant differences in risk of any outcome between the assessed and non-assessed groups. Compared with no assessment, fetal

assessment was associated with a non-significant increased RR of neonatal death (RR 3.07, 95% CI 0.5-18.83) and a non-significant decreased RR of maternal chorioamnionitis (RR 0.71, 95% CI 0.36-1.41), respiratory distress syndrome (RR 0.7, 95% CI 0.42-1.17) and intraventricular haemorrhage (RR 0.93, 95% CI 0.45-1.91). There were no fetal deaths in either group, so the RR is not calculable for this outcome. The two studies included in these meta-analyses (Carlan (1997)(351) and Cotton (1984) (352)) roughly agreed on the direction and degree of the RR, and statistical measures suggest low heterogeneity (I^2 was 0% for all outcomes, with P-values from Chi^2 tests for heterogeneity ranging from 0.32 to 0.93). The results are summarised in Figure 7.3.

Fetal death (antenatal)

Study or Subgroup	Assessment Events	Assessment Total	No assessment Events	No assessment Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Carlan 1997	0	45	0	47		Not estimable
Cotton 1984	0	22	0	22		Not estimable
Total (95% CI)		**67**		**69**		**Not estimable**
Total events	0		0			

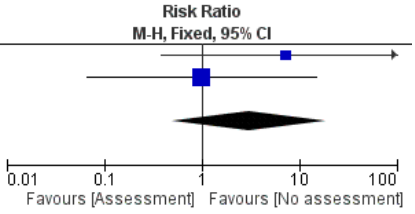
Heterogeneity: Not applicable
Test for overall effect: Not applicable



Neonatal death (in the first 28 days of life)

Study or Subgroup	Assessment Events	Assessment Total	No assessment Events	No assessment Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Carlan 1997	3	45	0	47	32.9%	7.30 [0.39, 137.54]
Cotton 1984	1	22	1	22	67.1%	1.00 [0.07, 15.00]
Total (95% CI)		**67**		**69**	**100.0%**	**3.07 [0.50, 18.83]**
Total events	4		1			

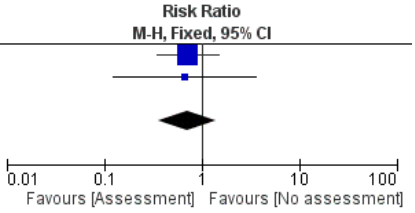
Heterogeneity: Chi² = 0.99, df = 1 (P = 0.32); I² = 0%
Test for overall effect: Z = 1.21 (P = 0.23)



Maternal chorioamnionitis

Study or Subgroup	Assessment Events	Assessment Total	No assessment Events	No assessment Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Carlan 1997	9	45	13	47	80.9%	0.72 [0.34, 1.52]
Cotton 1984	2	22	3	22	19.1%	0.67 [0.12, 3.61]
Total (95% CI)		**67**		**69**	**100.0%**	**0.71 [0.36, 1.41]**
Total events	11		16			

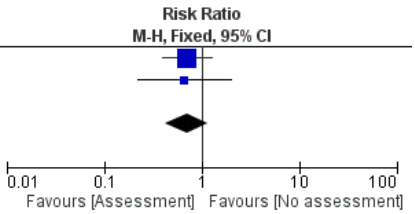
Heterogeneity: Chi² = 0.01, df = 1 (P = 0.93); I² = 0%
Test for overall effect: Z = 0.97 (P = 0.33)



Respiratory distress syndrome

Study or Subgroup	Assessment Events	Assessment Total	No assessment Events	No assessment Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Carlan 1997	13	45	19	47	75.6%	0.71 [0.40, 1.27]
Cotton 1984	4	22	6	22	24.4%	0.67 [0.22, 2.04]
Total (95% CI)		**67**		**69**	**100.0%**	**0.70 [0.42, 1.17]**
Total events	17		25			

Heterogeneity: Chi² = 0.01, df = 1 (P = 0.91); I² = 0%
Test for overall effect: Z = 1.35 (P = 0.18)



Intraventricular haemorrhage

Study or Subgroup	Assessment Events	Assessment Total	No assessment Events	No assessment Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Carlan 1997	5	45	4	47	32.8%	1.31 [0.37, 4.56]
Cotton 1984	6	22	8	22	67.2%	0.75 [0.31, 1.80]
Total (95% CI)		**67**		**69**	**100.0%**	**0.93 [0.45, 1.91]**
Total events	11		12			

Heterogeneity: Chi² = 0.51, df = 1 (P = 0.47); I² = 0%
Test for overall effect: Z = 0.19 (P = 0.85)

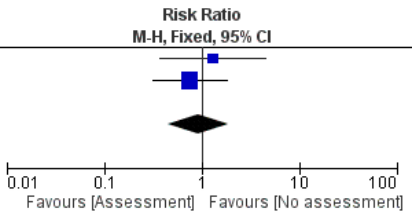


Figure 7.3. Forest plots comparing any method of assessment versus no assessment for five outcomes.

7.4 Discussion

7.4.1 Summary of main results

This review looked at fetal assessment methods for improving neonatal and maternal outcomes in PPROM. The search returned six reports describing four studies. Of the three studies that met the inclusion criteria of this review, all were unmasked single-centred randomised controlled trials and all investigated different fetal assessment methods. Carlan (1997)(351) investigated endovaginal ultrasound versus no assessment, Cotton (1984) (352) investigated amniocentesis versus no assessment, and Lewis (1999) (353) investigated daily nonstress tests versus daily modified biophysical profiling. Overall, there were few significant differences in maternal or fetal outcomes between the two groups in any of the studies. There were no significant differences in a meta-analysis of any assessment method versus no assessment. None of the included studies addressed all of the primary outcomes, although Carlan (1997)(351) and Cotton (1984) (352) found no significant difference in neonatal death between the assessed and non-assessed groups.

7.4.2 Overall completeness and applicability of evidence

There are too few trials in this area to provide enough data to draw any firm conclusions. Additionally, none of the identified studies reported all of the primary outcomes, and few of the secondary outcomes were reported. The "results" of fetal assessment (i.e. the decision to intervene to deliver the baby early) were also not reported in any study. None of the outcomes that were reported were significantly affected by any of the assessment methods, so the evidence cannot be applied in a clinical setting.

As outlined in the protocol, we planned to meta-analyse the results of trials comparing specific, individual assessments versus no assessment. However, only three trials were identified and each assessed a different method of fetal assessment, so we were only able to conduct a meta-analysis of "any assessment" versus "no assessment". The combined analysis of very different methods of fetal assessment (endovaginal ultrasound and amniocentesis) could be considered invalid.

7.4.3 Quality of the evidence

Overall, the quality of the three studies included in this review is mixed. In two (Carlan (1997)(351) and Cotton (1984) (352)) out of three studies, the method used to generate

allocation sequences was not described and in one study (Cotton (1984) (352)) the concealment of these allocations was also not described. Although the nature of the interventions prevented patient and clinician blinding, blinding of outcome assessors was also not attempted in any study. Two studies were free of selective reporting, but one study (Cotton (1984) (352)) did not report data for several outcomes mentioned in the methods section. For Cotton (1984) (352) and Lewis (1999) (353) there was an unclear risk of bias associated with the failure to perform a subgroup analysis of patients with multiple pregnancies. Attrition was not a problem in any of the studies and reasons for any patients lost to follow up were also described adequately.

The small number of studies and limited applicability of the evidence mean that findings of these studies and this review may not be generalisable to other populations.

7.4.3.1 Potential biases in the review process

The applicability of this review is limited by the small number of high quality studies identified. In an effort to reduce biases from other potential sources, study eligibility for inclusion, data extraction and risk of bias assessment were independently conducted by two review authors.

7.4.4 Agreements and disagreements with other studies or reviews

No other studies or reviews assessing fetal assessment methods for improving neonatal and maternal outcomes in PPROM were identified.

7.4.5 Conclusion

Currently, there is insufficient evidence on the benefits and harms of any fetal assessment method for improving neonatal and maternal outcomes in women with PPROM to make firm recommendations to guide clinical practice. However, this is an important clinical question and further randomised controlled trials are required. These should collect data on a wide range of neonatal and maternal outcomes, including fetal and maternal morbidity and mortality and the "results" of the fetal assessment (i.e. time from assessment to delivery and whether the delivery was medically indicated or spontaneous) and should include subgroup analyses to assess the effect of multiple pregnancies. Future systematic reviews in this area should be weary of combining analysis of disparate fetal

assessment methods, and should meta-analyse individual assessment methods versus no assessment to provide the most clinically useful evalutaion of their efficacy.

8 General discussion

8.1 Background and aims of PhD

The molecular mechanisms initiating human parturition are complex and difficult to study. Therefore, we have a limited understanding of the system, which seriously hampers attempts to predict, prevent and treat parturition complications. Advanced data analysis methods can help overcome barriers to parturition research by maximising the amount of useful data gleaned from existing data. The work for this PhD aimed to explore the use of special methods to analyse novel and published data to study the molecular mechanisms initiating human parturition and the epidemiology of parturition complications. The work has provided several original contributions. Specifically, these are:

- The largest myometrial microarray experiment to compare labouring and non-labouring samples to date. The study used 48 samples compared to the previous largest study of this kind, which used 39 samples.
    - This study provided further support for the hypothesis that labour is associated with an upregulation of inflammatory genes.
    - It also suggested that other patient characteristics such as BMI and maternal age do not play a large role in shaping the myometrial transcriptome.
- The first meta-analysis of microarray data comparing labouring and non-labouring tissue:
    - Providing further support for the hypothesis that labour is associated with an upregulation of inflammatory genes, while increasing the statistical power, reliability and generalisability of this finding.
- The first kinetic model of a signalling pathway to study parturition, and the first model of NF-kappa B activation to include reactions upstream of IKK activation.
- The first attempt to explore the network graph approach to study epidemiological data on non-communicable health outcomes, and the first use of computational epidemiology in pregnancy research.
- The first study to investigate intergenerational PPH, and the first to show that the odds of a woman experiencing PPH is not higher in women whose

mothers/grandmothers experienced PPH in their own pregnancies, compared to women whose mothers/grandmothers did not.

- The first systematic review to explore the efficacy of fetal assessment in improving maternal and neonatal outcomes following PPROM.

8.2 Summary of findings

8.2.1 Microarray analysis of the myometrial transcriptome

'Systems biology' provides a way to consolidate and extend current knowledge to characterise complex processes through describing the components of a system and their interactions at a global level. It considers biological systems from a holistic perspective and attempts to identify emergent properties that would not have been obvious using more conventional, reductionist techniques. One of the main techniques in systems biology involves generating and analysing high-throughput data, such as that generated by microarray experiments.

In Chapter 3, it was hypothesised that exploring myometrial microarray data using functional analysis and a network graph approach would reveal genes and molecular pathways associated with labour and other patient characteristics. To test this, a large microarray experiment was conducted using myometrial samples from 22 labouring and 26 non-labouring women. This study was 100% powered to detect a fold change of 1.2 with an alpha of 0.05. This means that there is a very low probability that true differences between labouring and non-labouring samples will not be discovered.1761 genes were significantly differentially expressed between the two groups. Network graph and Gene Ontology analysis showed that labour was associated with an upregulation of biological processes such as inflammation and cell movement/migration, and a downregulation of muscle-specific processes and pathways such as muscle development, smooth muscle contraction and calcium signalling. Both of these observations are in agreement with the results of previous myometrial microarray studies (115,140), and provide further evidence to support the hypothesis that parturition is associated with inflammation. However, it was also hypothesised that these observations may be in part due to a change in the composition of the myometrial tissue. It has previously been demonstrated that more leukocytes are present in the myometrium at labour (4), and this was confirmed in the samples used in this microarray through immunohistochemistry staining for the leukocyte marker CD45.

When choosing samples to include in the microarray, an effort was made to maximise the number of samples rather than to reduce inter-patient variation. Network graph and heatmap analysis suggested that other patient characteristics such as BMI, maternal age and gestational age were unlikely to have a significant impact on the results; samples clustered by labour status rather than any other characteristic. Parity appeared to have a small effect, with 421 genes being significantly differentially expressed in first deliveries (para=0) compared to subsequent deliveries (para≥1), although there were no clear differences in the biological processes identified as associated with para=0 and those identified as associated with para≥1. However, it is worth noting that "lumping" patients together to increase numbers is only useful if it increases the power of subsequent analyses. Network graphs of sample-sample relationships only show global, genome-wide differences in expression, which may mask significant differences in the expression of individual genes. Therefore interesting differences between individuals, other than labour-status, cannot be detected using this method. Network graphs built using individual gene expression profiles do allow these differences to be detected, but ignore genes that don't share a similar expression profile with any other genes. Ideally, samples should be chosen in an effort to maximise numbers, but *also* reduce inter-patient variability.

It was also hypothesised that the reliability and generalisability of the findings from this microarray analysis would be increased through meta-analysis with other, similar myometrial microarray datasets. Therefore, the first meta-analysis of microarray data to study parturition using raw, complete data was conducted. Meta-analysis methods similar to those described for more traditional meta-analyses of randomised controlled trials (as in Chapter 7) were used. The meta-analysis showed that although different studies show limited agreement on the precise myometrial gene signature associated with labour, studies do agree that labour is associated with an upregulation of inflammatory processes and a downregulation of muscle-specific processes.

Appropriate analysis of high throughput studies such as this are essential for improving our understanding of the molecular mechanisms underlying parturition, and will provide a basis for understanding differences between normal and dysfunctional or preterm labour.

8.2.2 Computer modelling of a signalling pathway relevant to preterm labour

In addition to high-throughput data generation and analysis, systems biology also involves building computational models to integrate complex interactions. Such models can be used to develop comprehensive *in silico* reproductions of "pregnant" tissues that demonstrate "emergent" properties, not obvious by conventional analysis (156).

In Chapter 4 it was hypothesised that published knowledge of a signalling pathway relevant to preterm labour could be integrated to build a computational model that successfully simulates *in vitro* behaviour. To test this, a pathway relevant to infection-induced preterm labour (LPS activation of NF-κB) was chosen to model. This pathway results in activation of NF-κB, which plays a key role in many of the inflammatory processes found to be associated with labour in Chapter 3. Detailed information about the structure and kinetics of this pathway were collated from the literature. A formalised pathway diagram and a kinetic model based on ordinary differential equations were built. The kinetic model successfully recapitulated *in silico* the published behaviour (205) of the system *in vitro* in mouse embryo fibroblasts (MEFs). The model extended previous models of LPS-induced NF-κB activation by using LPS as an input and explicitly modelling molecular interactions upstream of IKK activation, including LPS to IKK and the production and action of TNFα. This allowed closer analysis of the interactions that activate IKK and therefore affect downstream NF-kB activity.

Computational modelling is a good example of how the systems biology approach complements the traditional reductionist approach. It is essential that models are validated using wet-lab experiments to measure the activity of individual proteins over time. The model described in this thesis was built using data taken from only one paper (205) and could therefore be described as "underpowered" Although the model described here needs to be validated in uterine smooth muscle cells, it provides proof of concept that it is possible to build computational models of signalling pathways relevant to labour. When validated appropriately, such models could be used for drug testing *in silico*, providing a rapid, safe, economical and ethical strategy to identify candidate effective therapies for further testing. Thus, these models have the potential to improve our understanding of parturition and translate into improved pregnancy outcomes.

8.2.3 A NETWORK GRAPH APPROACH TO EXPLORE BIRTH RECORDS

Systems and computational approaches such as those explored in Chapters 3 and 4 have been increasingly applied to molecular research over the past decade. However, most epidemiological studies still use traditional statistical approaches, and 'computational epidemiology', which looks for patterns in unstructured sources of data, is in its infancy (170–172).

There are currently no reports of computational epidemiology being applied to pregnancy research, although it could be a useful, novel approach that may help generate interesting hypotheses about the epidemiology of parturition complications. These hypotheses could then be tested using traditional epidemiological approaches that analyse associations between exposures and outcomes, providing a good example of how the systems and reductionist approaches are complementary.

The network graph approach used to analyse gene expression data in Chapter 3 is an interesting, novel method that works well with large datasets and is not restricted to analysing molecular data. Therefore, in Chapter 5 it was hypothesised that a network graph approach could be used to visualise and explore epidemiological data relating to obstetric complications and birth outcomes.

Specifically, this work described the use of 3D, interactive, unbiased network graphs to visually explore the Walker Cohort – a database of nearly 50,000 birth records collected from 1952 to 1966 in Dundee. BioLayout *Express*^{3D} was used to build network graphs in which each 'node' represents either a clinical characteristic or a delivery episode, connected by weighted 'edges' based on Pearson correlation coefficients. The large size of the Walker cohort means the analyses are well powered to detect true relationships in the data.

By building graphs in which nodes to represent obstetric characteristics, relationships can be visualised and traced to a wider network to generate hypotheses about covariates. This approach copes well with missing data and allows discovery of 'emergent' relationships that could not be identified using a purely statistical approach. Graphs in which nodes represent deliveries are larger and can be explored to reveal factors that often occur together in different deliveries. This is useful for highlighting pregnancy risk factors for parturition complications.

Although the network graph approach did not highlight any novel associations, it provided an effective way to quickly explore relationships within the data and was particularly useful in identifying non-informative variables.

8.2.4 Record-linking to study intergenerational transmission of PPH

Records from clinical databases or epidemiological studies can often be linked to other datasets to increase the breadth of data available for each patient. Pooling similar datasets can also increase the power of analyses by increasing the number of participants. However, datasets must be directly comparable (i.e. containing similar variables and participants from similar populations). If datasets are not comparable, pooling data can in fact reduce the power of the study. The historical nature of the Walker cohort (analysed in Chapter 5) allows its records to be linked with a large number of current health-outcome databases, and allows linkage of records over generations. Therefore, in Chapter 6 it was hypothesised that data from Walker and another similar database of Scottish birth records (SMR02) could be record-linked to assess if the risk of postpartum haemorrhage (PPH) is higher in women whose mothers had PPH in their pregnancies (Chapter 6). PPH was chosen because previous PPH is a significant risk factor for subsequent PPH,(325–328) which raises the possibility that this predisposition could be heritable. No previous studies had addressed this.

The analyses did not support a large increased risk of PPH for women whose mothers/grandmothers had PPH. These analyses were underpowered to detect odds ratios above 1.3, so results should be interpreted cautiously. Caesarean delivery, genital trauma or episiotomy, high birth weight and nulliparity were identified as risk factors for PPH, which confirms the results of previous studies.(64,327,329–331) This work suggests that pregnant women whose mothers had PPH can be reassured that they are unlikely to be at any significantly increased risk, compared to those whose mothers did not have PPH.

8.2.5 Systematic review and meta-analysis to evaluate the efficacy of fetal assessment in improving outcomes following preterm prelabour rupture of membranes

In addition to studying patterns in disease as discussed in Chapters 5 and 6, epidemiology is also concerned with how clinical interventions affect the health of a population. The

efficacy of medical interventions is assessed through randomised controlled trials (RCTs), but the reliability of the evidence is greatly increased through systematic review and meta-analysis of all available data from comparable RCTs. Systematic reviews are most useful when there are enough comparable RCTs to allow meta-analysis on a range of outcomes. However, it is worth noting that the results of RCTs may be different in different populations, and meta-analysis (which treats all populations as alike) may mask such differences. Therefore, it is necessary for systematic reviews to discuss the design and results of individual RCTs as well as the results of any meta-analyses.

Complications at parturition are often difficult to predict, prevent and manage due to an incomplete understanding of their aetiology. Furthermore, the efficacy of current interventions to improve outcomes following parturition complications is often unclear. In Chapter 7, a systematic review was conducted to study fetal assessment methods for improving neonatal and maternal outcomes following a major obstetric complication, preterm prelabour rupture of membranes (PPROM). After PPROM, accurate methods for assessing fetal wellbeing and prognosis are needed to aid obstetricians' decisions in planning the time and mode of delivery. However, the value of the various methods of fetal assessment in ultimately improving neonatal and maternal outcomes had yet to be established.

The review was conducted in collaboration with the Cochrane group, who carried out the literature search. This returned six reports describing four studies. Of the three studies that met the inclusion criteria of the review, all investigated different fetal assessment methods: endovaginal ultrasound versus no assessment (351), amniocentesis versus no assessment, (351) and daily nonstress tests versus daily modified biophysical profiling. (352) Overall, there were few significant differences in maternal or fetal outcomes between the two groups in any of the studies, and there were no significant differences in a meta-analysis of any assessment method versus no assessment.

Therefore, currently there is insufficient evidence on the benefits and harms of fetal assessment methods for improving neonatal and maternal outcomes in women with PPROM to make firm recommendations to guide clinical practice. However, this is an important clinical question, so further randomised controlled trials are required.

8.3 Suggestions for future work

Several of the results in this thesis would benefit from further exploration.

The expression of genes and pathways of interest highlighted by the microarray experiment should be investigated at the protein level. Additionally, although more difficult to obtain, it would be interesting to carry out a similarly large microarray experiment using tissue from the uterine fundus, which is more contractile. It would also be interesting to study the myometrial transcriptome at different stages of labour. Cervical dilation could be used to estimate the stage of labour, however this estimate would be very rough and the experiment would require large numbers of samples at each stage in order to minimise the effect of noise in the data and maximise the chances of seeing any real effect. Our finding that labouring samples show more variation than non-labouring samples suggests that the myometrial gene expression signature does vary with the stage of labour. Future microarray experiments should use a sufficient number of samples in each group (we suggest 20 or more) in order to ensure analyses are well-powered. Finally, the quality of input material should be assessed in all molecular investigations, because it has a large effect on the results of downstream analyses. Details of how input quality is assessed should also be reported in papers describing such studies,

The LPS-NF-κB model could be extended to include more LPS targets and/or model inputs, which would allow the model to incorporate pathway crosstalk and more possible targets for pharmaceutical intervention. However, this would take a long time and the size of the model might make it difficult to run simulations. Different modelling techniques (such as Boolean modelling that uses qualitative rather than quantitative data) could be explored to produce a model that is less reliant on specific kinetic information and models molecular interactions in the uterus at a broader level. Either way, the model must be validated using wet-lab *in vitro* experiments in myometrial smooth muscle cells before it can be used to make predictions about the pathway in human parturition. Care should be taken to ensure that wet-lab experiments are designed to provide useful data at appropriate timepoints and that analyses are well-powered. The reliability and usefulness of the model hinges on the quality of this input data.

The network graph approach to explore epidemiological data could be applied to other datasets. The relationships between the factors collected for the Walker cohort are already

understood, we believe that the network graph approach could generate new insights if applied to a dataset in which less is known about the associations between its variables (for example, a dataset containing data about the intergenerational transmission of pregnancy complications).

As with all the analyses described in this thesis, care should be taken to ensure the data used to build network graphs is high quality. Datasets may need to be "cleaned" to remove erroneous values, and less useful/meaningless variables should be omitted. A range of MCL algorithm parameter values should also be explored. However, the size of the dataset (around 50,000 cases minimum) is the most important factor in producing useful network graphs.

Network graphs are useful for exploring associations in large datasets and are not intended to make any suggestions about causality or the "direction" of those associations. Directed Acyclic Graphs (DAGs) may be a more appropriate technique for assessing causality. DAGs are increasingly being used in epidemiological studies for this purpose.

The intergenerational transmission of PPH could be assessed using other large datasets. The intergenerational transmission of other pregnancy-related conditions could also be assessed by linking SMR02 to Walker. Again, the quality of data is a key factor in the reliability and generalisability of results. Larger datasets collected from several different centres are most likely to require some level of "cleaning", however large databases (>50,000 participants) are also most likely to provide sufficient power for analyses.

Finally, as discussed in Chapter 7, there is a need for more high quality randomised controlled trials to assess the effectiveness of fetal assessment methods in improving neonatal and maternal outcomes following PPROM. High quality, well-powered RCTs are essential for carrying out meta-analyses that yield useful, generalisable results.

8.4 General conclusion

Taken collectively, the results in this thesis suggest that in-depth analysis of molecular and epidemiological data related to parturition and associated complications is helpful in maximising the amount of useful information that can be gleaned. This provides useful insights into not only the molecular mechanisms initiating parturition, but also the aetiology and clinical management of parturition complications.

9 REFERENCES

1. Ledingham MA, Thomson AJ, Jordan F, Young A, Crawford M, Norman JE. Cell adhesion molecule expression in the cervix and myometrium during pregnancy and parturition. Obstetrics and gynecology [Internet]. 2001 Mar [cited 2013 Aug 30];97(2):235–42.

2. Winkler M, Fischer DC, Ruck P, Horny HP, Kemp B, Rath W. Cytokine concentrations and expression of adhesion molecules in the lower uterine segment during parturition at term: relation to cervical dilatation and duration of labor. Zeitschrift für Geburtshilfe und Neonatologie [Internet]. 1998 [cited 2013 Aug 30];202(4):172–5.

3. Osman I, Young A, Ledingham MA, Thomson AJ, Jordan F, Greer IA, et al. Leukocyte density and pro-inflammatory cytokine expression in human fetal membranes, decidua, cervix and myometrium before and during labour at term. Molecular human reproduction [Internet]. 2003 Jan 1 [cited 2013 Aug 30];9(1):41–5.

4. Thomson AJ, Telfer JF, Young A, Campbell S, Stewart CJ, Cameron IT, et al. Leukocytes infiltrate the myometrium during human parturition: further evidence that labour is an inflammatory process. Human reproduction (Oxford, England) [Internet]. 1999 Jan 1 [cited 2013 Aug 30];14(1):229–36.

5. Young A, Thomson AJ, Ledingham M, Jordan F, Greer IA, Norman JE. Immunolocalization of proinflammatory cytokines in myometrium, cervix, and fetal membranes during human parturition at term. Biology of reproduction [Internet]. 2002 Mar 1 [cited 2013 Aug 30];66(2):445–9.

6. Elliott CL, Slater DM, Dennes W, Poston L, Bennett PR. Interleukin 8 expression in human myometrium: changes in relation to labor onset and with gestational age. American journal of reproductive immunology [Internet]. 2000 May [cited 2013 Aug 7];43(5):272–7.

7. Sennström MB, Ekman G, Westergren-Thorsson G, Malmström A, Byström B, Endrésen U, et al. Human cervical ripening, an inflammatory process mediated by cytokines. Molecular human reproduction [Internet]. 2000 Apr [cited 2013 Aug 30];6(4):375–81.

8. Tribe RM, Moriarty P, Dalrymple A, Hassoni AA, Poston L. Interleukin-1beta induces calcium transients and enhances basal and store operated calcium entry in human myometrial smooth muscle. Biology of reproduction. 2003 May;68(5):1842–9.

9. Belt AR, Baldassare JJ, Molnár M, Romero R, Hertelendy F. The nuclear transcription factor NF-kappaB mediates interleukin-1beta-induced expression of cyclooxygenase-2 in human myometrial cells. American journal of obstetrics and gynecology [Internet]. 1999 Aug [cited 2013 Aug 30];181(2):359–66.

10. Hertelendy F, Romero R, Molnár M, Todd H, Baldassare JJ. Cytokine-initiated signal transduction in human myometrial cells. American journal of reproductive immunology [Internet]. 1993 [cited 2013 Aug 30];30(2-3):49–57.

11. Rauk PN, Chiao JP. Interleukin-1 stimulates human uterine prostaglandin production through induction of cyclooxygenase-2 expression. American journal of reproductive immunology [Internet]. 2000 Mar [cited 2013 Aug 30];43(3):152–9.

12. Johnston TA, Greer IA, Kelly RW, Calder AA. Plasma prostaglandin metabolite concentrations in normal and dysfunctional labour. British journal of obstetrics and gynaecology. 1993 May;100(5):483–8.

13. Molnár M, Romero R, Hertelendy F. Interleukin-1 and tumor necrosis factor stimulate arachidonic acid release and phospholipid metabolism in human myometrial cells. American journal of obstetrics and gynecology. 1993 Oct;169(4):825–9.

14. Johnson RF, Mitchell CM, Clifton V, Zakar T. Regulation of 15-hydroxyprostaglandin dehydrogenase (PGDH) gene activity, messenger ribonucleic acid processing, and protein abundance in the human chorion in late gestation and labor. The Journal of clinical endocrinology and metabolism. 2004 Nov;89(11):5639–48.

15. Sanborn BM, Ku C-Y, Shlykov S, Babich L. Molecular signaling through G-protein-coupled receptors and the control of intracellular calcium in myometrium. Journal of the Society for Gynecologic Investigation. 2005 Oct;12(7):479–87.

16. Asbóth G, Phaneuf S, López Bernal AL. Prostaglandin E receptors in myometrial cells. Acta physiologica Hungarica [Internet]. 1997 [cited 2013 Aug 30];85(1):39–50.

17. López Bernal A. Mechanisms of labour--biochemical aspects. BJOG : an international journal of obstetrics and gynaecology. 2003 Apr;110 Suppl 39–45.

18. Ledingham MA, Denison FC, Riley SC, Norman JE. Matrix metalloproteinases-2 and -9 and their inhibitors are produced by the human uterine cervix but their secretion is not regulated by nitric oxide donors. Human reproduction (Oxford, England). 1999 Aug;14(8):2089–96.

19. Xu P, Alfaidy N, Challis JRG. Expression of matrix metalloproteinase (MMP)-2 and MMP-9 in human placenta and fetal membranes in relation to preterm and term labor. The Journal of clinical endocrinology and metabolism. 2002 Mar;87(3):1353–61.

20. Smith R. Mechanisms of disease: Parturition. The New England Journal of Medicine2. 2007;356:271–83.

21. Boroditsky RS, Reyes FI, Winter JS, Faiman C. Maternal serum estrogen and progesterone concentrations preceding normal labor. Obstetrics and gynecology [Internet]. 1978 Jun [cited 2013 Aug 30];51(6):686–91.

22. Mathur RS, Landgrebe S, Williamson HO. Progesterone, 17-hydroxyprogesterone, estradiol, and estriol in late pregnancy and labor. American journal of obstetrics and gynecology. 1980 Jan;136(1):25–7.

23. Csapo AI, Knobil E, van der Molen HJ, Wiest WG. Peripheral plasma progesterone levels during human pregnancy and labor. American journal of obstetrics and gynecology [Internet]. 1971 Jul 1 [cited 2013 Aug 30];110(5):630–2.

24. Csapo AI, Pinto-Dantas CA. The effect of progesterone on the human uterus. Proceedings of the National Academy of Sciences of the United States of America [Internet]. 1965 Oct [cited 2013 Aug 30];54(4):1069–76.

25. Csapo AI. Progesterone block. The American journal of anatomy [Internet]. 1956 Mar [cited 2013 Aug 30];98(2):273–91.

26. Mesiano S, Chan E-C, Fitter JT, Kwek K, Yeo G, Smith R. Progesterone withdrawal and estrogen activation in human parturition are coordinated by progesterone receptor A expression in the myometrium. The Journal of clinical endocrinology and metabolism. 2002 Jun;87(6):2924–30.

27. Mitchell BF, Wong S. Changes in 17 beta,20 alpha-hydroxysteroid dehydrogenase activity supporting an increase in the estrogen/progesterone ratio of human fetal membranes at parturition. American journal of obstetrics and gynecology. 1993 May;168(5):1377–85.

28. Condon JC, Jeyasuria P, Faust JM, Wilson JW, Mendelson CR. A decline in the levels of progesterone receptor coactivators in the pregnant uterus at term may antagonize progesterone receptor function and contribute to the initiation of parturition. Proceedings of the National Academy of Sciences of the United States of America [Internet]. 2003 Aug 5 [cited 2013 Aug 30];100(16):9518–23.

29. Madsen G, Zakar T, Ku CY, Sanborn BM, Smith R, Mesiano S. Prostaglandins differentially modulate progesterone receptor-A and -B expression in human myometrial cells: evidence for prostaglandin-induced functional progesterone withdrawal. The Journal of clinical endocrinology and metabolism. 2004 Feb;89(2):1010–3.

30. Smith R, Paul J, Maiti K, Tolosa J, Madsen G. Recent advances in understanding the endocrinology of human birth. Trends in endocrinology and metabolism: TEM. Elsevier Ltd; 2012 Jun;1–8.

31. Pinto RM, Lerner U, Glauberman M, Pontelli H. Influence of estradiol-17-beta upon the oxytocic action of oxytocin in the pregnant human uterus. American journal of obstetrics and gynecology. 1966 Nov;96(6):857–62.

32. Nissenson R. Opposing Effects of Estradiol and Progesterone on Oxytocin Receptors in Rabbit Uterus. Proceedings of the National Academy of Sciences. 1978 Apr;75(4):2044–8.

33. Fuchs AR, Fuchs F, Husslein P, Soloff MS. Oxytocin receptors in the human uterus during pregnancy and parturition. American journal of obstetrics and gynecology. 1984 Nov;150(6):734–41.

34. Orsino A, Taylor C V, Lye SJ. Connexin-26 and connexin-43 are differentially expressed and regulated in the rat myometrium throughout late pregnancy and with the onset of labor. Endocrinology. 1996 May;137(5):1545–53.

35. Petrocelli T, Lye SJ. Regulation of transcripts encoding the myometrial gap junction protein, connexin-43, by estrogen and progesterone. Endocrinology. 1993 Jul;133(1):284–90.

36. Welsh T, Johnson M, Yi L, Tan H, Rahman R, Merlino A, et al. Estrogen receptor (ER) expression and function in the pregnant human myometrium: estradiol via ERα activates ERK1/2 signaling in term myometrium. The Journal of endocrinology. 2012 Feb;212(2):227–38.

37. Tulchinsky D, Hobel CJ, Yeager E, Marshall JR. Plasma estrone, estradiol, estriol, progesterone, and 17-hydroxyprogesterone in human pregnancy. I. Normal pregnancy. American journal of obstetrics and gynecology. 1972 Apr;112(8):1095–100.

38. Turnbull AC, Patten PT, Flint AP, Keirse MJ, Jeremy JY, Anderson AB. Significant fall in progesterone and rise in oestradiol levels in human peripheral plasma before onset of labour. Lancet. 1974 Jan;1(7848):101–3.

39. Makrigiannakis a, Semmler M, Briese V, Eckerle H, Minas V, Mylonas I, et al. Maternal serum corticotropin-releasing hormone and ACTH levels as predictive markers of premature labor. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 2007 May;97(2):115–9.

40. Campbell EA, Linton EA, Wolfe CD, Scraggs PR, Jones MT, Lowry PJ. Plasma corticotropin-releasing hormone concentrations during pregnancy and parturition. The Journal of clinical endocrinology and metabolism [Internet]. 1987 May [cited 2013 Aug 30];64(5):1054–9.

41. Kalantaridou SN, Makrigiannakis A, Mastorakos G, Chrousos GP. Roles of reproductive corticotropin-releasing hormone. Annals of the New York Academy of Sciences. 2003 Nov;997:129–35.

42. Linton EA, Perkins A V, Woods RJ, Eben F, Wolfe CD, Behan DP, et al. Corticotropin releasing hormone-binding protein (CRH-BP): plasma levels decrease during the third trimester of normal human pregnancy. The Journal of clinical endocrinology and metabolism. 1993 Jan;76(1):260–2.

43. McLean M, Bisits A, Davies J, Woods R, Lowry P, Smith R. A placental clock controlling the length of human pregnancy. Nature medicine. 1995 May;1(5):460–3.

44. Alvi S, Brown N, Bennett P. Corticotrophin-releasing hormone and platelet-activating factor induce transcription of the type-2 cyclo-oxygenase gene in human fetal membranes. Molecular human … [Internet]. 1999 May [cited 2013 Aug 30];5(5):476–80.

45. Bugg GJ, Siddiqui F, Thornton JG. Oxytocin versus no treatment or delayed treatment for slow progress in the first stage of spontaneous labour. The Cochrane database of systematic reviews [Internet]. 2011 Jan [cited 2013 Aug 30];(7):CD007123.

46. Papatsonis D, Flenady V, Cole S, Liley H. Oxytocin receptor antagonists for inhibiting preterm labour. The Cochrane database of systematic reviews [Internet]. 2005 Jan [cited 2013 Aug 30];(3):CD004452.

47. Thornton S, Davison JM, Baylis PH. Plasma oxytocin during the first and second stages of spontaneous human labour. Acta endocrinologica. 1992 May;126(5):425–9.

48. Blanks AM, Vatish M, Allen MJ, Ladds G, de Wit NCJ, Slater DM, et al. Paracrine oxytocin and estradiol demonstrate a spatial increase in human intrauterine tissues with labor. The Journal of clinical endocrinology and metabolism [Internet]. 2003 Jul [cited 2013 Aug 30];88(7):3392–400.

49. Wilson T, Liggins GC, Whittaker DJ. Oxytocin stimulates the release of arachidonic acid and prostaglandin F2 alpha from human decidual cells. Prostaglandins. 1988 May;35(5):771–80.

50. Fuchs AR, Husslein P, Fuchs F. Oxytocin and the initiation of human parturition. II. Stimulation of prostaglandin production in human decidua by oxytocin. American journal of obstetrics and gynecology. 1981 Nov;141(6):694–7.

51. Pasetto N, Zicari A, Piccione E, Lenti L, Pontieri G, Ticconi C. Influence of labor and oxytocin on in vitro leukotriene release by human fetal membranes and uterine decidua at term gestation. American journal of obstetrics and gynecology. 1992 May;166(5):1500–6.

52. Terzidou V, Blanks AM, Kim SH, Thornton S, Bennett PR. Labor and inflammation increase the expression of oxytocin receptor in human amnion. Biology of reproduction. 2011 Mar;84(3):546–52.

53. Blanks AM, Thornton S. The role of oxytocin in parturition. BJOG : an international journal of obstetrics and gynaecology [Internet]. 2003 Apr [cited 2013 Aug 30];110 Suppl 46–51.

54. Mitchell BF, Taggart MJ. Are animal models relevant to key aspects of human parturition? American journal of physiology. Regulatory, integrative and comparative physiology. 2009 Sep;297(3):R525–45.

55. Bukowski R, Hankins GD V, Saade GR, Anderson GD, Thornton S. Labor-associated gene expression in the human uterine fundus, lower segment, and cervix. PLoS medicine [Internet]. 2006 Jun [cited 2013 Aug 30];3(6):e169.

56. Sparey C, Robson SC, Bailey J, Lyall F, Europe-Finner GN. The differential expression of myometrial connexin-43, cyclooxygenase-1 and -2, and Gs alpha proteins in the upper and lower segments of the human uterus during pregnancy and labor. The Journal of clinical endocrinology and metabolism. 1999 May;84(5):1705–10.

57. Liggins GC, Fairclough RJ, Grieves SA, Forster CS, Knox BS. Parturition in the sheep. Ciba Foundation symposium. 1977 Jan;(47):5–30.

58. Ratajczak CK, Fay JC, Muglia LJ. Preventing preterm birth: the past limitations and new potential of animal models. Disease models & mechanisms. 2010 Jan;3(7-8):407–14.

59. Bethin KE, Nagai Y, Sladek R, Asada M, Sadovsky Y, Hudson TJ, et al. Microarray analysis of uterine gene expression in mouse and human pregnancy. Molecular endocrinology (Baltimore, Md.) [Internet]. 2003 Aug [cited 2013 Aug 30];17(8):1454–69.

60. National Health Service. Hospital Episode Statistics: NHS Maternity Statistics, 2011-12 [Internet]. The Health and Social Care Information Centre. 2012.

61. Maxwell GL. Preterm premature rupture of membranes. Obstetrical & gynecological survey. 1993 Aug;48(8):576–83.

62. Algovik M, Nilsson E, Cnattingius S, Lichtenstein P, Nordenskjöld A, Westgren M. Genetic influence on dystocia. Acta obstetricia et gynecologica Scandinavica [Internet]. 2004 Sep [cited 2013 Aug 30];83(9):832–7.

63. Calvert C, Thomas SL, Ronsmans C, Wagner KS, Adler AJ, Filippi V. Identifying regional variation in the prevalence of postpartum haemorrhage: a systematic review and meta-analysis. PloS one [Internet]. 2012 Jan [cited 2013 Aug 30];7(7):e41114.

64. Bateman BT, Berman MF, Riley LE, Leffert LR. The epidemiology of postpartum hemorrhage in a large, nationwide sample of deliveries. Anesthesia and analgesia [Internet]. 2010 May 1 [cited 2013 Aug 30];110(5):1368–73.

65. Cantwell R, Clutton-Brock T, Cooper G, Dawson A, Drife J, Garrod D, et al. Saving Mothers' Lives: Reviewing maternal deaths to make motherhood safer: 2006-2008. The Eighth Report of the Confidential Enquiries into Maternal Deaths in the United Kingdom. BJOG : an international journal of obstetrics and gynaecology [Internet]. 2011 Mar [cited 2013 Aug 30];118 Suppl 1–203.

66. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet [Internet]. 2008 Jan 5 [cited 2013 Aug 20];371(9606):75–84.

67. Berkowitz GS, Papiernik E. Epidemiology of preterm birth. Epidemiologic reviews [Internet]. 1993 Jan [cited 2013 Aug 30];15(2):414–43.

68. Murphy DJ. Epidemiology and environmental factors in preterm labour. Best practice & research. Clinical obstetrics & gynaecology. 2007 Oct;21(5):773–89.

69. Honest H, Bachmann LM, Sundaram R, Gupta JK, Kleijnen J, Khan KS. The accuracy of risk scores in predicting preterm birth--a systematic review. Journal of obstetrics and gynaecology [Internet]. 2004 Jun [cited 2013 Aug 30];24(4):343–59.

70. Honest H, Forbes C a, Durée KH, Norman G, Duffy SB, Tsourapas A, et al. Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling. Health technology assessment (Winchester, England) [Internet]. 2009 Sep [cited 2013 Aug 30];13(43):1–627.

71. Duley L, Bennett PR. Tocolysis for women in preterm labour. Royal College of Obstetricians and Gynaecologists, editor. RCOG Green-top guidelines no. 1b. RCOG Press; 2011.

72. Gyetvai K, Hannah ME, Hodnett ED, Ohlsson A. Tocolytics for preterm labor: a systematic review. Obstetrics and gynecology. 1999 Nov;94(5 Pt 2):869–77.

73. Haas DM, Imperiale TF, Kirkpatrick PR, Klein RW, Zollinger TW, Golichowski AM. Tocolytic therapy: a meta-analysis and decision analysis. Obstetrics and gynecology [Internet]. 2009 Mar [cited 2013 Aug 30];113(3):585–94.

74. Savitz DA, Blackmore CA, Thorp JM. Epidemiologic characteristics of preterm delivery: etiologic heterogeneity. American journal of obstetrics and gynecology. 1991 Feb;164(2):467–71.

75. Hadley CB, Main DM, Gabbe SG. Risk factors for preterm premature rupture of the fetal membranes. American journal of perinatology [Internet]. 1990 Oct [cited 2013 Aug 30];7(4):374–9.

76. American College of Obstetricians and Gynecologists. ACOG practice bulletin. Premature rupture of membranes. Clinical management guidelines for obstetrician-gynecologists. Number 1, June 1998. American College of Obstetricians and Gynecologists. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 1998 Oct;63(1):75–84.

77. Mercer BM. Preterm premature rupture of the membranes. Obstetrics and gynecology. 2003 Jan;101(1):178–93.

78. Parry S, Strauss JF. Premature rupture of the fetal membranes. The New England journal of medicine. 1998 Mar;338(10):663–70.

79. Simhan HN, Canavan TP. Preterm premature rupture of membranes: diagnosis, evaluation and management strategies. BJOG : an international journal of obstetrics and gynaecology. 2005 Mar;112 Suppl 32–7.

80. Mercer BM, Goldenberg RL, Meis PJ, Moawad AH, Shellhaas C, Das A, et al. The Preterm Prediction Study: prediction of preterm premature rupture of membranes through clinical findings and ancillary testing. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. American journal of obstetrics and gynecology. 2000 Sep;183(3):738–45.

81. Odibo AO, Talucci M, Berghella V. Prediction of preterm premature rupture of membranes by transvaginal ultrasound features and risk factors in a high-risk population. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2002 Sep;20(3):245–51.

82. Hofmeyr GJ, Essilfie-Appiah G, Lawrie TA. Amnioinfusion for preterm premature rupture of membranes. The Cochrane database of systematic reviews [Internet]. 2011 Jan [cited 2013 Aug 30];(12):CD000942.

83. Mackeen AD, Seibel-Seamon J, Grimes-Dennis J, Baxter JK, Berghella V. Tocolytics for preterm premature rupture of membranes. The Cochrane database of systematic reviews [Internet]. 2011 Jan [cited 2013 Aug 30];(10):CD007062.

84. Kenyon S, Boulvain M, Neilson JP. Antibiotics for preterm rupture of membranes. The Cochrane database of systematic reviews [Internet]. 2010 Jan [cited 2013 Aug 30];(8):CD001058.

85. Nesheim BI. Duration of labor. An analysis of influencing factors. Acta obstetricia et gynecologica Scandinavica [Internet]. 1988 Jan [cited 2013 Aug 30];67(2):121–4.

86. Garfield RE. Cellular and molecular bases for dystocia. Clinical obstetrics and gynecology [Internet]. 1987 Mar [cited 2013 Aug 30];30(1):3–18.

87. Wray S. Uterine contraction and physiological mechanisms of modulation. The American journal of physiology. 1993 Jan;264(1 Pt 1):C1–18.

88. Berg-Lekås ML, Högberg U, Winkvist A. Familial occurrence of dystocia. American journal of obstetrics and gynecology [Internet]. 1998 Jul [cited 2013 Aug 30];179(1):117–21.

89. Sandström A, Cnattingius S, Wikström AK, Stephansson O. Labour dystocia--risk of recurrence and instrumental delivery in following labour--a population-based cohort study. BJOG : an international journal of obstetrics and gynaecology. 2012 Dec;119(13):1648–56.

90. Treacy A, Robson M, O'Herlihy C. Dystocia increases with advancing maternal age. American journal of obstetrics and gynecology. 2006 Sep;195(3):760–3.

91. Vahratian A, Zhang J, Troendle JF, Savitz DA, Siega-Riz AM. Maternal prepregnancy overweight and obesity and the pattern of labor progression in term nulliparous women. Obstetrics and gynecology. 2004 Nov;104(5 Pt 1):943–51.

92. Zhu B-P, Grigorescu V, Le T, Lin M, Copeland G, Barone M, et al. Labor dystocia and its association with interpregnancy interval. American journal of obstetrics and gynecology. 2006 Jul;195(1):121–8.

93. Kjaergaard H, Dykes AK, Ottesen B, Olsen J. Risk indicators for dystocia in low-risk nulliparous women: a study on lifestyle and anthropometrical factors. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology [Internet]. Informa UK Ltd London, UK; 2010 Jan [cited 2013 Aug 30];30(1):25–9.

94. Feinstein U, Sheiner E, Levy A, Hallak M, Mazor M. Risk factors for arrest of descent during the second stage of labor. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics [Internet]. 2002 Apr [cited 2013 Aug 30];77(1):7–14.

95. Sheiner E, Levy A, Feinstein U, Hershkovitz R, Hallak M, Mazor M. Obstetric risk factors for failure to progress in the first versus the second stage of labor. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians. 2002 Jun;11(6):409–13.

96. Kjaergaard H, Olsen J, Ottesen B, Nyberg P, Dykes A-K. Obstetric risk indicators for labour dystocia in nulliparous women: a multi-centre cohort study. BMC pregnancy and childbirth. 2008 Jan;8(1):45.

97. Mori R, Tokumasu H, Pledge D, Kenyon S. High dose versus low dose oxytocin for augmentation of delayed labour. The Cochrane database of systematic reviews [Internet]. 2011 Jan [cited 2013 Aug 30];(10):CD007201.

98. Wei S, Wo BL, Qi H-P, Xu H, Luo Z-C, Roy C, et al. Early amniotomy and early oxytocin for prevention of, or therapy for, delay in first stage spontaneous labour compared with routine care. The Cochrane database of systematic reviews [Internet]. 2012 Jan [cited 2013 Aug 30];9:CD006794.

99. Macara LM, Murphy KW. The contribution of dystocia to the cesarean section rate. American journal of obstetrics and gynecology. 1994 Jul;171(1):71–7.

100. Knight M, Callaghan WM, Berg C, Alexander S, Bouvier-Colle M-H, Ford JB, et al. Trends in postpartum hemorrhage in high resource countries: a review and recommendations from the International Postpartum Hemorrhage Collaborative Group. BMC pregnancy and childbirth. 2009 Jan;9:55.

101. Carroli G, Cuesta C, Abalos E, Gulmezoglu a M. Epidemiology of postpartum haemorrhage: a systematic review. Best practice & research. Clinical obstetrics &

gynaecology [Internet]. Elsevier Ltd; 2008 Dec [cited 2013 Aug 30];22(6):999–1012.

102. Tsu VD. Antenatal screening: its use in assessing obstetric risk factors in Zimbabwe. Journal of epidemiology and community health. 1994 Jun;48(3):297–305.

103. Prata N, Hamza S, Bell S, Karasek D, Vahidnia F, Holston M. Inability to predict postpartum hemorrhage: insights from Egyptian intervention data. BMC pregnancy and childbirth. BioMed Central Ltd; 2011 Jan;11(1):97.

104. Mousa HA, Alfirevic Z. Treatment for primary postpartum haemorrhage. The Cochrane database of systematic reviews [Internet]. 2007 Jan [cited 2013 Aug 30];(1):CD003249.

105. Alexander J, Thomas P, Sanghera J. Treatments for secondary postpartum haemorrhage. The Cochrane database of systematic reviews [Internet]. 2002 Jan [cited 2013 Aug 30];(1):CD002867.

106. Begley CM, Gyte GML, Devane D, McGuire W, Weeks A. Active versus expectant management for women in the third stage of labour. The Cochrane database of systematic reviews [Internet]. 2011 Jan [cited 2013 Aug 30];(11):CD007412.

107. Hofmeyr GJ, Ferreira S, Nikodem VC, Mangesi L, Singata M, Jafta Z, et al. Misoprostol for treating postpartum haemorrhage: a randomized controlled trial [ISRCTN72263357]. BMC pregnancy and childbirth [Internet]. 2004 Aug 6 [cited 2013 Aug 30];4(1):16.

108. Walraven G, Dampha Y, Bittaye B, Sowe M, Hofmeyr J. Misoprostol in the treatment of postpartum haemorrhage in addition to routine management: a placebo randomised controlled trial. BJOG : an international journal of obstetrics and gynaecology. 2004 Sep;111(9):1014–7.

109. Anderson G, Sprott H, Olsen B. Opinion: Publish Negative Results | The Scientist Magazine® [Internet]. The Scientist Magazine. 2013.

110. Gravett MG, Rubens CE, Nunes TM. Global report on preterm birth and stillbirth (2 of 7): discovery science. BMC pregnancy and childbirth [Internet]. 2010 Jan [cited 2013 Aug 30];10 Suppl 1(Suppl 1):S2.

111. Romero R, Kuivaniemi H, Tromp G. Functional genomics and proteomics in term and preterm parturition. The Journal of clinical endocrinology and metabolism. 2002 Jun;87(6):2431–4.

112. Aguan K, Carvajal J. Application of a functional genomics approach to identify differentially expressed genes in human myometrium during pregnancy and labour. Molecular human … [Internet]. 2000 Dec [cited 2013 Aug 30];6(12):1141–5.

113. Breuiller-Fouche M, Charpigny G, Germain G. Functional genomics of the pregnant uterus: from expectations to reality, a compilation of studies in the myometrium.

BMC pregnancy and childbirth [Internet]. 2007 Jan [cited 2013 Aug 30];7 Suppl 1:S4.

114. Charpigny G, Leroy M-J, Breuiller-Fouché M, Tanfin Z, Mhaouty-Kodja S, Robin P, et al. A functional genomic study to identify differential gene expression in the preterm and term human myometrium. Biology of reproduction [Internet]. 2003 Jun [cited 2013 Aug 30];68(6):2289–96.

115. Weiner CPP, Mason CWW, Dong Y, Buhimschi I a A, Swaan PWW, Buhimschi CSS. Human effector/initiator gene sets that regulate myometrial contractility during term and preterm labor. American journal of obstetrics and gynecology [Internet]. Elsevier Inc.; 2010 May [cited 2014 Jan 10];202(5):474.e1–20.

116. Buhimschi CS, Bhandari V, Hamar BD, Bahtiyar M-O, Zhao G, Sfakianaki AK, et al. Proteomic profiling of the amniotic fluid to detect inflammation, infection, and neonatal sepsis. PLoS medicine [Internet]. 2007 Jan [cited 2013 Aug 30];4(1):e18.

117. Buhimschi CS, Weiner CP, Buhimschi I a. Clinical proteomics: a novel diagnostic tool for the new biology of preterm labor, part I: proteomics tools. Obstetrical & gynecological survey [Internet]. 2006 Jul [cited 2013 Aug 30];61(7):481–6.

118. Buhimschi CS, Rosenberg VA, Dulay AT, Thung S, Sfakianaki AK, Bahtiyar M-O, et al. Multidimensional system biology: genetic markers and proteomic biomarkers of adverse pregnancy outcome in preterm birth. American journal of perinatology [Internet]. 2008 Mar [cited 2013 Aug 30];25(3):175–87.

119. O'Brien M, Morrison JJ, Smith TJ. Upregulation of PSCDBP, TLR2, TWIST1, FLJ35382, EDNRB, and RGS12 gene expression in human myometrium at labor. Reproductive sciences (Thousand Oaks, Calif.). 2008 Apr;15(4):382–93.

120. Rehman KS, Yin S, Mayhew BA, Word RA, Rainey WE. Human myometrial adaptation to pregnancy: cDNA microarray gene expression profiling of myometrium from non-pregnant and pregnant women. Molecular human reproduction. 2003 Nov;9(11):681–700.

121. Equils O, Nambiar P, Hobel CJ, Smith R, Simmons CF, Vali S. A computer simulation of progesterone and Cox2 inhibitor treatment for preterm labor. Bader JS, editor. PloS one [Internet]. Public Library of Science; 2010 Jan [cited 2013 Aug 30];5(1):e8502.

122. Sokolowski P, Saison F, Giles W, McGrath S, Smith D, Smith J, et al. Human uterine wall tension trajectories and the onset of parturition. Bader JS, editor. PloS one. Public Library of Science; 2010 Jan;5(6):e11037.

123. Aslanidi O, Atia J, Benson a P, van den Berg H a, Blanks a M, Choi C, et al. Towards a computational reconstruction of the electrodynamics of premature and full term human labour. Progress in biophysics and molecular biology [Internet]. 2011 Oct [cited 2013 Aug 30];107(1):183–92.

124. Tong W-C, Choi CY, Kharche S, Karche S, Holden A V, Zhang H, et al. A computational model of the ionic currents, Ca2+ dynamics and action potentials underlying contraction of isolated uterine smooth muscle. PloS one. 2011 Jan;6(4):e18685.

125. Young RC. A computer model of uterine contractions based on action potential propagation and intercellular calcium waves. Obstetrics and gynecology. 1997 Apr;89(4):604–8.

126. Wanner O, Cranshaw DJ, Pliska V. The use of dynamic models to study the role of calcium in the oxytocin-induced contractions of the uterus. Molecular and Cellular Endocrinology. 1977;6:281–92.

127. Romero R, Espinoza J, Gotsch F, Kusanovic JP, Friel L a, Erez O, et al. The use of high-dimensional biology (genomics, transcriptomics, proteomics, and metabolomics) to understand the preterm parturition syndrome. BJOG : an international journal of obstetrics and gynaecology [Internet]. 2006 Dec [cited 2013 Dec 3];113 Suppl 118–35.

128. Taggart MJ, Blanks A, Kharche S, Holden A, Wang B, Zhang H. Towards understanding the myometrial physiome: approaches for the construction of a virtual physiological uterus. BMC pregnancy and childbirth. 2007 Jan;7 Suppl 1:S3.

129. Claverie JM. Computational methods for the identification of differential and coordinated gene expression. Human molecular genetics [Internet]. 1999 Jan [cited 2013 Aug 30];8(10):1821–32.

130. Lim S, MacIntyre DA, Lee YS, Khanjani S, Terzidou V, Teoh TG, et al. Nuclear factor kappa B activation occurs in the amnion prior to labour onset and modulates the expression of numerous labour associated genes. PloS one. 2012 Jan;7(4):e34707.

131. Lee KJ, Shim SH, Kang KM, Kang JH, Park DY, Kim SH, et al. Global gene expression changes induced in the human placenta during labor. Placenta. 2010 Aug;31(8):698–704.

132. Kim SY, Romero R, Tarca AL, Bhatti G, Lee J, Chaiworapongsa T, et al. miR-143 regulation of prostaglandin-endoperoxidase synthase 2 in the amnion: implications for human parturition at term. PloS one. 2011 Jan;6(9):e24131.

133. Li R, Ackerman WE, Summerfield TL, Yu L, Gulati P, Zhang J, et al. Inflammatory gene regulatory networks in amnion cells following cytokine stimulation: translational systems approach to modeling human parturition. PloS one. 2011 Jan;6(6):e20560.

134. Lee D-C, Romero R, Kim CJ, Chaiworapongsa T, Tarca AL, Lee J, et al. Surfactant protein-A as an anti-inflammatory component in the amnion: implications for human pregnancy. Journal of immunology (Baltimore, Md. : 1950). 2010 Jun;184(11):6479–91.

135. Khanjani S, Kandola MK, Lindstrom TM, Sooranna SR, Melchionda M, Lee YS, et al. NF-κB regulates a cassette of immune/inflammatory genes in human pregnant myometrium at term. Journal of cellular and molecular medicine. 2011 Apr;15(4):809–24.

136. Hassan SS, Romero R, Haddad R, Hendler I, Khalek N, Tromp G, et al. The transcriptome of the uterine cervix before and after spontaneous term parturition. American journal of obstetrics and gynecology [Internet]. 2006 Sep [cited 2013 Aug 30];195(3):778–86.

137. Vu T-D, Yun Feng, Placido J, Reznik SE. Placental matrix metalloproteinase--1 expression is increased in labor. Reproductive sciences (Thousand Oaks, Calif.). 2008 Apr;15(4):420–4.

138. Havelock JC, Keller P, Muleba N, Mayhew BA, Casey BM, Rainey WE, et al. Human myometrial gene expression before and during parturition. Biology of reproduction [Internet]. 2005 Mar [cited 2013 Aug 30];72(3):707–19.

139. Tromp G, Kuivaniemi H, Romero R, Chaiworapongsa T, Kim YM, Kim MR, et al. Genome-wide expression profiling of fetal membranes reveals a deficient expression of proteinase inhibitor 3 in premature rupture of membranes. American journal of obstetrics and gynecology. 2004 Oct;191(4):1331–8.

140. Mittal P, Romero R, Tarca AL, Gonzalez J, Draghici S, Xu Y, et al. Characterization of the myometrial transcriptome and biological pathways of spontaneous human labor at term. Journal of perinatal medicine. 2010 Nov;38(6):617–43.

141. Esplin MS, Fausett MB, Peltier MR, Hamblin S, Silver RM, Branch DW, et al. The use of cDNA microarray to identify differentially expressed labor-associated genes within the human myometrium during labor. American journal of obstetrics and gynecology [Internet]. 2005 Aug [cited 2013 Aug 30];193(2):404–13.

142. Chan EC, Fraser S, Yin S, Yeo G, Kwek K, Fairclough RJ, et al. Human myometrial genes are differentially expressed in labor: a suppression subtractive hybridization study. The Journal of clinical endocrinology and metabolism [Internet]. 2002 Jun [cited 2013 Aug 30];87(6):2435–41.

143. Gorowiec MR, Catalano RD, Norman JE, Denison FC, Jabbour HN. Prokineticin 1 induces inflammatory response in human myometrium: a potential role in initiating term and preterm parturition. The American journal of pathology [Internet]. 2011 Dec [cited 2013 Aug 30];179(6):2709–19.

144. Helguera G, Eghbali M, Sforza D, Minosyan TY, Toro L, Stefani E. Changes in global gene expression in rat myometrium in transition from late pregnancy to parturition. Physiological genomics [Internet]. 2009 Jan 8 [cited 2013 Dec 3];36(2):89–97.

145. Renthal NE, Chen C-C, Williams KC, Gerard RD, Prange-Kiel J, Mendelson CR. miR-200 family and targets, ZEB1 and ZEB2, modulate uterine quiescence and

contractility during pregnancy and labor. Proceedings of the National Academy of Sciences of the United States of America [Internet]. 2010 Nov 30 [cited 2013 Dec 3];107(48):20828–33.

146. Zhao B, Koon D, Curtis AL, Soper J, Bethin KE. Identification of 9 uterine genes that are regulated during mouse pregnancy and exhibit abnormal levels in the cyclooxygenase-1 knockout mouse. Reproductive biology and endocrinology : RB&E [Internet]. 2007 Jan [cited 2013 Dec 3];5:28.

147. Girotti M, Zingg HH. Gene expression profiling of rat uterus at different stages of parturition. Endocrinology [Internet]. 2003 Jun [cited 2013 Dec 3];144(6):2254–65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12746283

148. Mittal P, Romero R, Tarca AL, Draghici S, Nhan-Chang C-L, Chaiworapongsa T, et al. A molecular signature of an arrest of descent in human parturition. American journal of obstetrics and gynecology. 2011 Feb;204(2):177.e15–33.

149. Brennan DJ, McGee SF, Rexhepaj E, O'Connor DP, Robson M, O'Herlihy C. Identification of a myometrial molecular profile for dystocic labor. BMC pregnancy and childbirth [Internet]. BioMed Central Ltd; 2011 Jan [cited 2013 Aug 30];11(1):74.

150. Theocharidis A, van Dongen S, Enright AJ, Freeman TC. Network visualization and analysis of gene expression data using BioLayout Express(3D). Nature protocols. 2009 Jan;4(10):1535–50.

151. Dawson HD, Loveland JE, Pascal G, Gilbert JGR, Uenishi H, Mann KM, et al. Structural and functional annotation of the porcine immunome. BMC genomics [Internet]. 2013 Jan [cited 2013 Aug 30];14(1):332.

152. Galea S, Riddle M, Kaplan GA. Causal thinking and complex system approaches in epidemiology. International journal of epidemiology [Internet]. 2010 Mar [cited 2013 Aug 29];39(1):97–106.

153. Kobayashi A, Donaldson DS, Kanaya T, Fukuda S, Baillie JK, Freeman TC, et al. Identification of novel genes selectively expressed in the follicle-associated epithelium from the meta-analysis of transcriptomics data from multiple mouse cell and tissue populations. DNA research : an international journal for rapid publication of reports on genes and genomes. 2012 Oct;19(5):407–22.

154. Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, et al. InnateDB: systems biology of innate immunity and beyond--recent updates and continuing curation. Nucleic acids research [Internet]. 2013 Jan [cited 2013 Aug 30];41(Database issue):D1228–33.

155. Vora S, Abbas A, Kim CJ, Summerfield TL, Kusanovic JP, Iams JD, et al. Nuclear factor-kappa B localization and function within intrauterine tissues from term and preterm labor and cultured fetal membranes. Reproductive Biology and Endocrinology. 2010 Jan;8(1):8.

156. Smolen PD, Baxter DA, Byrne JH. Mathematical modeling and analysis of intracellular signaling pathways. From Molecules to Networks: An introduction to cellular and molecular neuroscience. New York: Academic Press; 2004. p. 391–430.

157. Aldridge BB, Burke JM, Lauffenburger D a, Sorger PK. Physicochemical modelling of cell signalling pathways. Nature cell biology [Internet]. 2006 Nov [cited 2013 Aug 29];8(11):1195–203.

158. Kitano H. Computational systems biology. Nature. 2002 Nov;420(6912):206–10.

159. Li C, Donizelli M, Rodriguez N, Dharuri H, Endler L, Chelliah V, et al. BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. BMC systems biology. 2010 Jan;4(1):92.

160. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics (Oxford, England). 2003 Mar;19(4):524–31.

161. Sharp GC, Saunders PTK, Norman JE. Computer models to study uterine activation at labour. Molecular human reproduction [Internet]. 2013 Jun 30.

162. Bursztyn L, Eytan O, Jaffa AJ, Elad D. Mathematical model of excitation-contraction in a uterine smooth muscle cell. American journal of physiology. Cell physiology [Internet]. 2007 May [cited 2013 Aug 30];292(5):C1816–29.

163. Rihana S, Terrien J, Germain G, Marque C. Mathematical modeling of electrical activity of uterine muscle cells. Medical & biological engineering & computing. 2009 Jun;47(6):665–75.

164. Barclay M, Andersen H, Simon C. Emergent behaviors in a deterministic model of the human uterus. Reproductive sciences (Thousand Oaks, Calif.) [Internet]. 2010 Oct [cited 2013 Aug 30];17(10):948–54.

165. Andersen H, Barclay M. A computer model of uterine contractions based on discrete contractile elements. Obstetrics & Gynecology [Internet]. 1995 [cited 2013 Aug 30];86:108–11.

166. Vauge C, Mignot T-M, Paris B, Breuiller-Fouché M, Chapron C, Attoui M, et al. A mathematical model for the spontaneous contractions of the isolated uterine smooth muscle from patients receiving progestin treatment. Acta biotheoretica. 2003 Jan;51(1):19–34.

167. Laforet J, Rabotti C, Terrien J, Mischi M, Marque C. Toward a multiscale model of the uterine electrical activity. IEEE transactions on bio-medical engineering. 2011 Dec;58(12):3487–90.

168. La Rosa PS, Eswaran H, Preissl H, Nehorai A. Multiscale forward electromagnetic model of uterine contractions during pregnancy. BMC medical physics. 2012 Nov;12(1):4.

169. Bisits AM, Smith R, Mesiano S, Yeo G, Kwek K, MacIntyre D, et al. Inflammatory aetiology of human myometrial activation tested using directed graphs. PLoS computational biology [Internet]. 2005 Jul [cited 2013 Aug 30];1(2):132–6.

170. University of North Texas. Welcome to CERL | CERL [Internet]. Computational epidemiology research laboratory at the University of North Texas. 2013.

171. Tameru B, Nganwa D, Bogale A, Robnett V, Habtemariam T. The Role of Computational Epidemiology and Risk Analysis in the Fight Against HIV/AIDS. Journal of AIDS & clinical research. 2012 Jul;3(6).

172. Wikipedia community. Computational epidemiology [Internet]. Wikipedia, the free encyclopedia. 2012. Available from: http://en.wikipedia.org/wiki/Computational_epidemiology

173. Brauer F. Mathematical epidemiology is not an oxymoron. BMC public health [Internet]. 2009 Jan [cited 2013 Aug 30];9 Suppl 1:S2.

174. Barendregt J. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. Population health … [Internet]. 2003 Apr [cited 2013 Aug 30];1(1):4.

175. Murphy DJ, Libby G, Chien P, Forsyth S, Greene S, Morris A. Cohort study of forceps delivery and the risk of epilepsy in adulthood. American journal of obstetrics and gynecology. 2004 Aug;191(2):392–7.

176. Uma R, Libby G, Murphy J. Obstetric management of a woman ' s first delivery and the implications for pelvic floor surgery in later life. Journal of Obstetrics and Gynaecology. 2005;112(August):1043–6.

177. Vera-Badillo FE, Shapiro R, Ocana A, Amir E, Tannock IF. Bias in reporting of end points of efficacy and toxicity in randomized, clinical trials for women with breast cancer. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO [Internet]. 2013 May [cited 2013 Aug 13];24(5):1238–44.

178. Ramasamy A, Mondry A, Holmes CC, Altman DG. Key issues in conducting a meta-analysis of gene expression microarray datasets. PLoS medicine. 2008 Sep;5(9):e184.

179. Stevens JR, Doerge RW. Meta-analysis combines affymetrix microarray results across laboratories. Comparative and functional genomics. 2005 Jan;6(3):116–22.

180. Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer research. 2002 Aug;62(15):4427–33.

181. Choi JK, Yu U, Kim S, Yoo OJ. Combining multiple microarray studies and modeling interstudy variation. Bioinformatics [Internet]. 2003 Jan [cited 2013 Aug 30];19 Suppl 1:i84–90.

182. Hu P, Wang X, Haitsma JJ, Furmli S, Masoom H, Liu M, et al. Microarray meta-analysis identifies acute lung injury biomarkers in donor lungs that predict development of primary graft failure in recipients. PloS one [Internet]. 2012 Jan [cited 2013 Aug 30];7(10):e45506.

183. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nature genetics [Internet]. 2001 Dec [cited 2013 Aug 30];29(4):365–71.

184. Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. Nature. 2000 Jun;405(6788):827–36.

185. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. Nature genetics [Internet]. 1999 Jan [cited 2013 Aug 30];21(1 Suppl):33–7.

186. Noordewier MO, Warren P V. Gene expression microarrays and the integration of biological knowledge. Trends in biotechnology. 2001 Oct;19(10):412–5.

187. Mutter GL, Zahrieh D, Liu C, Neuberg D, Finkelstein D, Baker HE, et al. Comparison of frozen and RNALater solid tissue storage methods for use in RNA expression microarrays. BMC genomics. 2004 Nov;5(1):88.

188. Du P, Kibbe W, Lin S. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008;24(13):1547–8.

189. Du P, Kibbe WA, Lin SM. nuID: a universal naming scheme of oligonucleotides for illumina, affymetrix, and other microarrays. Biology direct [Internet]. 2007 Jan [cited 2013 Aug 30];2:16.

190. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. Nucleic acids research. 2008 Feb;36(2):e11.

191. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC bioinformatics [Internet]. 2010 Jan [cited 2013 Aug 30];11:587.

192. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics (Oxford, England) [Internet]. 2003 Jan 22 [cited 2013 Aug 9];19(2):185–93.

193. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, et al. gplots: Various R programming tools for plotting data. R package. [Internet]. 2012. Available from: cran.r-project.org/package=gplots

194. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, S D, Irizarry W, editors. Bioinformatics and Computational Biology Solutions using R and Bioconductor. New York: Springer; 2005. p. 397–420.

195. Fieller EC. Some problems in interval estimation. Journal of the Royal Statistical Society Series B Methodological. JSTOR; 1954;16(2):175–85.

196. Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS letters [Internet]. 2004 Aug 27 [cited 2013 Aug 30];573(1-3):83–92.

197. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols [Internet]. 2009 Jan [cited 2013 Aug 9];4(1):44–57.

198. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research [Internet]. 2009 Jan [cited 2013 Aug 9];37(1):1–13.

199. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics [Internet]. 2000 May [cited 2013 Aug 6];25(1):25–9.

200. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research. 2000 Jan;28(1):27–30.

201. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic acids research. 2012 Jan;40(Database issue):D109–14.

202. Gilbert DR, Heiner M, Breitling R, Orton R. MAP Kinase Signaling Protocols. In: Seger R, editor. Life Sciences. Second. New York: Springer Science; 2010. p. 369–84.

203. Breitling R, Gilbert D, Heiner M, Orton R. A structured approach for the engineering of biochemical network models, illustrated for signalling pathways. Briefings in bioinformatics [Internet]. 2008 Sep [cited 2013 Aug 30];9(5):404–21.

204. Hoffmann A, Levchenko A, Scott ML, Baltimore D. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. Science [Internet]. 2002 Nov 8 [cited 2013 Aug 30];298(5596):1241–5.

205. Covert MW, Leung TH, Gaston JE, Baltimore D. Achieving stability of lipopolysaccharide-induced NF-kappaB activation. Science [Internet]. 2005 Sep 16 [cited 2013 Aug 30];309(5742):1854–7.

206. Nelson DE, Ihekwaba a EC, Elliott M, Johnson JR, Gibney C a, Foreman BE, et al. Oscillations in NF-kappaB signaling control the dynamics of gene expression. Science (New York, N.Y.). 2004 Oct;306(5696):704–8.

207. Moodie S, Le Novere N, Sorokin A, Mi H, Schreiber F. Systems Biology Graphical Notation: Process Description language Level 1. Nature Precedings. 2009 Sep;

208. Freeman TC, Raza S, Theocharidis A, Ghazal P. The mEPN scheme: an intuitive and flexible graphical system for rendering biological pathways. BMC systems biology [Internet]. 2010 Jan [cited 2013 Aug 30];4(1):65.

209. Funahashi A, Matsuoka Y, Jouraku A, Morohashi M, Kikuchi N, Kitano H. CellDesigner 3.5 : A Versatile Modeling Tool for Biochemical Networks. Proceedings of the IEEE. Institute of Electrical and Electronics Engineers; 2008;96(8):1254–65.

210. yWorks the diagramming company. yEd Graph Editor [Internet]. 2013. Available from: http://www.yworks.com/en/products_yed_about.html

211. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, et al. COPASI--a COmplex PAthway SImulator. Bioinformatics (Oxford, England) [Internet]. 2006 Dec 15 [cited 2013 Aug 30];22(24):3067–74.

212. Libby G, Smith A, McEwan NF, Chien PFW, Greene S a, Forsyth JS, et al. The Walker Project: a longitudinal study of 48,000 children born 1952-1966 (aged 36-50 years in 2002) and their families. Paediatric and perinatal epidemiology. 2004 Jul;18(4):302–12.

213. Hellerstein J. Quantitative cleaning for large databases. White Paper, United Nations Economic Commission for Europe. 2008;

214. Bates D, Maechler M, Bolker B. lme4: Linear mixed-effects models using S4 classes. R package [Internet]. 2012 [cited 2013 Aug 30]. Available from: http://cran.r-project.org/package=lme4

215. Venkatraman ES. clinfun: Clinical Trial Design and Data Analysis Functions. R package. [Internet]. 2013. Available from: http://cran.r-project.org/package=clinfun

216. GraphPad Software. Advice: Don't compute the power to detect the difference actually observed [Internet]. 2013. Available from: http://www.graphpad.com/guides/prism/5/user-guide/prism5help.html?stat_why_it_isnt_helpful_to_compute.htm

217. Sharp G, Stock S, Norman J. Fetal assessment methods for improving neonatal and maternal outcomes in preterm prelabour rupture of membranes (Protocol). Cochrane Database of Systematic Reviews. 2012;(11).

218. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. The Cochrane Collaboration; 2011.

219. Gravett MG, Novy MJ, Rosenfeld RG, Reddy AP, Jacob T, Turner M, et al. Diagnosis of intra-amniotic infection by proteomic profiling and identification of novel biomarkers. JAMA : the journal of the American Medical Association [Internet]. 2004 Jul 28 [cited 2013 Aug 30];292(4):462–9.

220. Larsson E, McLean SE, Mecham RP, Lindahl P, Nelander S. Do two mutually exclusive gene modules define the phenotypic diversity of mammalian smooth muscle? Molecular genetics and genomics : MGG. 2008 Aug;280(2):127–37.

221. Shabo I, Svanvik J. Expression of macrophage antigens by tumor cells. Advances in experimental medicine and biology [Internet]. 2011 Jan [cited 2013 Aug 7];714:141–50.

222. Webb RC. Smooth muscle contraction and relaxation. Advances in physiology education [Internet]. 2003 Dec [cited 2013 Aug 7];27(1-4):201–6.

223. Bollapragada S, Bollopragada S, Youssef R, Jordan F, Greer I, Norman J, et al. Term labor is associated with a core inflammatory response in human fetal membranes, myometrium, and cervix. American journal of obstetrics and gynecology [Internet]. Mosby, Inc.; 2009 Jan [cited 2013 Aug 30];200(1):104.e1–11.

224. Greig PC, Herbert WN, Robinette BL, Teot LA. Amniotic fluid interleukin-10 concentrations increase through pregnancy and are elevated in patients with preterm labor associated with intrauterine infection. American journal of obstetrics and gynecology. 1995 Oct;173(4):1223–7.

225. Simpson KL, Keelan JA, Mitchell MD. Labor-associated changes in interleukin-10 production and its regulation by immunomodulators in human choriodecidua. The Journal of clinical endocrinology and metabolism. 1998 Dec;83(12):4332–7.

226. Phillips RJ, Al-Zamil H, Hunt LP, Fortier M a, López Bernal a. Genes for prostaglandin synthesis, transport and inactivation are differentially expressed in human uterine tissues, and the prostaglandin F synthase AKR1B1 is induced in myometrial cells by inflammatory cytokines. Molecular human reproduction. 2011 Jan;17(1):1–13.

227. Astle S, Newton R, Thornton S, Vatish M, Slater DM. Expression and regulation of prostaglandin E synthase isoforms in human myometrium with labour. Molecular human reproduction [Internet]. 2007 Jan [cited 2013 Aug 7];13(1):69–75.

228. Li L, Kang J, Lei W. Role of Toll-like receptor 4 in inflammation-induced preterm delivery. Molecular Human Reproduction. 2010;16(4):267–72.

229. Tattersall M, Engineer N, Khanjani S, Sooranna SR, Roberts VH, Grigsby PL, et al. Pro-labour myometrial gene expression: are preterm labour and term labour the same? Reproduction (Cambridge, England). 2008 Apr;135(4):569–79.

230. Challis JRG, Matthews SG, Gibb W, Lye SJ. Endocrine and paracrine regulation of birth at term and preterm. Endocrine reviews [Internet]. 2000 Oct [cited 2013 Aug 30];21(5):514–50.

231. Challis J, Sloboda D, Alfaidy N, Lye S, Gibb W, Patel F, et al. Prostaglandins and mechanisms of preterm birth. Reproduction [Internet]. 2002 Jul 1 [cited 2013 Aug 7];124(1):1–17.

232. Sehringer B, Schäfer WR, Wetzka B, Deppert WR, Brunner-Spahr R, Benedek E, et al. Formation of proinflammatory cytokines in human term myometrium is stimulated by lipopolysaccharide but not by corticotropin-releasing hormone. The Journal of clinical endocrinology and metabolism. 2000 Dec;85(12):4859–65.

233. Elovitz MA, Wang Z, Chien EK, Rychlik DF, Phillippe M. A new model for inflammation-induced preterm birth: the role of platelet-activating factor and Toll-like receptor-4. American Journal of Pathology [Internet]. 2003 Nov [cited 2013 Dec 17];163(5):2103–11.

234. Fidel PL, Romero R, Ramirez M, Cutright J, Edwin SS, LaMarche S, et al. Interleukin-1 receptor antagonist (IL-1ra) production by human amnion, chorion, and decidua. American journal of reproductive immunology. 1994 Aug;32(1):1–7.

235. Wang H, Hirsch E. Bacterially-induced preterm labor and regulation of prostaglandin-metabolizing enzyme expression in mice: the role of toll-like receptor 4. Biology of reproduction. 2003 Dec;69(6):1957–63.

236. Youssef RE, Ledingham MA, Bollapragada SS, O'Gorman N, Jordan F, Young A, et al. The role of toll-like receptors (TLR-2 and -4) and triggering receptor expressed on myeloid cells 1 (TREM-1) in human term and preterm labor. Reproductive sciences (Thousand Oaks, Calif.) [Internet]. 2009 Sep 1 [cited 2013 Dec 17];16(9):843–56.

237. Kim YM, Romero R, Chaiworapongsa T, Kim GJ, Kim MR, Kuivaniemi H, et al. Toll-like receptor-2 and -4 in the chorioamniotic membranes in spontaneous labor at term and in preterm parturition that are associated with chorioamnionitis. American journal of obstetrics and gynecology [Internet]. 2004 Oct [cited 2013 Dec 17];191(4):1346–55.

238. Li L, Kang J, Lei W. Role of Toll-like receptor 4 in inflammation-induced preterm delivery. Molecular human reproduction [Internet]. 2010 Apr [cited 2013 Dec 17];16(4):267–72.

239. Ilievski V, Hirsch E. Synergy between viral and bacterial toll-like receptors leads to amplification of inflammatory responses and preterm labor in the mouse. Biology of reproduction [Internet]. 2010 Nov [cited 2013 Dec 17];83(5):767–73.

240. Equils O, Moffatt-Blue C, Ishikawa T, Simmons CF, Ilievski V, Hirsch E. Pretreatment with pancaspase inhibitor (Z-VAD-FMK) delays but does not prevent intraperitoneal heat-killed group B Streptococcus-induced preterm delivery in a pregnant mouse model. Infectious diseases in obstetrics and gynecology [Internet]. 2009 Jan [cited 2013 Dec 17];2009:749432.

241. Kawai T, Akira S. Signaling to NF-kappaB by Toll-like receptors. Trends in molecular medicine [Internet]. 2007 Nov [cited 2013 Dec 17];13(11):460–9.

242. Croston GE, Cao Z, Goeddel D V. NF-kappa B activation by interleukin-1 (IL-1) requires an IL-1 receptor-associated protein kinase activity. The Journal of biological chemistry [Internet]. 1995 Jul 14 [cited 2013 Aug 30];270(28):16514–7.

243. Dokter WH, Tuyt L, Sierdsema SJ, Esselink MT, Vellenga E. The spontaneous expression of interleukin-1 beta and interleukin-6 is associated with spontaneous expression of AP-1 and NF-kappa B transcription factor in acute myeloblastic leukemia cells. Leukemia [Internet]. 1995 Mar [cited 2013 Aug 30];9(3):425–32.

244. Wu GD, Lai EJ, Huang N, Wen X. Oct-1 and CCAAT/enhancer-binding protein (C/EBP) bind to overlapping elements within the interleukin-8 promoter. The role of Oct-1 as a transcriptional repressor. The Journal of biological chemistry. 1997 Jan;272(4):2396–403.

245. Newton R, Kuitert LM, Bergmann M, Adcock IM, Barnes PJ. Evidence for involvement of NF-kappaB in the transcriptional control of COX-2 gene expression by IL-1beta. Biochemical and biophysical research communications. 1997 Aug;237(1):28–32.

246. Allport VC. Human labour is associated with nuclear factor-kappaB activity which mediates cyclo-oxygenase-2 expression and is involved with the "functional progesterone withdrawal."Molecular Human Reproduction [Internet]. 2001 Jun 1 [cited 2013 Aug 30];7(6):581–6.

247. Chapman NR, Europe-Finner GN, Robson SC. Expression and deoxyribonucleic acid-binding activity of the nuclear factor kappaB family in the human myometrium during pregnancy and labor. The Journal of clinical endocrinology and metabolism [Internet]. 2004 Nov [cited 2013 Aug 30];89(11):5683–93.

248. Golightly E, Jabbour HN, Norman JE. Endocrine immune interactions in human parturition. Molecular and cellular endocrinology [Internet]. Elsevier Ireland Ltd; 2011 Mar 15 [cited 2013 Aug 30];335(1):52–9.

249. Schumann RR, Leong SR, Flaggs GW, Gray PW, Wright SD, Mathison JC, et al. Structure and function of lipopolysaccharide binding protein. Science (New York, N.Y.). 1990 Sep;249(4975):1429–31.

250. Wright SD, Ramos RA, Tobias PS, Ulevitch RJ, Mathison JC. CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. Science (New York, N.Y.). 1990 Sep;249(4975):1431–3.

251. Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, et al. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science (New York, N.Y.). 1998 Dec;282(5396):2085–8.

252. Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, et al. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the LPS gene product. Journal of immunology (Baltimore, Md. : 1950) [Internet]. 1999 Apr 1 [cited 2013 Aug 30];162(7):3749–52.

253. Takeuchi O, Hoshino K, Kawai T, Sanjo H, Takada H, Ogawa T, et al. Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. Immunity. 1999 Oct;11(4):443–51.

254. Fenton MJ, Golenbock DT. LPS-binding proteins and receptors. Journal of leukocyte biology. 1998 Jul;64(1):25–32.

255. Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR, et al. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. Nature [Internet]. 2001 Apr 26 [cited 2013 Aug 30];410(6832):1099–103.

256. Hemmi H, Kaisho T, Takeuchi O, Sato S, Sanjo H, Hoshino K, et al. Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. Nature immunology [Internet]. 2002 Mar [cited 2013 Aug 30];3(2):196–200.

257. Schnare M, Holt AC, Takeda K, Akira S, Medzhitov R. Recognition of CpG DNA is mediated by signaling pathways dependent on the adaptor protein MyD88. Current biology : CB. 2000 Sep;10(18):1139–42.

258. Häcker H, Vabulas RM, Takeuchi O, Hoshino K, Akira S, Wagner H. Immune cell activation by bacterial CpG-DNA through myeloid differentiation marker 88 and tumor necrosis factor receptor-associated factor (TRAF)6. The Journal of experimental medicine. 2000 Aug;192(4):595–600.

259. Takeuchi O, Takeda K, Hoshino K, Adachi O, Ogawa T, Akira S. Cellular responses to bacterial cell wall components are mediated through MyD88-dependent signaling cascades. International immunology. 2000 Jan;12(1):113–7.

260. Kawai T, Adachi O, Ogawa T, Takeda K, Akira S. Unresponsiveness of MyD88-deficient mice to endotoxin. Immunity. 1999 Jul;11(1):115–22.

261. Kawai T, Takeuchi O, Fujita T, Inoue J, Mühlradt PF, Sato S, et al. Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IFN-regulatory factor 3 and the expression of a subset of lipopolysaccharide-inducible genes. Journal of immunology (Baltimore, Md. : 1950). 2001 Nov;167(10):5887–94.

262. Doyle S, Vaidya S, O'Connell R, Dadgostar H, Dempsey P, Wu T, et al. IRF3 mediates a TLR3/TLR4-specific antiviral gene program. Immunity. 2002 Sep;17(3):251–63.

263. Toshchakov V, Jones BW, Perera P-Y, Thomas K, Cody MJ, Zhang S, et al. TLR4, but not TLR2, mediates IFN-beta-induced STAT1alpha/beta-dependent gene expression in macrophages. Nature immunology. 2002 Apr;3(4):392–8.

264. Hoshino K, Kaisho T, Iwabe T, Takeuchi O, Akira S. Differential involvement of IFN-beta in Toll-like receptor-stimulated dendritic cell activation. International immunology. 2002 Oct;14(10):1225–31.

265. Lee MS, Kim Y-J. Signaling pathways downstream of pattern-recognition receptors and their cross talk. Annual review of biochemistry [Internet]. 2007 Jan [cited 2013 Aug 30];76(Il):447–80.

266. Yamamoto M, Sato S, Hemmi H, Sanjo H, Uematsu S, Kaisho T, et al. Essential role for TIRAP in activation of the signalling cascade shared by TLR2 and TLR4. Nature. 2002 Nov;420(6913):324–9.

267. Horng T, Barton GM, Flavell RA, Medzhitov R. The adaptor molecule TIRAP provides signalling specificity for Toll-like receptors. Nature [Internet]. 2002 Nov 21 [cited 2013 Aug 30];420(6913):329–33.

268. Wesche H, Henzel WJ, Shillinglaw W, Li S, Cao Z. MyD88: an adapter that recruits IRAK to the IL-1 receptor complex. Immunity. 1997 Dec;7(6):837–47.

269. Li S, Strelow A, Fontana EJ, Wesche H. IRAK-4: a novel member of the IRAK family with the properties of an IRAK-kinase. Proceedings of the National Academy of Sciences of the United States of America. 2002 Apr;99(8):5567–72.

270. Suzuki N, Suzuki S, Yeh W-C. IRAK-4 as the central TIR signaling mediator in innate immunity. Trends in immunology. 2002 Oct;23(10):503–6.

271. Gottipati S, Rao NL, Fung-Leung W-P. IRAK1: a critical signaling mediator of innate immunity. Cellular signalling [Internet]. 2008 Mar [cited 2013 Aug 30];20(2):269–76.

272. Ross K, Yang L, Dower S, Volpe F, Guesdon F. Identification of threonine 66 as a functionally critical residue of the interleukin-1 receptor-associated kinase. The Journal of biological chemistry. 2002 Oct;277(40):37414–21.

273. Muroi M, Tanamoto K. TRAF6 distinctively mediates MyD88- and IRAK-1-induced activation of NF-kappaB. Journal of leukocyte biology. 2008 Mar;83(3):702–7.

274. Dauphinee SM, Karsan A. Lipopolysaccharide signaling in endothelial cells. Laboratory investigation; a journal of technical methods and pathology [Internet]. 2006 Jan [cited 2013 Aug 12];86(1):9–22.

275. Kollewe C, Mackensen A-C, Neumann D, Knop J, Cao P, Li S, et al. Sequential autophosphorylation steps in the interleukin-1 receptor-associated kinase-1 regulate its availability as an adapter in interleukin-1 signaling. The Journal of biological chemistry. 2004 Feb;279(7):5227–36.

276. Jiang Z, Ninomiya-Tsuji J, Qian Y, Matsumoto K, Li X. Interleukin-1 (IL-1) receptor-associated kinase-dependent IL-1-induced signaling complexes phosphorylate TAK1 and TAB2 at the plasma membrane and activate TAK1 in the cytosol. Molecular and cellular biology. 2002 Oct;22(20):7158–67.

277. Yamin TT, Miller DK. The interleukin-1 receptor-associated kinase is degraded by proteasomes following its phosphorylation. The Journal of biological chemistry. 1997 Aug;272(34):21540–7.

278. Irie T, Muta T, Takeshige K. TAK1 mediates an activation signal from toll-like receptor(s) to nuclear factor-kappaB in lipopolysaccharide-stimulated macrophages. FEBS letters. 2000 Mar;467(2-3):160–4.

279. Wang C, Deng L, Hong M, Akkaraju GR, Inoue J, Chen ZJ. TAK1 is a ubiquitin-dependent kinase of MKK and IKK. Nature. 2001 Jul;412(6844):346–51.

280. Arch R, Gedrich R, Thompson C. Tumor necrosis factor receptor-associated factors (TRAFs)—a family of adapter proteins that regulates life and death. Genes & development [Internet]. 1998 Sep [cited 2013 Aug 30];12(18):2821–30.

281. Adhikari A, Xu M, Chen ZJ. Ubiquitin-mediated activation of TAK1 and IKK. Oncogene [Internet]. 2007 May 14 [cited 2013 Aug 26];26(22):3214–26.

282. Yamamoto M, Sato S, Mori K, Hoshino K, Takeuchi O, Takeda K, et al. Cutting edge: a novel Toll/IL-1 receptor domain-containing adapter that preferentially activates the IFN-beta promoter in the Toll-like receptor signaling. Journal of immunology (Baltimore, Md. : 1950). 2002 Dec;169(12):6668–72.

283. Yamamoto M, Sato S, Hemmi H, Uematsu S, Hoshino K, Kaisho T, et al. TRAM is specifically involved in the Toll-like receptor 4-mediated MyD88-independent signaling pathway. Nature immunology. 2003 Nov;4(11):1144–50.

284. Fitzgerald KA, Rowe DC, Golenbock DT. Endotoxin recognition and signal transduction by the TLR4/MD2-complex. Microbes and infection / Institut Pasteur [Internet]. 2004 Dec [cited 2013 Aug 30];6(15):1361–7.

285. Yamamoto M, Sato S, Hemmi H, Hoshino K, Kaisho T, Sanjo H, et al. Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. Science. 2003 Aug;301(5633):640–3.

286. Fitzgerald KA, McWhirter SM, Faia KL, Rowe DC, Latz E, Golenbock DT, et al. IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. Nature immunology [Internet]. 2003 May [cited 2013 Aug 20];4(5):491–6.

287. Hemmi H, Takeuchi O, Sato S, Yamamoto M, Kaisho T, Sanjo H, et al. The roles of two IkappaB kinase-related kinases in lipopolysaccharide and double stranded RNA signaling and viral infection. The Journal of experimental medicine [Internet]. 2004 Jun 21 [cited 2013 Aug 19];199(12):1641–50.

288. Häcker H, Karin M. Regulation and function of IKK and IKK-related kinases. Science's STKE : signal transduction knowledge environment [Internet]. 2006 Oct 17 [cited 2013 Aug 12];2006(357):re13.

289. Mori M, Yoneyama M, Ito T, Takahashi K, Inagaki F, Fujita T. Identification of Ser-386 of interferon regulatory factor 3 as critical target for inducible phosphorylation that determines activation. The Journal of biological chemistry. 2004 Mar;279(11):9698–702.

290. Stauber GB, Aiyer RA, Aggarwal BB. Human tumor necrosis factor-alpha receptor. Purification by immunoaffinity chromatography and initial characterization. The Journal of biological chemistry. 1988 Dec;263(35):19098–104.

291. Kelliher MA, Grimm S, Ishida Y, Kuo F, Stanger BZ, Leder P. The death domain kinase RIP mediates the TNF-induced NF-kappaB signal. Immunity. 1998 Mar;8(3):297–303.

292. Zhang J, Cado D, Chen A, Kabra NH, Winoto A. Fas-mediated apoptosis and activation-induced T-cell proliferation are defective in mice lacking FADD/Mort1. Nature. 1998 Mar;392(6673):296–300.

293. Yeh WC, Shahinian A, Speiser D, Kraunus J, Billia F, Wakeham A, et al. Early lethality, functional NF-kappaB activation, and increased sensitivity to TNF-induced cell death in TRAF2-deficient mice. Immunity. 1997 Nov;7(5):715–25.

294. Liu ZG, Hsu H, Goeddel D V, Karin M. Dissection of TNF receptor 1 effector functions: JNK activation is not linked to apoptosis while NF-kappaB activation prevents cell death. Cell. 1996 Nov;87(3):565–76.

295. Ea C-K, Deng L, Xia Z-P, Pineda G, Chen ZJ. Activation of IKK by TNFalpha requires site-specific ubiquitination of RIP1 and polyubiquitin binding by NEMO. Molecular cell [Internet]. 2006 Apr 21 [cited 2013 Aug 30];22(2):245–57.

296. Gil J, Alcamí J, Esteban M. Activation of NF-kappa B by the dsRNA-dependent protein kinase, PKR involves the I kappa B kinase complex. Oncogene [Internet]. 2000 Mar 9 [cited 2013 Aug 30];19(11):1369–78.

297. Brown K, Park S, Kanno T, Franzoso G, Siebenlist U. Mutual regulation of the transcriptional activator NF-kappa B and its inhibitor, I kappa B-alpha. Proceedings of the National Academy of Sciences of the United States of America [Internet]. 1993 Mar 15 [cited 2013 Aug 30];90(6):2532–6.

298. Scott ML, Fujita T, Liou HC, Nolan GP, Baltimore D. The p65 subunit of NF-kappa B regulates I kappa B by two distinct mechanisms. Genes & development. 1993 Jul;7(7A):1266–76.

299. Sun SC, Ganchi PA, Ballard DW, Greene WC. NF-kappa B controls expression of inhibitor I kappa B alpha: evidence for an inducible autoregulatory pathway. Science (New York, N.Y.). 1993 Mar;259(5103):1912–5.

300. Selvarajoo K. Discovering differential activation machinery of the Toll-like receptor 4 signaling pathways in MyD88 knockouts. FEBS letters. 2006 Feb;580(5):1457–64.

301. Romero R, Mazor M, Tartakovsky B. Systemic administration of interleukin-1 induces preterm parturition in mice. American journal of obstetrics and gynecology [Internet]. 1991 Oct [cited 2013 Dec 11];165(4 Pt 1):969–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1951564

302. Sadowsky DW, Novy MJ, Witkin SS, Gravett MG. Dexamethasone or interleukin-10 blocks interleukin-1beta-induced uterine contractions in pregnant rhesus monkeys. American journal of obstetrics and gynecology [Internet]. 2003 Jan [cited 2013 Dec 11];188(1):252–63.

303. Zhang G, Ghosh S. Toll-like receptor-mediated NF-kappaB activation: a phylogenetically conserved paradigm in innate immunity. The Journal of clinical investigation. 2001 Jan;107(1):13–9.

304. Dallot E, Méhats C, Oger S, Leroy M-J, Breuiller-Fouché M. A role for PKCzeta in the LPS-induced translocation NF-kappaB p65 subunit in cultured myometrial cells. Biochimie [Internet]. 2005 Jun [cited 2013 Dec 11];87(6):513–21.

305. Sharp GC, Ma H, Saunders PTK, Norman JE. A computational model of lipopolysaccharide-induced nuclear factor kappa B activation: a key signalling pathway in infection-induced preterm labour. PloS one [Internet]. 2013 Jan [cited 2013 Oct 7];8(7):e70180.

306. Natividad A, Freeman TC, Jeffries D, Burton MJ, Mabey DCW, Bailey RL, et al. Human conjunctival transcriptome analysis reveals the prominence of innate defense in Chlamydia trachomatis infection. Infection and immunity. 2010 Nov;78(11):4895–911.

307. Mabbott NA, Kenneth Baillie J, Hume DA, Freeman TC. Meta-analysis of lineage-specific gene expression signatures in mouse leukocyte populations. Immunobiology [Internet]. 2010 [cited 2013 Aug 30];215(9-10):724–36.

308. Freeman TC, Goldovsky L, Brosch M, van Dongen S, Mazière P, Grocock RJ, et al. Construction, visualisation, and clustering of transcription networks from microarray expression data. PLoS computational biology [Internet]. 2007 Oct [cited 2013 Aug 30];3(10):2032–42.

309. Yuan P, Qiao L, Dai L, Wang Y-P, Zhou G-X, Han Y, et al. Spatial distribution patterns of anorectal atresia/stenosis in China: use of two-dimensional graph-theoretical clustering. World journal of gastroenterology : WJG. 2009 Jun;15(22):2787–93.

310. Salathé M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH. A high-resolution human contact network for infectious disease transmission. Proceedings of the National Academy of Sciences of the United States of America. 2010 Dec;107(51):22020–5.

311. Groendyke C, Welch D, Hunter DR. A network-based analysis of the 1861 Hagelloch measles data. Biometrics [Internet]. 2012 Sep [cited 2013 Aug 30];68(3):755–65.

312. Lazer S, Biale Y, Mazor M, Lewenthal H, Insler V. Complications associated with the macrosomic fetus. The Journal of reproductive medicine. 1986 Jun;31(6):501–5.

313. Low J. Caesarean section--past and present. Journal of obstetrics and gynaecology Canada. 2009 Dec;31(12):1131–6.

314. Menacker F, Declercq E, Macdorman MF. Cesarean delivery: background, trends, and epidemiology. Seminars in perinatology. 2006 Oct;30(5):235–41.

315. Royal College of Obstetricians and Gynaecologists. Induction of labour. Evidence-based Clinical Guideline. London: RCOG Press; 2001.

316. Royal College of Obstetricians and Gynaecologists. Hyptertension in pregnancy: the management of hyptertensive disorders during pregnancy. NICE Clinical guideline. London: RCOG Press; 2010.

317. Tsvieli O, Sergienko R, Sheiner E. Risk factors and perinatal outcome of pregnancies complicated with cephalopelvic disproportion: a population-based study. Archives of gynecology and obstetrics. 2012 Apr;285(4):931–6.

318. Stewart PJ, Dulberg C, Arnill AC, Elmslie T, Hall PF. Diagnosis of dystocia and management with cesarean section among primiparous women in Ottawa-Carleton. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 1990 Mar;142(5):459–63.

319. Ponkey SE, Cohen AP, Heffner LJ, Lieberman E. Persistent fetal occiput posterior position: obstetric outcomes. Obstetrics and gynecology. 2003 May;101(5 Pt 1):915–20.

320. Fitzpatrick M, McQuillan K, O'Herlihy C. Influence of persistent occiput posterior position on delivery outcome. Obstetrics and gynecology. 2001 Dec;98(6):1027–31.

321. Hofmeyr GJ, Hannah ME. Planned caesarean section for term breech delivery. The Cochrane database of systematic reviews [Internet]. 2003 Jan [cited 2013 Aug 22];(3):CD000166.

322. Royal College of Obstetricians and Gynaecologists. Operative vaginal delivery. Green-top guideline number 26. London: RCOG Press; 2011.

323. Lowe NK. A review of factors associated with dystocia and cesarean section in nulliparous women. Journal of midwifery & women's health [Internet]. 2007 [cited 2013 Aug 30];52(3):216–28.

324. Selo-Ojeme DO. Primary postpartum haemorrhage. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology. 2002 Sep;22(5):463–9.

325. Ford JB, Roberts CL, Bell JC, Algert CS, Morris JM. Postpartum haemorrhage occurrence and recurrence: a population-based study. The medical journal of Australia [Internet]. 2007 Oct 1 [cited 2013 Aug 30];187(7):391–3.

326. Magann EF, Evans S, Hutchinson M, Collins R, Howard BC, Morrison JC. Postpartum hemorrhage after vaginal birth: an analysis of risk factors. Southern medical journal. 2005 Apr;98(4):419–22.

327. Combs CA, Murphy EL, Laros RK. Factors associated with postpartum hemorrhage with vaginal birth. Obstetrics and gynecology [Internet]. 1991 Jan [cited 2013 Aug 30];77(1):69–76.

328. Hall MH, Halliwell R, Carr-Hill R. Concomitant and repeated happenings of complications of the third stage of labour. British journal of obstetrics and gynaecology. 1985 Jul;92(7):732–8.

329. Sosa CG, Althabe F, Belizán JM, Buekens P. Risk factors for postpartum hemorrhage in vaginal deliveries in a Latin-American population. Obstetrics and gynecology. 2009 Jun;113(6):1313–9.

330. Sheiner E, Sarid L, Levy A, Seidman DS, Hallak M. Obstetric risk factors and outcome of pregnancies complicated with early postpartum hemorrhage: a population-based study. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians. 2005 Sep;18(3):149–54.

331. Brinsden PR, Clark AD. Postpartum haemorrhage after induced and spontaneous labour. British medical journal [Internet]. 1978 Sep 23 [cited 2013 Aug 30];2(6141):855–6.

332. Oyelese Y, Ananth C V. Postpartum hemorrhage: epidemiology, risk factors, and causes. Clinical obstetrics and gynecology. 2010 Mar;53(1):147–56.

333. Lutomski JE, Byrne BM, Devane D, Greene R a. Increasing trends in atonic postpartum haemorrhage in Ireland: an 11-year population-based cohort study. BJOG : an international journal of obstetrics and gynaecology. 2012 Feb;119(3):306–14.

334. Louis GB, Dukic V, Heagerty PJ, Louis TA, Lynch CDCD, Ryan LM, et al. Analysis of repeated pregnancy outcomes. Statistical methods in medical research. 2006 Apr;15(2):103–26.

335. Bhattacharya S, Raja EA, Mirazo ER, Campbell DM, Lee AJ, Norman JE, et al. Inherited predisposition to spontaneous preterm delivery. Obstetrics and gynecology [Internet]. 2010 Jun [cited 2013 Aug 30];115(6):1125–33.

336. Mogren I, Högberg U, Winkvist A, Stenlund H. Familial occurrence of preeclampsia. Epidemiology (Cambridge, Mass.). 1999 Sep;10(5):518–22.

337. Douvas SG, Brewer MJ, McKay ML, Rhodes PG, Kahlstorf JH, Morrison JC. Treatment of premature rupture of the membranes. The Journal of reproductive medicine. 1984 Oct;29(10):741–4.

338. Merenstein GB, Weisman LE. Premature rupture of the membranes: neonatal consequences. Seminars in perinatology. 1996 Oct;20(5):375–80.

339. Kenyon S, Pike K, Jones DR, Brocklehurst P, Marlow N, Salt a, et al. Childhood outcomes after prescription of antibiotics to pregnant women with spontaneous preterm labour: 7-year follow-up of the ORACLE II trial. Lancet. Elsevier Ltd; 2008 Oct;372(9646):1319–27.

340. Abou El Senoun G, Dowswell T, Mousa HA. Planned home versus hospital care for preterm prelabour rupture of the membranes (PPROM) prior to 37 weeks' gestation. The Cochrane database of systematic reviews [Internet]. 2010 Jan [cited 2013 Aug 30];(4):CD008053.

341. Alfirevic Z, Stampalija T, Gyte GM. Fetal and umbilical Doppler ultrasound in high-risk pregnancies. The Cochrane database of systematic reviews [Internet]. 2010 Jan [cited 2013 Aug 30];(1):CD007529.

342. Alfirevic Z, Stampalija T, Gyte GM. Fetal and umbilical Doppler ultrasound in normal pregnancy. The Cochrane database of systematic reviews [Internet]. 2010 Jan [cited 2013 Aug 30];(8):CD001450.

343. Grivell RM, Alfirevic Z, Gyte GM, Devane D. Antenatal cardiotocography for fetal assessment. The Cochrane database of systematic reviews [Internet]. 2010 Jan [cited 2013 Aug 30];(1):CD007863.

344. Lalor JG, Fawole B, Alfirevic Z, Devane D. Biophysical profile for fetal assessment in high risk pregnancies. The Cochrane database of systematic reviews [Internet]. 2008 Jan [cited 2013 Aug 30];(1):CD000038.

345. Mangesi L, Hofmeyr GJ. Fetal movement counting for assessment of fetal wellbeing. The Cochrane database of systematic reviews [Internet]. 2007 Jan [cited 2013 Aug 30];(1):CD004909.

346. Carroll SG, Papaioannou S, Nicolaides KH. Assessment of fetal activity and amniotic fluid volume in the prediction of intrauterine infection in preterm prelabor amniorrhexis. American journal of obstetrics and gynecology [Internet]. 1995 May [cited 2013 Aug 30];172(5):1427–35.

347. Gauthier DW, Meyer WJ, Bieniarz A. Biophysical profile as a predictor of amniotic fluid culture results. Obstetrics and gynecology. 1992 Jul;80(1):102–5.

348. Goldstein I, Romero R, Merrill S, Wan M, O'Connor TZ, Mazor M, et al. Fetal body and breathing movements as predictors of intraamniotic infection in preterm premature rupture of membranes. American journal of obstetrics and gynecology. 1988 Aug;159(2):363–8.

349. Vintzileos AM, Campbell WA, Nochimson DJ, Weinbaum PJ, Mirochnick MH, Escoto DT. Fetal biophysical profile versus amniocentesis in predicting infection in preterm premature rupture of the membranes. Obstetrics and gynecology. 1986 Oct;68(4):488–94.

350. Yücel N, Yücel O, Yekeler H. The relationship between umbilical artery Doppler findings, fetal biophysical score and placental inflammation in cases of premature rupture of membranes. Acta obstetricia et gynecologica Scandinavica. 1997 Jul;76(6):532–5.

351. Carlan SJ, Richmond LB, O'Brien WF. Randomized trial of endovaginal ultrasound in preterm premature rupture of membranes. Obstetrics and gynecology [Internet]. 1997 Mar [cited 2013 Aug 30];89(3):458–61.

352. Cotton DB, Gonik B, Bottoms SF. Conservative versus aggressive management of preterm rupture of membranes. A randomized trial of amniocentesis. American journal of perinatology [Internet]. 1984 Jul [cited 2013 Aug 30];1(4):322–4.

353. Lewis DF, Adair CD, Weeks JW, Barrilleaux PS, Edwards MS, Garite TJ. A randomized clinical trial of daily nonstress testing versus biophysical profile in the management of preterm premature rupture of membranes. American journal of obstetrics and gynecology [Internet]. 1999 Dec [cited 2013 Aug 30];181(6):1495–9.

354. Li J, Ma Q, Chen R, Li H, Zhang H. Effect of different pre-labor positions for premature rupture of membranes with vertex and engaged presentation on the maternal and neonatal outcomes. Chinese Journal of Evidence-Based Medicine. 2010;10(12):1415–8.

355. Harper M. Nonstress tests versus biophysical profiles [letter]. American Journal of Obstetrics and Gynecology1. 1999;183(3):777–8.

356. Chamberlain PF, Manning FA, Morrison I, Harman CR, Lange IR. Ultrasound evaluation of amniotic fluid volume. I. The relationship of marginal and decreased amniotic fluid volumes to perinatal outcome. American journal of obstetrics and gynecology. 1984 Oct;150(3):245–9.

357. Snegovskikh V, Park JS, Norwitz ER. Endocrinology of parturition. Endocrinology and metabolism clinics of North America. 2006 Mar;35(1):173–91, viii.

10 APPENDICES

10.1 APPENDIX 1

Table 10.1. A list of computer software applications and online tools used for data analysis.

	Application	Version	Developer/location	Purpose
Bioinformatics tools	ArrayExpress		http://www.ebi.ac.uk/arrayexpress/	An online repository of microarray data.
	Biolayout Express	3.0	http://www.biolayout.org/	Software for drawing and analysing network graphs.
	BioModels		http://www.ebi.ac.uk/biomodels-main/	An online repository of computational models of biological processes.
	Cell Designer	4.2	http://www.celldesigner.org/	A diagram editor for drawing gene-regulatory and biochemical networks.
	Copasi	4.8	http://www.copasi.org/	A software application for simulation and analysis of biochemical networks and their dynamics.
	CVS Bioinformatics Data Mining Tool		http://www.bioinf.mvm.ed.ac.uk/projects/sharp/search.php	An online tool to mine microarray data, built by the CVS Bioinformatics team.
	DAVID (The Database for Annotation, Visualization and Integrated Discovery)	6.7	National Institute of Allergy and Infectious Diseases http://david.abcc.ncifcrf.gov/	A set of functional annotation tools for investigators to understand biological meaning behind large lists of genes.
	Gene Expression Omnibus		http://www.ncbi.nlm.nih.gov/geo/	An online repository of microarray data.
	KEGG Pathway		http://www.genome.jp/kegg/pathway.html	An online collection of manually drawn pathway maps representing knowledge on molecular interaction and reaction networks.
	Reactome		http://www.reactome.org/	An online collection of manually drawn pathway maps representing knowledge on molecular interaction and reaction networks.
	yED	3.11	http://www.yworks.com/en/products_yed_about.html	A software application to produce high quality diagrams.

General use	Review Manager	5.2	http://ims.cochrane.org/revman	Software for preparing and maintaining Cochrane reviews.
	"Safe Haven"		https://informatics.dundee.ac.uk/Citrix/DesktopWeb/auth/login.aspx	Used to access servers at the Health Informatics Centre at the University of Dundee.
	PubMed		http://www.ncbi.nlm.nih.gov/pubmed/	An online repository of abstracts used to search for literature.
	UniProt		http://www.uniprot.org/	An online resource of protein sequence and functional information.
	AxioVision	Rel 4.8	Carl Zeiss Microscopy, Germany	Software for viewing and taking photographs through the PROVIS microscope.
Data and statistical analysis	Area-Proportional Venn Diagram Creator		http://apps.bioinforx.com/bxaf6/tools/app_venndiagram.php	An online tool for creating proportional Venn diagrams.
	GraphPad Prism	6.0	http://www.graphpad.com/scientific-software/prism/	A software application for organising data, performing statistical analyses and drawing scientific graphs.
	Microsoft Excel	2010	Microsoft Office	A spreadsheet application for organising data, performing calculations and drawing graphs.
	R	2.15.1	http://www.r-project.org/	R is a free software programming language and a software environment for statistical computing and graphics.
	SPSS statistics	19.0	http://www-01.ibm.com/software/uk/analytics/spss/	An application for organising data, performing calculations and drawing graphs.
	Venny		http://bioinfogp.cnb.csic.es/tools/venny/	An online tool for creating four-way Venn diagrams.

Table 10.2. PCR results.

Gene	Labouring				Non-labouring				P-value of t-test comparing delta CT values	Fold change - NL to L samples (s.d.)
	Mean raw CT value (s.d.)	Mean delta CT value (s.d.)	Mean cDNA input (s.d.)	Normalised input (s.d.)	Mean raw CT value (s.d.)	Mean delta CT value (s.d.)	Mean cDNA input (s.d.)	Normalised input (s.d.)		
18s	11.79 (5.19)	na	24595.33 (10500.91)	na	10.79 (4.03)	na	26357.22 (10296.49)	na	na	na
IL-8	25.23 (3.36)	15.02 (3.16)	1.02 (1.69)	4.1E-05 (7.1E-05)	29.48 (1.65)	19.47 (1.54)	0.02 (0.02)	7.7E-07 (8.7E-07)	< 0.0001	53.68 (91.99)
IL-6	28.17 (1.90)	17.95 (1.73)	1.49 (1.58)	6.1E-05 (6.9E-05)	30.74 (1.41)	20.73 (1.14)	0.20 (0.20)	7.6E-06 (8.1E-06)	< 0.0001	7.98 (9.12)
MT1E	29.35 (1.88)	19.14 (1.54)	1.83(1.47)	7.4E-05 (6.8E-05)	30.99 (1.08)	20.98 (0.90)	0.43 (0.29)	1.6E-05 (1.3E-05)	< 0.0001	4.57 (4.15)
RBM42	27.47 (1.00)	17.26 (0.46)	0.05 (0.02)	2.2E-06 (1.3E-06)	27.28 (0.72)	17.27 (0.33)	0.06 (0.02)	2.2E-06 (1.1E-06)	0.90	1.01 (0.59)
SHROOM4	27.79 (1.39)	17.57 (0.83)	0.90 (0.58)	3.7E-05 (2.8E-05)	27.64 (0.98)	17.63 (0.63)	0.87 (0.32)	3.3E-05 (1.8E-05)	0.78	1.12 (0.86)
TPM1	21.44 (3.73)	10.44 (0.73)	0.98 (0.55)	4.0E-05 (2.8E-05)	20.59 (0.77)	10.57 (0.44)	0.93 (0.31)	3.5E-05 (1.8E-05)	0.44	1.13 (0.79)
OXTR	22.98 (3.54)	12.15 (2.53)	0.54 (0.46)	2.2E-06 (1.3E-06)	20.94 (2.00)	10.93 (1.97)	1.13 (0.58)	4.3E-05 (2.8E-05)	0.06	-1.97 (0.485)
FABP4	29.24 (1.88)	19.02(1.10)	1.05 (0.89)	4.3E-05 (4.1E-05)	27.98 (1.27)	17.97 (0.95)	2.26 (2.00)	8.6E-05 (8.3E-05)	0.001	-2.01 (0.47)
MYH11	19.56 (3.81)	8.57 (1.05)	0.76 (0.51)	3.1E-05 (2.4E-05)	17.79 (0.77)	7.78 (0.45)	1.18 (0.33)	4.5E-05 (2.2E-05)	0.001	-1.45 (0.55)
IGFBP5	21.80 (3.68)	9.53 (5.74)	0.73 (0.75)	3.0E-05 (3.3E-05)	20.27 (1.00)	9.49 (4.35)	0.89 (0.42)	3.4E-05 (2.1E-05)	0.98	-1.14 (0.99)

10.3 Appendix 3

R Code

Creating a correlation matrix

Compute the correlation between all columns in a dataframe and return a symmetrical matrix of these correlation values.

```
DATA <- read.csv("file path.csv") #read the data into R
CMATRIX <- cor(DATA) #create a correlation matrix using
the Pearson method (matrix is classed as a dataframe)
```

Creating a heatmap based on a correlation matrix

Create a heatmap to visualise a correlation matrix (as in Figure 3.3). The function is part of the gplots package.

```
data.matrix(cmatrix)->cmatrix #make the correlation matrix
object into a matrix
tiff("hm.tif", width=4000, height=3500, res=400) #tell R
how to save heatmap
heatmap.2(cmatrix, Rowv=NA, Colv=NA trace=c("none"),
scale="none", revC=TRUE, density.info="none") #create
heatmap without sorting rows or columns
dev.off() #tell R to save heatmap to the working directory
```

Calculating fold changes and confidence intervals using Fieller's theorem

As described in Chapter 2, Section 2.1.14.10.

```
# x is a vector of expression values for the labouring
group
# y is a vector of expression values for the non-labouring
group
FC <- function(x, y){
x.mu <- mean(x)
y.mu <- mean(y)
FC <- x.mu/y.mu #calculate the fold change (FC)
df <-length(x) + length(y) - 2
t<- -qt(0.05/2, df)
# standard error of x (square root of variance/length)
SEx <- sqrt(var(x)/length(x))
# standard error of y (square root of variance/length)
SEy <- sqrt(var(y)/length(y))
##### If SEy > y.mu, then: ######
g <- (t * SEy/y.mu)^2
SEFC2 <- (FC/(1-g)) * (sqrt((1-g)*((SEx^2)/x.mu^2) +
(SEy^2)/y.mu^2))
lower2 <- (FC/(1-g)) - t * SEFC2
upper2 <- (FC/(1-g)) + t * SEFC2
data.frame(FC, lower2, upper2)
}
##### If SEy < y.mu, then: ######
SEFC1 <- FC * sqrt(((SEx^2)/x.mu^2)+((SEy^2)/y.mu^2))
lower1 <- FC - (t*SEFC1)
```

```
upper1 <- FC + (t*SEFC1)
data.frame(FC, lower1, upper1)
FC(nominator, denominator)
# where 'nominator' is a vector of numbers belonging to one group, and 'denominator' is a vector of numbers belonging to another.
```

MULTIVARIATE LOGISTIC REGRESSION MODELLING

Perform a logistic regression model to assess the adjusted effect of exposure variables on an outcome variable.

```
DATA <- read.csv("file path.csv") #read the data into R
log_model <- glm(OUTCOME ~ EXPOSURE1 + EXPOSURE2 + EXPOSURE3 + EXPOSURE4, data = DATA, family = binomial(logit)) #calculate the logistic model
OR_log_model <- exp(1)^coef(summary(log_model)) #calculate the odds ratios for each exposure variable using the coefficients produced by the model.
lowerCI_OR_log_model <- exp(coef(summary(log_model))[,1] + qnorm(0.025) * coef(summary(log_model))[,2]) #calculate the lower confidence limits of the ORs
upperCI_OR_log_model <- exp(coef(summary(log_model))[,1] + qnorm(0.975) * coef(summary(log_model))[,2]) #calculate the upper confidence limits of the ORs
log_model_OR_results <- cbind(OR_log_model[,1], lowerCI_OR_log_model, upperCI_OR_log_model) #create a table of final results
```

GENERALISED LINEAR MIXED MODELLING (GLMM)

Perform a generalised linear mixed model to assess the effect of exposure variables on an outcome variable, while adjusting for fixed and random effects.

```
DATA <- read.csv("file path.csv") #read the data into R
GLMM <- glmer(outcome~ FIXEDEFFECTEXPOSURE1 + FIXEDEFFECTEXPOSURE2 + FIXEDEFFECTEXPOSURE3 + FIXEDEFFECTEXPOSURE4 + (1|RANDOMEFFECT1) + (1|RANDOMEFFECT2), data = DATA, family = binomial(logit)) #calculate the model
OR_GLMM<- exp(1)^coef(summary(log_model)) #calculate the odds ratios for each exposure variable using the coefficients produced by the model.
lowerCI_OR_GLMM <- exp(coef(summary(GLMM))[,1] + qnorm(0.025) * coef(summary(GLMM))[,2]) #calculate the lower confidence limits of the ORs
upperCI_OR_GLMM <- exp(coef(summary(GLMM))[,1] + qnorm(0.975) * coef(summary(GLMM))[,2]) #calculate the upper confidence limits of the ORs
GLMM_OR_results <- cbind(OR_GLMM[,1], lowerCI_OR_GLMM, upperCI_OR_GLMM) #create a table of final results
```

Calculating the lowest odds ratio an analysis is powered to detect

This calculation returns the lowest odds ratio a Fisher's exact test would be powered to detect. It uses the "clinfun" R package.

```
fe.mdor(ncase, ncontrol, pcontrol, alpha=0.05, power=0.8)
# where ncase is the number of cases, ncontrols is the number of controls, pcontrol is the probability of the outcome in the control group, alpha is the size of the test and power is the power of the test.
```

10.4 APPENDIX 4

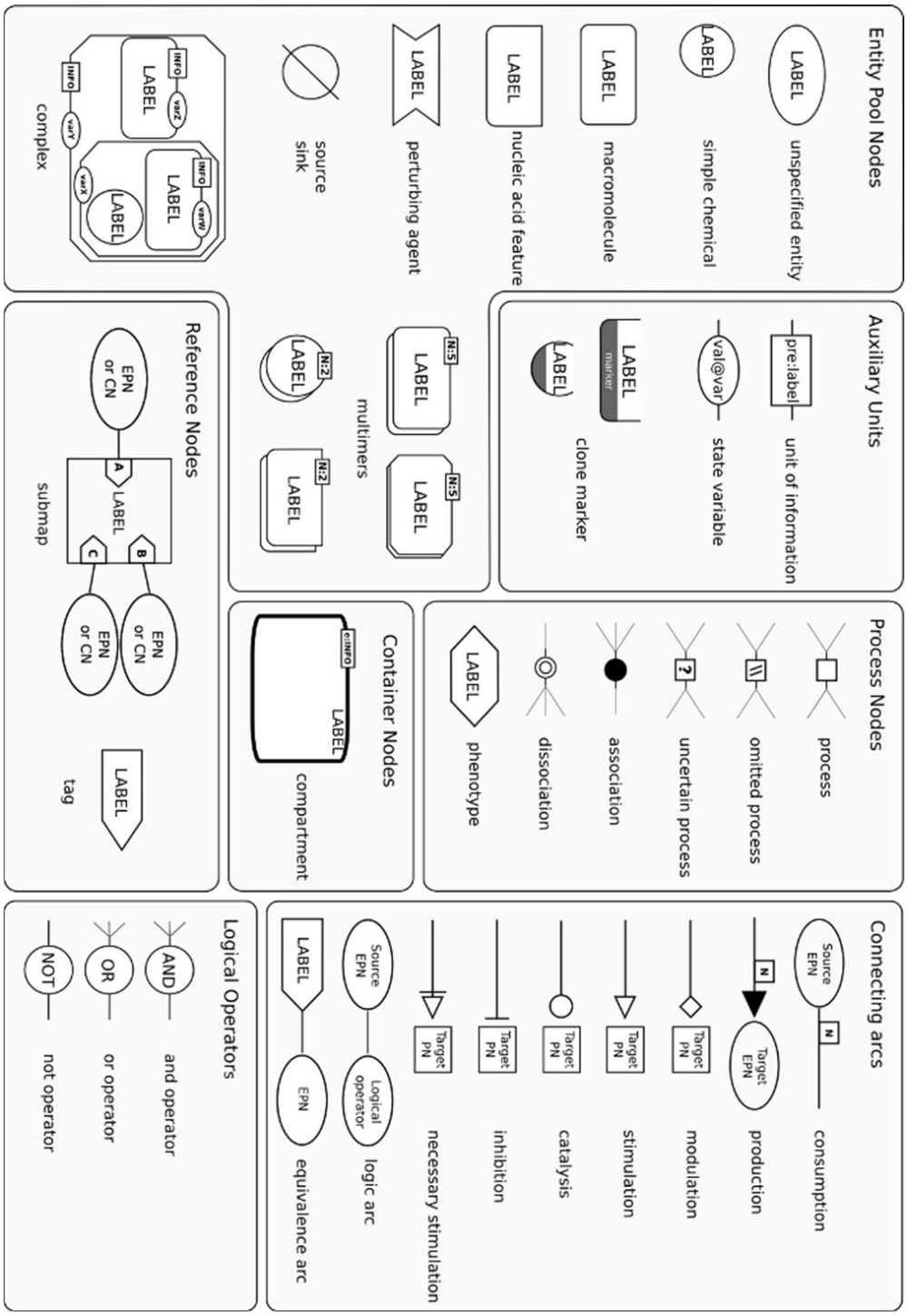


Figure 10.1. A key to Systems Biology Graphical Notation (SBGN).

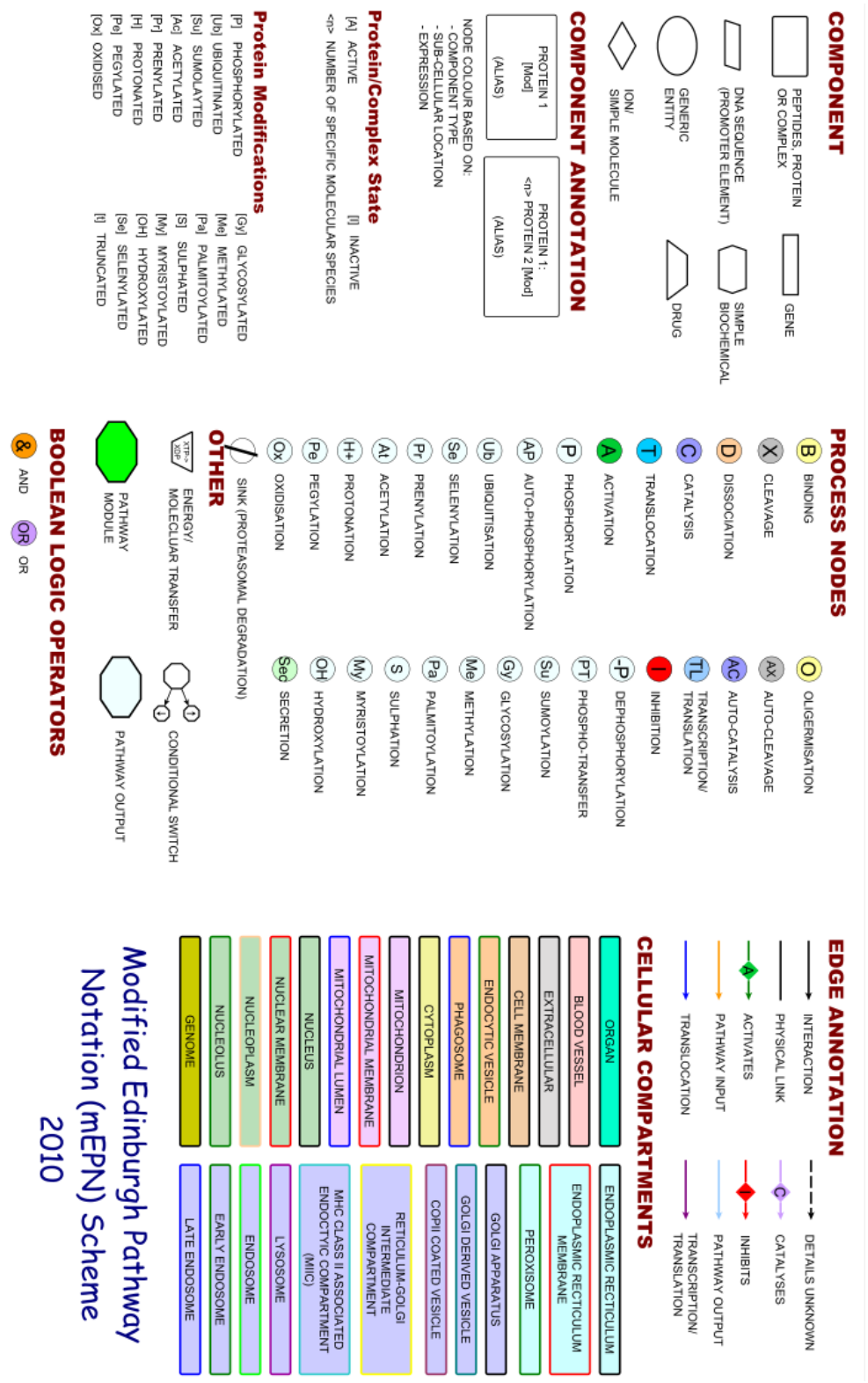


Figure 10.2. A key to modified Edinburgh Pathway Notation (mEPN).

10.4 APPENDIX 5

Table 10.3 Walker variables used in the development of the network graph (Chapter 5). Variables highlighted in red were derived from continuous variables.

Variable name	Description	Percentage of complete records (total = **49585)**
B_C_H_LARGE	Large crown heel length (≥90th centile [≥53.34cm])	48%
B_C_H_CM_SMALL	Small crown heel length (≤10th centile [≤38cm])	48%
B_C_R_LARGE	Large crown rump length (≥90th centile [≥53.34cm])	17%
B_C_R_SMALL	Small crown rump length (≤10th centile [≤28.6cm])	17%
B_EXCH_TRANSF	Fetal exchange transfusion	27%
B_FETAL_DEATH	Fetal death	31%
B_FETAL_DISTRESS	Fetal distress	57%
B_GESTATION_MOD_PT_32-37	Moderately preterm (32-37 weeks – WHO definition)	25%
B_GESTATION_V_PT_28-31	Very preterm (28-31 – WHO definition)	25%
B_GESTATION_E_PT_UNDER28	Extremely preterm (under 28 weeks – WHO definition)	25%
B_GESTATION_POST_OVER42	Postterm (over 42 weeks – WHO definition)	25%
B_HAEMORRHAGIC	Haemorrhagic baby	42%
B_HEAD_CIRCUM_LARGE	Large head circumference (≥90th centile [≥36.8cm])	24%
B_HEAD_CIRCUM_SMALL	Small head circumference (<10th centile [≤32.9cm])	24%
B_MORBIDITY	Neonatal morbidity	31%
B_NND	Neonatal death	31%
B_RH_DISEASE	Rhesus disease	72%
B_SEV_ASPHYXIA	Severe asphyxia (baby)	27%
B_SEX_F	Female baby	99%
B_SEX_M	Male baby	99%
B_SL_ASPHYXIA	Slight asphyxia (baby)	27%
B_WEIGHT_BIRTH_UNDER2.5kg	Birth weight under 2.5kg (WHO definition)	99%
B_WEIGHT_BIRTH_OVER4kg	Birth weight over 4kg (WHO definition)	99%
B_WEIGHT_DISCHARGE_LARGE	Large discharge weight (<10th centile [≥3912.2cm])	18%
B_WEIGHT_DISCHARGE_SMALL	Small discharge weight (>90th centile [≤2749.9cm])	18%
CB_ PLAC_RETAINED_PPH	Retained placenta or postpartum haemorrhage	31%
CB_EMERGENCY	Mother admitted as an emergency	31%
CB_INDUCTION_ARM	Artificial rupture of membranes	31%
CB_INDUCTION_MED	Medical induction of labour	31%
CB_INDUCTION_SURG	Surgical induction of labour	42%
CB_LABOUR_OVER_24HRs	Labour lasting over 24 hours	94%
CB_MOD_CAESAREAN	Delivery via Caesarean section	99%
CB_MOD_FORCEPS_VENT	Delivery using forceps or ventouse	99%
CB_PERINEAL_SUTURE	Perineal suture	42%

CB_PITOCIN_DRIP	Pitocin drip	31%
CB_PLAC_MAN_REMOVAL	Manual removal of placenta	42%
CB_PR_BREECH	Breech presentation	99%
CB_PR_BROW_FACE	Brow face presentation	42%
CB_PR_COMPOUND	Compound presentation	42%
CB_PR_TRANSVERSE	Transverse presentation	42%
CB_TEARS_EPIS	Tears or episiotomy	99%
CB_VERSION_ANT	External version	42%
CB_VERSION_INT	Internal version	42%
CB_VERSION_UA	Version under anaesthetic	31%
F_AGE_YOUNG	Young paternal age (≤20)	49%
F_AGE_OLD	Old paternal age (≥50)	49%
F_HEIGHT_SHORT	Short paternal height (≤1st centile [≤154.94cm])	19%
F_HEIGHT_TALL	Tall paternal height (≥99th centile [≥187.96cm])	19%
F_HSC_1_2	Husband's Social Class - 1 or 2	57%
F_HSC_3	Husband's Social Class - 3	57%
F_HSC_4_5	Husband's Social Class - 4 or 5	57%
M_ABORTION	Maternal history of abortion	42%
M_AGE_UNDER20	Young maternal age (under 20)	99%
M_AGE_35_39	Older maternal age (35 to 39)	99%
M_AGE_OVER40	Very old maternal age (over 40)	99%
M_ALBUMINURIA	Maternal albuminuria	31%
M_ANAEMIA_ANY	Maternal anaemia	99%
M_ANTENATAL_ADMISSION	Antenatal admission	31%
M_CEP_PEL_DISP	Cephalopelvic disproportion	42%
M_ECLAMPSIA	Eclampsia	72%
M_EXCESS_VOMITING	Excess vomiting	72%
M_HAEM_APH	Antepartum haemorrhage	57%
M_HAEM_PPH	Postpartum haemorrhage	42%
M_HEIGHT_SHORT	Short maternal height (<1st centile [≤147.32cm])	52%
M_HEIGHT_TALL	Tall maternal height (>99th centile [≥171.45cm])	52%
M_HISTORY_CS	History of Caesarean section	31%
M_HISTORY_DIFF_LAB	History of difficult labour	57%
M_HISTORY_GYN_OP	History of gynaecological surgery	31%
M_HISTORY_SB_NND	History of stillbirth or neonatal death	57%
M_HYDRAMNIOS	Polyhydramnios	42%
M_HYPERTENSION	Maternal hypertension	72%
M_MAT_DEATH	Maternal death	42%
M_MEMB_RAGGED	Ragged membranes	57%
M_MULTI_PREG	Multiple pregnancy	99%
M_NEPHRITIS	Maternal nephritis	42%
M_OEDEMA	Maternal oedema	31%
M_PARITY_0	Nulliparity	95%

M_PARITY_1plus	Parity ≥1	95%
M_PELVIC_DEFORMITY	Pelvic deformity	42%
M_PLAC_ACCID_HAEM	Placental abruption	42%
M_PLAC_BATTLEDORE	Battledore placenta	57%
M_PLAC_CALCIFIED	Calcified placenta	57%
M_PLAC_CIRCUMVALLATE	Circumvallate placenta	57%
M_PLAC_INCOMPLETE	Incomplete placenta	57%
M_PLAC_INFARCTED	Infarcted placenta	57%
M_PLAC_PRAEVIA	Placenta praevia	68%
M_PLAC_RETROP_CLOT	Retroperineal clot	57%
M_PLACENTA_HEAVY	Heavy placenta (>90th centile [≥830g])	51%
M_PLACENTA_LIGHT	Light placenta (<10th centile [≤481.94g])	51%
M_PRE_ECLAMPSIA	Preeclampsia	42%
M_PROLAPSED_TANGLED_CORD	Prolapsed or tangled umbilical cord	68%
M_PU_BREAST	Puerperal breast complications	27%
M_PU_GENITAL	Puerperal genital complications	27%
M_PU_MORBIDITY	Puerperal morbidity	42%
M_PU_NOTI_PYREXIA	Notification of puerperal pyrexia	42%
M_PU_PPS	Postpartum sterilisation	27%
M_PU_RESPIRATORY	Puerperal maternal respiratory distress	27%
M_PYELITIS	Pyelitis	42%
M_RESP_DISEASE	Maternal respiratory disease	42%
M_RH_SENSITISED	Rhesus disease sensitisation	27%
M_SCIMD_4_5	Scottish Index of Multiple Deprivation - 4 or 5	93%
M_SCIMD_3	Scottish Index of Multiple Deprivation - 3	93%
M_SCSIMD_1_2	Scottish Index of Multiple Deprivation - 1 or 2	93%
M_SEUR6_RURAL	ScottishGovernment 2 fold Urban Rural Classification – Rural (SEUR6 code 5 or 6)	93%
M_SEUR6_URBAN	ScottishGovernment 2 fold Urban Rural Classification – Urban (SEUR6 code 1, 2, 3 or 4)	93%
M_TUMOURS	Uterine tumours	42%
M_UTERINE_DYS	Uterine dysfunction	42%
M_VAGINITIS	Vaginitis	42%

10.5 APPENDIX 6

THE COCHRANE COLLABORATION®

COCHRANE PREGNANCY AND CHILDBIRTH GROUP

Title Registration Form

Email the completed form to Frances Kellie (f.kellie@liverpool.ac.uk) Deputy Managing Editor, or send to the Cochrane Pregnancy and Childbirth Group, Women's and Children's Health, Institute of Translational Medicine, University of Liverpool, First Floor, Liverpool Women's NHS Foundation Trust, Crown Street, Liverpool, L8 7SS, UK. Tel: +44 151 7959571. Fax: +44 151 7959598.

Before completing this form:

- Make sure your proposal falls within the scope of the Pregnancy and Childbirth group, and that it has not already been covered in another Cochrane review published in *The Cochrane Library* (http://www.thecochranelibrary.com/view/0/index.html). Also, check existing registered titles at http://pregnancy.cochrane.org/whats-new.
- Note that all authors must follow the Cochrane Handbook for Systematic Reviews of Interventions (see http://www.cochrane.org/training/cochrane-handbook).
- Be aware that preparing a Cochrane review requires a significant, long-term commitment. At least two authors are required before a title can be registered and the review team must include at least one experienced Cochrane review author.
- Read the 'Notes for authors completing the Title Registration Form' which is available from http://pregnancy.cochrane.org/how-propose-new-cochrane-review .
- *(To activate a yes/no check boxes in this form double click the box and change the default value to 'checked')*

Proposed title (using standard format)

(Include the word 'for' in the title; for example, [intervention] FOR [health problem]; '[Intervention A] versus [intervention B] FOR [health problem]' (*see* Handbook section 4.2.1))

Fetal assessment for preterm prelabour rupture of membranes

Contact person

(This is the author who will be taking responsibility for the development of the proposal and ensuring the continuity of the review once published (*see* Handbook section 4.2.3)

Name: Gemma Sharp

Reason for writing this review

Following preterm prelabour rupture of membranes, accurate methods for predicting fetal wellbeing and prognosis are needed to aid obstetrician's decisions in planning the time and mode of delivery. However, the value of the various methods of fetal assessment in improving neonatal and labour outcomes has yet to be established.

Description of proposal: *(see Handbook chapter 5)*

Background: Preterm prelabour rupture of membranes (pPROM) is a multifactorial disorder and the causes are usually unknown. Gestational age is the main predictor of neonatal outcome, with pPROM at earlier gestational ages being associated with a higher risk of preterm birth and poor neonatal prognosis. Accurate methods to assess prognosis following pPROM are needed to aid obstetricians' decisions in planning the time and mode of delivery. Accurate prognosis is also important for counselling expecting parents, and preparing neonatologists. There are several methods for assessing fetal wellbeing, but the effectiveness of these in improving prognosis is uncertain.

Objective: To systematically review the literature to compare methods of fetal assessment in predicting outcomes following preterm prelabour rupture of membranes.	
Types of study: (section 5.5)	Randomised Control Trials
Participants: (section 5.2)	Women with preterm prelabour rupture of the membranes (PPROM) before 37 weeks' gestation with no specific maternal or fetal contraindications to expectant management (defined by trialists). We will include singleton and multiple pregnancies of any parity, but the effects of these factors will be assessed through subgroup analyses.
Intervention and comparison(s): (section 5.3)	• Fetal MRI for lung volume • Fetal movement counting • Fetal and umbilical artery or venous Doppler ultrasound • Biophysical profile or modified biophysical profile • Fetal cardiotocography • Screening for choriomanionitis • Amniocentesis for fetal lung maturity • Other methods of fetal assessment • Any method of fetal assessment • No fetal assessment.
Outcomes: (section 5.4)	Primary: • Perinatal mortality (deaths between fetal viability (24 weeks' gestation) and the end of the 7th day after delivery, or as defined by trial authors). • Neonatal infection/sepsis (proven neonatal infection with positive blood culture within 48 hours or more after birth, or culture proven neonatal pneumonia or meningitis). • Maternal mortality. • Serious maternal morbidity (including septicaemia, admission to intensive care unit, organ failure, major postpartum haemorrhage, hysterectomy).
	Maternal • Maternal chorioamnionitis (variously defined by authors). • Maternal endometritis (variously defined by authors). • Vaginal birth/caesarean section. • Emergency caesarean section/elective caesarean section. • Induction of labour. • Postpartum maternal pyrexia (variously defined by authors). • Days of antenatal hospitalisation. • Days of postnatal hospitalisation. • Breastfeeding initiated in hospital. • Breastfeeding at hospital discharge. • Maternal satisfaction

	Fetal • Gestational age at birth. • Days from randomisation to birth. • Birth within 48 hours after rupture membranes. • Birth within seven days of rupture of membranes. • Birth before 37 weeks' gestation. • Midtrimester miscarriage. • Intrauterine death after 20 weeks. • Stillbirth (as defined by trial authors). Neonatal • Neonatal mortality only includes deaths in the first 28 days of life. • Postneonatal mortality only includes deaths after 28 days of life but before one year. • Infant death (at 12 months of age). • Respiratory distress syndrome. • Use of surfactant. • Use of mechanical ventilation. • Days of mechanical ventilation. • Days of oxygen therapy. • Oxygen treatment greater than 28 days. • Oxygen therapy at 36 weeks' postmenstrual age. • Birthweight. • Birthweight <10th centile for gestational age • Birthweight less than 2500 grams. • Birthweight less than 1500 grams. • IUGR as defined by triallists. • Admission to neonatal intensive care unit. • Length of stay in neonatal intensive care unit. • Days from birth to discharge home from hospital. • Major cerebral abnormalities on ultrasound prior to discharge. • Necrotising enterocolitis. • Neonatal encephalopathy (as described by authors). • Presumed neonatal infection up to 48 hours of birth. • Presumed neonatal infection 48 hours or more after birth. • Postural deformities (as defined by authors). • Disability at time of childhood follow up (as defined by authors). • Serious disability (as defined by authors) after two years. • Diagnosis of fetal distress in labour (as defined by triallist), • CTG abnormality in labour (as defined by triallists) • Apgar scores <7 at 5 min,

	• Cord pH less than 7.00, • Neonatal encephalopathy, Other • Caregiver satisfaction
Subgroup analyses: (section 9.6)	(Please note that subgroup analyses are usually restricted to the review's primary outcomes). • Singleton/multiple pregnancy • Gestational age at assessment • Gestational age at pPROM
Other information:	

Cross referencing other Cochrane reviews.

Proposals for new Cochrane reviews should not overlap with current publications or work underway. Cochrane reviews should cite other relevant Cochrane reviews (for example, when describing alternative interventions and any effects of the interventions when used for other health conditions or in relation to how the findings of the review compare and contrast with other Cochrane reviews). Please cite here other Cochrane reviews, in progress or published that are relevant to your proposal.

Planned home versus hospital care for preterm prelabour rupture of the membranes (PPROM) prior to 37 weeks' gestation
Antenatal cardiotocography for fetal assessment
Biophysical profile for fetal assessment in high risk pregnancies
Fetal and umbilical Doppler ultrasound in high-risk pregnancies
Fetal and umbilical Doppler ultrasound in normal pregnancy
Fetal movement counting for assessment of fetal wellbeing
Regimens of fetal surveillance for impaired fetal growth
Utero-placental Doppler ultrasound for improving pregnancy outcome

Author's responsibilities

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Cochrane Collaboration policy.

A draft protocol and full review must be submitted to the Pregnancy and Childbirth Group within the timescales you specify on this form. If the drafts are not submitted by the agreed deadlines, or if we are unable to contact you for an extended period, the Pregnancy and Childbirth Group has the right to de-register the title or transfer the title to alternative authors.

The Pregnancy and Childbirth Group has the right to de-register or transfer the title if the draft protocol or review does not meet the standards of the Group and/or The Cochrane Collaboration.

By completing this form, you accept responsibility for maintaining the review in the light of new evidence, comments and criticisms and other developments, and updating the review at least once every two years or, if requested, transferring responsibility for maintaining the review to others as agreed with the Pregnancy and Childbirth Group.

Publication in the Cochrane Database of Systematic Reviews

The support of the CRG in preparing your review is conditional upon your agreement to publish the protocol, finished review and subsequent updates the *Cochrane Database of Systematic Reviews*. By completing this form you undertake to publish this review in the *Cochrane Database of Systematic Reviews* before publishing elsewhere (concurrent publication in other journals may be allowed in certain circumstances with prior permission from the CRG).

I understand the commitment required to undertake a Cochrane review, and agree to publish first in the Cochrane Database of Systematic Reviews.

Signed on behalf of the authors:

Form completed by: **Gemma Sharp** **Date: 19th March 2012**

Do the authors have any potential conflict of interest in this topic? Yes ☐ No ☒

If yes, please give details. Authors should declare any present or past affiliations or other involvement in any organisation or entity with an interest in the review which might lead to a real or perceived conflict of interest. This includes acting as an investigator of a study that might be included in this review. Authors should declare potential conflicts even if they are confident that their judgement is not influenced (see Handbook section 2.6 and www.cochrane.org/docs/commercialsponsorship.htm).

Review context

Is the review subject to any specific funding?	Funding is available from the Alexander McKern Bequest
Has the review already been completed or published?	No

Proposed deadlines

Please specify the dates by which you plan to submit drafts to the Review Group:

Protocol (within 6 months):	July 2012
Review (within 12 months of completion of the protocol):	December 2012

Review authors (see Handbook section 4.2.2.)

Each person named as an author must make a substantial contribution to the conception and design, or analysis and interpretation of the data in the review. Please attach a brief CV for each author.

Contact person / Author 1 (see Handbook section 4.2.3)

(To activate a yes/no check boxes in this form double click the box and change the default value to 'checked')

Prefix (e.g. Ms, Dr):	Ms	First name (名字 míngzi):	Gemma
Middle initial(s):	C	Family name (姓 xìng):	Sharp
Suffix (e.g. MD, PhD):		Web address:	

Preferred full name for review byline: e.g. John Smith = Smith JB; Chen Ming Yu = Chen MY
Sharp GC

Do you already have a user account for the Archie database? Yes ☐ No ☒

Email address:	1) gemma.sharp@ed.ac.uk		
Email address:	2)		
Job Title/Position:	PhD Student		
Department:	MRC Centre for Reproductive Health		
Organisation:	University of Edinburgh		
Street/Address:	QMRI, 47 Little France Crescent		
City:	Edinburgh	Post/Zip code:	EH16 4TJ

State/Province: | Country: UK

Telephone number: 01312426602 | Fax number:

Mobile/cell number: 07837194901

Privacy: As the contact person, your address and email (even if marked as 'hidden') will be published with the completed protocol or review. Your details will be stored on our central database, known as 'Archie', and may be accessed by members of The Cochrane Collaboration. Details of our privacy policy are available at www.cc-ims.net/Archie/archie-privacy-policy.

Would you like to: Hide your address and phone numbers: ☐ Hide your email address: ☐

Country of origin: UK | Gender: Female ☒ Male ☐

Preferred contact: Do you have ready access to email/internet? Yes ☒ No ☐

If no, how you would like us to contact you?

What expertise do you bring to the review? (*e.g. clinical, review methods, statistics*)

Review methods

Have you written a systematic review before? Yes ☒ No ☐

If yes, was it a Cochrane review? (*please state title*)

Yes ☐ No ☒

Are you already a member of another Cochrane Review Group? Which one(s)? Yes ☐ No ☒

What is your first language?

English

At what level are you able to speak and write English?

Fluent, experienced in writing scientific texts.

Note: if your proposal is accepted, we will automatically subscribe you to PCG-info, our Group's email information/discussion list.

Author 2

(To activate a yes/no check boxes in this form double click the box and change the default value to 'checked')

Prefix (e.g. Ms, Dr): Dr | First name (名字 míngzi): Sarah

Middle initial(s): JE | Family name (姓 xìng): Stock

Suffix (e.g. MD, PhD): MRCOG PhD | Web address: Sarah.stock@ed.ac.uk

Preferred full name for review byline: e.g. John Smith = Smith JB; Chen Ming Yu = Chen MY

Stock, SJ

Do you already have a user account for the Archie database? Yes ☒ No ☐

Email address: 1)sarah.stock@ed.ac.uk

Email address: 2)

Job Title/Position: Clinical Lecturer and subspecialty trainee Maternal fetal medicine

Department: MRC Centre for Reproductive Health

Organisation: University of Edinburgh

Street/Address:	Queens Medical Research Institute		
	Little France		
City:	Edinburgh	Post/Zip code:	EH16 4TJ
State/Province:	Lothian	Country:	UK
Telephone number:	+44131 242 2691	Fax number:	
Mobile/cell number:	+44 7894629934		
Privacy:	As the contact person, your address and email (even if marked as 'hidden') will be published with the completed protocol or review. Your details will be stored on our central database, known as 'Archie', and may be accessed by members of The Cochrane Collaboration. Details of our privacy policy are available at www.cc-ims.net/Archie/archie-privacy-policy.		
Would you like to:	Hide your address and phone numbers: ☐	Hide your email address: ☐	
Country of origin:	UK	Gender:	Female ☒ Male ☐
Preferred contact:	Do you have ready access to email/internet?		Yes ☒ No ☐
	If no, how you would like us to contact you?		

What expertise do you bring to the review? (*e.g. clinical, review methods, statistics*)
Clinical expertise

Have you written a systematic review before? Yes ☒ No ☐

If yes, was it a Cochrane review? (*please state title*)

Immediate versus deferred delivery of the preterm baby with suspected fetal compromise for improving outcomes Yes ☒ No ☐

Are you already a member of another Cochrane Review Group? Which one(s)? Yes ☒ No ☐

PCG

What is your first language?

English

At what level are you able to speak and write English?

Tertiart Education

Note: if your proposal is accepted, we will automatically subscribe you to PCG-info, our Group's email information/discussion list.

Roles and responsibilities

Please advise who has agreed to undertake each of the following tasks:

For the protocol:

Draft the protocol	Gemma Sharp
Develop the search strategy for searches (if relevant) in addition to those developed by the Pregnancy and Childbirth Group's Trials Search Co-ordinator.	Gemma Sharp

For the review:

Run the additional searches	Gemma Sharp
Obtain copies of trials	Gemma Sharp
Select which trials to include (2 people)	Gemma Sharp, Sarah Stock
Extract data from trials (2 people)	Gemma Sharp, Sarah Stock
Enter data into RevMan	Gemma Sharp
Carry out the analysis	Gemma Sharp, with help of statistician
Interpret the analysis	Gemma Sharp, with help of statistician
Draft the final review	Gemma Sharp

Experience in systematic reviewing and other information

Have you seen the Cochrane Handbook for Systematic Reviews of Interventions and read the sections mentioned on this form? (*see* www.cochrane-handbook.org) Yes ☒ No ☐

Have you attended a Cochrane Review training workshop? (*see* www.cochrane.org/news/workshops.htm) Yes ☒ No ☐

If yes, which one? Aberdeen (GS),

If no, which one are you planning to attend?

What type of computer do you use? Mac ☐ PC ☒ Linux ☐

Have you downloaded and installed RevMan, the Cochrane review software? (*see* www.cc-ims.net/RevMan) Yes ☒ No ☐

Do you have access to:

The Cochrane Library Yes ☒ No ☐

MEDLINE Yes ☒ No ☐

PubMed Yes ☒ No ☐

EMBASE Yes ☒ No ☐

Do you have access to a medical library? Yes ☒ No ☐

If yes, can you order journal articles not held in the library? Yes ☒ No ☐

Do you have access to advice from a medical librarian? Yes ☒ No ☐

Do you have access to reference management software Yes ☒ No ☐

If yes, which software, and what version? Mendeley 1.3.1

Do you have access to a statistician (strongly recommended)? Yes ☒ No ☐

If yes, whom? expertise at the Wellcome Trust Clinical Research Facility, Edinburgh

Do you have contact with consumer groups relevant to this review? Yes ☐ No ☒

Have you identified appropriate time and resources to complete the review? Yes ☒ No ☐

Experience in systematic reviewing and other information contd.

Currently it is the Pregnancy and Childbirth Group's policy that new Cochrane review teams must have the support of an experienced Cochrane author to help with the process of preparing a Cochrane review. Please

identify who in your review team has direct experience of Cochrane review preparation and describe the extent and nature of their experience.

Name	Experience (e.g. extracting data)
Dr Sarah Stock	Has written review: Immediate versus deferred delivery of the preterm infant with suspected fetal compromise

For office use only

1. **Approved title:**

2. **Approved by:**

   (a) Name:

   Role: Contact Editor **Date approved:**

   (b) Name:

   Role: Co-ordinating Editor **Date approved:**

   (c) Name:

   Role: Deputy Managing Editor/ Managing Editor **Date approved:**

   (c) Name: Lynn Hampson

   Role: Trials Search Co-ordinator **Date approved:**

3. **Date registered in Archie:** **Registered by:**

4. **Notes:**

10.6 APPENDIX 7

The Cochrane Pregnancy and Childbirth Group

Data Extraction Template

Review title:

Review ID:	**Study ID:**	**Reference ID:**
Person extracting data:	**Date of date extraction:**	**Year of study publication:**
Title:		
Author:		
Reference:		

Study design

Type of study design (cluster RCT; block randomisation; stratified randomisation; multi-arm; factorial etc):

Country where the trial was conducted:

Was the trial multi-centre? If so, how many centres were there?

Participants and setting

Describe setting:
Inclusion criteria:
Exclusion criteria:
Total number of randomised participants:
Total available for analysis:
Reason for difference between number of participants and number for analysis:
Information on the age of participants:
Information on parity of participants:
Information on the ethnicity of participants:

Intervention

Experimental intervention:
Any other treatment:

Comparison

Control/Comparison intervention: **Any other treatment:**

Outcomes:

Outcomes:

Study methods
Risk of bias

Adequate sequence generation Was the allocation sequence adequately generated?	Yes / Unclear / No Describe:
Allocation concealment Was allocation concealment adequate?	Yes / Unclear / No Describe:
Blinding Was knowledge of the allocated intervention adequately prevented during the study?	Participant: Yes / Unclear / No Clinician: Yes / Unclear / No Outcome assessor : Yes / Unclear / No Describe:
Incomplete outcome data addressed Were complete outcome data adequately addressed?	Yes / Unclear / No Describe any loss of participants to follow-up at each data collection point: Describe any exclusion of participants after randomisation:
Free of selective reporting bias Are reports of study free of suggestions of selective reporting bias?	Yes / Unclear / No Describe:

Comparison

Control/Comparison intervention: **Any other treatment:**

Outcomes:

Outcomes:

Study methods
Risk of bias

Adequate sequence generation Was the allocation sequence adequately generated?	Yes / Unclear / No Describe:
Allocation concealment Was allocation concealment adequate?	Yes / Unclear / No Describe:
Blinding Was knowledge of the allocated intervention adequately prevented during the study?	Participant: Yes / Unclear / No Clinician: Yes / Unclear / No Outcome assessor : Yes / Unclear / No Describe:
Incomplete outcome data addressed Were complete outcome data adequately addressed?	Yes / Unclear / No Describe any loss of participants to follow-up at each data collection point: Describe any exclusion of participants after randomisation:
Free of selective reporting bias Are reports of study free of suggestions of selective reporting bias?	Yes / Unclear / No Describe:

	Oxygen treatment over 28 days				
	Oxygen therapy at 36+0 weeks				
	BW less than 10th centile for gestational age				
	BW <2500g				
	BW <1500g				
	Admission to neonatal ICU				
	Major cerebral abnormalities on US prior to discharge				
	Necrotising enterocolitis				
	Neonatal encephalopathy				
	Postural deformities				
	Disability at time of childhood follow up				
	Serious disability				
	Diagnosis of fetal distress in labour				
	Cardiotocographic abnormality in labour				
	Apgar scores less than 7 at 5 minutes				
	Cord pH less than 7.00				

	Outcome Measures (Continuous)	Total number of participants in study =					
		Intervention group Total no. in study =			**Control group** Total no. in study =		
		total	Mean (CI)	SD	total	mean (CI)	SD
	Days of antenatal hospitalisation						
	Days of postnatal hospitalisation						
	Total days hospitalisation						
	Maternal satisfaction						
	Gestational age at birth						
	Days from randomisation to birth						
	Days of mechanical ventilation						
	Days of oxygen therapy						
	Birthweight						
	Days from birth to discharge home						
	Caregiver satisfaction						

ADDITIONAL OUTCOMES DESCRIBED IN REPORT BUT NOT IN OUR PROTOCOL:

	Outcome Measures (Dichotomous)	Total number of participants in study =			
		Intervention group total no. in study =		**Control group** Total no. in study =	
		events	Total	events	total

		Total number of participants in study =

	Outcome Measures (Continuous)	**Intervention group** Total no. in study =			**Control group** Total no. in study =		
		total	mean +-CI	SD	total	mean +-CI	SD

Outcomes for sub-group analyses

	Outcome Measures (Dichotomous)	Total number of participants in study =			
		Intervention group total no. in study =		**Control group** Total no. in study =	
		events	Total	events	total
	Primary:				
1					
2					
	Secondary:				
3					
4					
5					

	Outcome Measures (Continuous)	Total number of participants in study =					
		Intervention group Total no. in study =			**Control group** Total no. in study =		
		total	mean	SD	total	mean	SD
	Primary:						
1							
2							
	Secondary:						
3							
4							
5							

General conclusions

Very brief summary of <u>study authors</u> main findings/conclusions:

Exclusion after data extraction

Reasons for exclusion: (study design? participants? interventions/ outcomes? attrition? bias?)

Dates:

Date entered into RevMan and by whom?

Date checked and by whom?

Date copy sent to editorial base and by whom?