Statistical Analysis Plan

Version 4.0

Date: 10th February 2023

Protocol version: 11.0, 31 May 2022

Sponsor: University Hospitals Coventry and Warwickshire NHS Trust
Funding body: National Institute for Health Research – HTA 16/77/02
Ethics Approval Date: 23/03/18

DoH Disclaimer - The views expressed are those of the author and not necessarily those of the NIHR or the Department of Health and Social Care.

FUNDED BY NIHR | National Institute for Health and Care Research

TABLE OF CONTENTS

Abbreviations

AE	Adverse Event
BMI	Body Mass Index
CCD	Comprehensive Cohort Design
CI	Confidence Interval
CONSORT	The Consolidated Standards of Reporting Trials
COVID-19	Coronavirus
CRF	Case report form
DMC	Data Monitoring Committee
ECMO	Extracorporeal Membrane Oxygenation
EFW	Estimated Fetal Weight
EQ-5D-5L	Health-related quality of life
GLM	Generalised Linear Model
GROW	Gestation related optimal weight
HES-ONS	Hospital episode statistics – Office for national statistics
ICIQ-UI	Urinary incontinence short form
iNO	Nitric Oxide
IPD	Individual patient data
IQR	Interquartile range
ITT	Intention to treat
LGA	Large for gestational age
MICE	Multivariate Imputation by Chained Equations
NIHR-HTA	National institute of Health Research – Health technology assessment
NHS	National Health Service
PPI	Patient and Public Involvement
RCOG	Royal College of Obstetricians and Gynaecologists
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard deviation
SOP	Standard operating procedure
SSQ	Six Simple Questions
TMF	Trial master file

UHCW	University Hospitals Coventry and Warwickshire
WCTU	Warwick Clinical Trials Unit
WHO	World Health Organisation

Authorship & Approval

Name	Role	Date	Signature
Katie Booth	Trial Statistician (WCTU)		
George Bouliotis	Senior Trial Statistician (WCTU)		
Siobhan Quenby	Chief Investigator (UHCW)		
Jason Gardosi	Chief investigator (Birmingham Perinatal Institute)		
	Independent Statistician		

Document History

Version	Date	Comment/justification	Timing in Relation to Final Analysis
1.0	18/2/2021	First draft by Miss Katie Booth & prof. Ranjit Lall	Prior (early recruitment stage)
2.0	20/11/2021	Second draft by Dr. George Bouliotis and Miss Katie Booth	Prior (recruitment stage)
3.0	15/11/2022	Third draft by Dr. Bouliotis and Miss Katie Booth	Prior (towards end of the recruitment)
4.0	10/2/2023	Third draft (final) by Dr. Bouliotis and Miss Katie Booth	Prior (towards end of the recruitment)

Study Team

Professor Siobhan Quenby, Professor of Obstetrics, Clinical Sciences Research Lab, University Hospitals Coventry and Warwickshire (UHCW) - Chief Investigator

Professor Jason Gardosi, Director of the Perinatal Institute and Professor of Maternal and Perinatal Health, University of Warwick – Chief Investigator

Professor Martin Underwood, Professor of Primary Care Medicine, WCTU - Warwick Lead PI and Co-applicant

Data Monitoring Committee (DMC) statisticians

Dr Victoria Cornelius, Imperial Clinical Trials Unit, Imperial College London

Trial IT systems & Programmers

Adrian Willis, WCTU, University of Warwick

Role: To generate and prepare reports monitoring the randomisation schedule. To supply data locks for interim and final analysis. Responsibility for randomisation system, clinical databases and related activities.

1 INTRODUCTION

This document details the proposed presentation and analysis for the main paper(s) reporting results from the multicentre randomised controlled trial Big Baby (ISRCTN18229892) to investigate potential benefits and harms of induction of labour in large for gestational age (LGA) fetuses at $38+^{0}$ to $38+^{4}$ weeks gestation.

Suggestions for subsequent analyses by oversight committees, journal editors or referees, will be considered carefully in line with the principles of this analysis plan. This document will serve as the final guidance for all the statistical analysis for this study and will supersede the Statistical Method section in the protocol if there are any discrepancies.

Any deviations from the statistical analysis plan will be described and justified in the final report to the funder. The analysis will be carried out by an identified, appropriately qualified and experienced statisticians, who will ensure the integrity of the data during their processing e.g. by parallel programming.

This SAP was prepared by the trial statisticians in accordance with the WCTU SOPs and the published guidelines related to analysis plans[1 2], after discussion with the Trial Monitoring Committee and it is approved and signed by an independent statistician.

The Big Baby trial is sponsored by the NIHR – Health Technology Assessment (HTA 16/77/02) programme. None of the funding source provided any influence nor controlled any planned analyses. All analyses will be developed and interpreted independently of the study sponsors

2 BACKGROUND INFORMATION

Shoulder dystocia occurs when the fetal head has been born but one, or both, of the shoulders become stuck behind the mother's pubic bone. For the purpose of this study shoulder dystocia is defined as a vaginal cephalic delivery that requires additional obstetric manoeuvres to deliver the fetus after the head has delivered and gentle traction has failed.[1,2] Potential complications for impacted shoulder for the mother include haemorrhage and third and fourth-degree laceration, and for the neonate include fracture of the clavicle or humerus,

temporary or permanent brachial plexus injury, hypoxic ischaemic encephalopathy, and neonatal death.[2]

Most, but not all, cases of shoulder dystocia occur in pregnancies where babies are macrosomic, variously defined as above 4kg, 4.5kg, or >90th weight for gestational age centile. Appropriate management of the condition includes clinical awareness, trained staff and appropriate management protocols and emergency drills. Preventive measures start with antenatal awareness of risk factors, including maternal obesity and diabetes and fetal growth and size.

Earlier delivery should reduce the baby's weight at birth and hence mitigate the main risk factor. However, it is uncertain whether this strategy would actually work, whether shoulder dystocia and its associated complications for mother and baby would actually be reduced, and whether there would be an increase in important side effects such as caesarean sections and associated maternal morbidity. Induction of labour can also be traumatic as it can be associated with prolonged painful labour and may lead to unplanned operative delivery. It is also unknown how many women would accept such a protocol of earlier delivery, or indeed how many would be content to proceed with vaginal delivery rather that requesting caesarean section, once informed about the increased risk of a large for gestational age baby and associated risk.

2.1 Rationale

It is important that a randomized control trial is performed to generate the data needed for women with large babies to make informed choices about their labour onset, likely mode of birth and potential short and longer-term impacts which may be associated with the option selected. This will support the need to explain all potential risks and benefits of management, highlighted by the recent Montgomery judgement,[4] and current maternity policy, which emphasises the importance of involving women in all decisions about their care, to ensure that real 'choice' is truly offered (National Maternity Review England 2016).

2.2 Objectives of the trial

2.2.1 *Primary objective*

To assess the effectiveness of induction at $38+^{0}$ (266 days) to $38+^{4}$ (270 days) weeks gestation in reducing the incidence of shoulder dystocia.

2.2.2 *Secondary objective*

To evaluate whether: induction of labour at 38+0 to 38+4, reduces the risk of neonatal birth injury, increases the risk of infant complications related to prematurity and induction increases the risk of birth injury to the mother.

2.3 Trial design

This is a prospective, multicentre, individually randomised controlled trial, for comparing active treatment (induction of labour at $38+^{0}$ to $38+^{4}$ weeks gestation) against standard care, of fetuses that are estimated to be large for gestational age ($>90^{th}$ customised centile estimated fetal weight (EFW) according to ultrasound at $35+^{0}$ to $38+^{0}$ weeks).

In addition to that, the trial operates a cohort group of those women who decline randomisation but are still interested and consented. The reasoning behind the use of the cohort group is to enhance generalisability of both the baseline data, the primary analysis and the safety analysis. For the purpose of this trial the cohort group consists of two sub-groups: That of the women requesting a planned caesarean section and that one of women without a planned caesarean. The cohort group will be involved in a small number of prespecified exploratory analyses and only as described in detail in this SAP.

2.4 Eligibility

Potential participants are women with fetuses estimated to be LGA at $38+^{0}$ to $38+^{4}$ weeks gestation, who meet the following eligibility criteria:

2.4.1 *Inclusion criteria*

- women aged 18 years or over

- women with a fetus above 90th customised estimated fetal weight centile on ultrasound scan at 35+^{0} to 38+^{0} weeks gestation
- women with a cephalic presentation.

2.4.2 *Exclusion criteria*

- multiple pregnancy
- fetus is in a breech or transverse lie presentation
- induction of labour contra-indicated
- fetus with known serious abnormality
- home birth or elective caesarean section already planned*
- caesarean section or induction indicated due to health conditions such as cardiac disease or hypertensive disorders*
- women taking medications and/or insulin therapy for diabetes or gestational diabetes; women with these conditions who are not taking medication are eligible
- current diagnosis of major psychiatric disorder requiring antipsychotic medication.
- women unable to give informed consent e.g. learning or communication difficulties that prevent understanding of the information provided
- prisoners
- previous stillbirth
- previous neonatal death ≤28 days
- current intrauterine fetal death.

** If the woman is otherwise eligible for the trial and was given the Participant Information Sheet prior to booking a planned caesarean section or induction (for suspected LGA baby), she is eligible to be in the cohort group.*

2.5 Intervention

Women will be randomised to either the booking of induction of labour at 38+^{0} – 38+^{4} (intervention) or standard care (control)

The CONSORT and the flow diagrams are presented in appendix 2. The intervention is the booking of induction of labour at 38+^{0} to 38+^{4} weeks gestational age (266-270 days); method of induction to follow standard practice at participating obstetric unit.

2.6 Outcomes

2.6.1 *Primary outcome*

Incidence of shoulder dystocia, definition by (RCOG) as, '*a vaginal cephalic delivery that requires additional obstetric manoeuvres to deliver the fetus after the head has delivered and gentle traction has failed*'. Shoulder dystocia will be assessed by a notes review, undertaken by an independent expert panel. The complete panel consists of a midwife, a senior obstetrician, a junior obstetrician, and a neonatologist. At least two members of the panel will review each set of delivery notes and indicate the presence or absence of shoulder dystocia as per RCOG definition. The independent panel is unaware of the trial allocation. In cases of discrepancy all members of the panel will be consulted.

2.6.2 *Secondary outcomes*

The trial also plans to follow secondary outcomes, categorised as fetal, maternal and long-term outcomes.

Fetal outcomes

Intrapartum:

- time (in minutes) recorded between delivery of the head and delivery of the body
- time (in days) spent in labour ward (health economic outcome)
- time from commencement of start of active second stage of labour until fetal expulsion
- number of stillbirths

Neonatal outcomes:

- neonatal death
- birth weight
- gestation at birth
- Apgar score at five minutes
- fractures
- brachial plexus injuries
- admission to the neonatal unit, (neonatal intensive care, special care baby unit, high dependency unit, transitional care) / duration of stay
- hypoxic-ischaemic encephalopathy

- need for phototherapy
- respiratory morbidity
- hypoglycaemia
- Clinically defined sepsis/given antiobiotics

Infants:

- proportion under specialist medical care at two months for a problem related to intra-partum experience
- maternal report of infant health concerns at six months
- in-hospital health care costs
- hospital readmission within 30 days of postnatal inpatient discharge.

Maternal outcomes

Intrapartum:

- total duration of hospital stay prior to delivery – (health economics)
- mode of delivery (vaginal, instrumental, planned caesarean section, emergency caesarean section)
- perineal tear (episiotomy and/or spontaneous 1st, 2nd, 3rd or 4th degree perineal tear)
- ~~vaginal tear~~
- cervical laceration or tear
- primary postpartum haemorrhage (≥500ml)
- retained placenta
- death

Post-partum:

- sepsis
- fever (>38.0°c) given antibiotics
- duration of hospital stay after delivery
- uptake of breastfeeding
- hospital readmission within 30 days of postnatal inpatient discharge.

Longer term outcomes:

Women's physical and psychological health and satisfaction with delivery:

- Experience; six simple questions (SSQ) at two months[5]
- Duration of exclusive breast feeding as assessed at two and six months
- Health-related quality of life (EQ-5D-5L) at baseline*, two and six months[6]
- Edinburgh post-natal depression scale at baseline*, two and six months[7]
- Impact of Events Scale two months[8]
- Post-partum bonding questionnaire at two months[9]
- Maternal report of infant health at two and six months
- Urinary incontinence ICIQ-UI short form assessed at baseline*, two and six months[10]
- Faecal incontinence assessed at baseline*, two and six months [11]
- Sexual function at baseline and six months [11]
- Maternal and infant death at six months from HES-ONS linked mortality data. Obtain if the six month follow-up is not completed.
- Participant health resource used for economic analysis for mother and baby at two and six months

Composite outcomes

- Intra-partum birth injury: one or both of fractures of clavicle/long bones of upper extremity or brachial plexus injury in baby.
- Prematurity associated problems: one or both of use of phototherapy or respiratory support.
- Maternal intra-partum complications: one or more of 3rd or 4th degree perineal tear, cervical laceration or tear, or primary postpartum haemorrhage (≥500ml)

2.7 Sample size

2.7.1 *Incidence of the primary outcome*

As the data are not included as part of NHS digital's summary of national maternity statistics, the true incidence of shoulder dystocia in our population of interest is uncertain. The target sample size for this trial is 4,000 participants. This is based on the incidence of "serious shoulder dystocia" in the control arm of the most recent and largest previous trial 16/411 (3.9%). In that trial incidence of shoulder dystocia was defined as: 'difficulty with delivery of the shoulders not resolved by McRobert's manoeuvre,' which is close to our definition of

shoulder dystocia: 'a vaginal cephalic delivery that requires additional obstetric manoeuvres to deliver the fetus after the head has delivered and gentle traction has failed.' The average gestation at randomisation in the Boulvain trial[12] was <38 weeks; we might, therefore expect a slightly higher incidence of shoulder dystocia in our population, where we expect delivery to be at a later gestational age and hence babies will be larger, so we have rounded this to 4%.

To show a 50% reduction at the intervention arm (i.e., to 2%), at alpha significance level 5% with power 90%, requires data on 1,626 women in each arm; 3252 in total. In the Boulvain study,[12] relative risk for "significant shoulder dystocia" was 0.32 (95% CI 0.12 to 0.85). Thus, a 50% reduction is a plausible target that would be considered clinically worthwhile. *Finally, we decided to inflate the sample estimate of 3252 by 23% to 4,000, for a number of reasons explained below*.

In the Boulvain study[12] 7.6% (31/408) of those in the intervention arm went into spontaneous labour prior to induction. This is commensurate with our prediction that 7% of our participants will go into spontaneous labour prior to induction, giving further reassurance that we are seeking a plausible effect size.

We are using a more stringent definition of shoulder dystocia than the composite primary outcome used by Boulvain et al., in their primary analysis and the relevant Cochrane review that reported the incidence of shoulder dystocia to be 6.8% in the control group.

There is considerable uncertainly around our sample size estimate. An allowance is needed for loss to the primary outcome; this should be very small. There may be effects from clustering by site that need to be accounted for; although our analysis of data from the Perinatal Institute indicates that the intra-cluster correlation coefficient for being large for gestational age is <0.00055 suggesting that any effect will be negligible. Most importantly, however, the sample size calculation is very dependent on the baseline rate of shoulder dystocia in our population of interest. For uncommon events such as shoulder dystocia even quite small differences in incidence can have substantial impact on size.

Given the uncertainties around this estimate we performed a key event analysis once we had primary outcome data on 1,000 deliveries. We asked the DMC to advise on whether any sample adjustment was needed, based on the incidence of shoulder dystocia in the control arm.

2.8 Randomisation & Blinding

Randomisation will be provided by WCTU based on a minimisation algorithm and using an on-line web application accessible to all recruiting sites. In case of any issue with the on-line service, a backup telephone service will operate. Minimisation is used for balancing site, estimated fetal weight centile (≤95th EFW centile, >95th EFW centile) and maternal age (≤35 years of age, >35 years of age).

To ensure allocation concealment, randomisation will only take place once all baseline data have been collected. Women will be randomised to either the booking of induction ($38+^{0}$ - $38+^{4}$) or 'standard care' and will be informed immediately of the randomisation outcome. Thus, participants will not be blinded and will be informed of their treatment allocation at randomisation. The trial team will be blinded where possible.

2.9 Assessments

All of the within hospital outcomes will be obtained from routinely collected data in each unit. At baseline, prior to randomisation we will collect routine demographic data; age, ethnicity, parity, height, weight and smoking status. Primary and secondary outcomes will be collected at delivery, 2- and 6-months follow-up. The delivery outcomes assessed by the adjudication panel will be assessed for the primary analysis. Demographic variables and EQ-5D will also be collected at baseline. Details are provided in the appendix 3.

All participants will be asked to complete questionnaires, at two and six months post-delivery, if they have not previously withdrawn. Participants will receive reminders to complete the questionnaires, either by text or email. If participants have not responded to the questionnaires within 6 weeks of the first questionnaire being sent to the participant, efforts will be made to collect a core set of data by telephone. These core data will include:

- Breastfeeding status at two and six months
- Health-related quality of life (EQ-5D-5L) at two and six months [6]
- Maternal report of infant health at two and six months
- Maternal report of her own health at two and six months
- COVID-19 related health question, mother and baby, at two months

Responses from the Two Month Questionnaire will be reviewed to identify babies who have potentially sustained harm relating to a birth injury. We will request relevant data from site for those babies identified and, blind to treatment allocation, an adjudication committee will classify these as delivery related/not delivery related and for those that are delivery related those likely to have a substantial long-term impact and those that are minor or likely to be short lived.

In the event of an unplanned home birth, or birth at another unit, we will collect data from the Ambulance Trust, General Practitioner, or Hospital Trust as appropriate.

In the event of the death of a baby, no questionnaires will be sent to the bereaved family.

For the duration of the COVID-19 pandemic the two- and six- month questionnaires will be collected by telephone, using the reduced dataset detailed above, if it is not possible to send out postal questionnaires.

2.10 Data monitoring

The Data Monitoring Committee (DMC) comprises independent experts in statistics, obstetrics and gynaecology, urogynaecology and paediatrics. They will ensure close monitoring of outcomes during the trial. Analyses of the accumulating data will be presented to the committee who will advise of any excess of adverse events, including shoulder dystocia, which in either group would justify early closure of the study. Frequency of reporting will be at the discretion of the DMC. The trial statistician will attend all DMC meetings and the Co-Chief Investigators and Trial Co-ordinator will attend the open part of the meeting.

2.11 Trial reporting

The trial will be reported in accordance with the Consolidated Standards of Reporting Trials (CONSORT) guidelines (www.consort-statement.org). Authorship of all trial publications will be agreed in accordance with the WCTU SOP 22 ‘Publication and Dissemination’. All

publications will be submitted to the NIHR-HTA Programme for approval prior to submission for publication.

Links to all findings, reports, publications and events will be available via the project website (https://www2.warwick.ac.uk/fac/med/research/hscience/ctu/trials/bigbaby).

3 GENERAL CONSIDERATIONS

3.1 Summaries and Presentation

Tabular summaries will be presented by allocation group. Continuous variables will be summarised by descriptive statistics including mean, standard deviation, standard error, median, maximum and minimum. Categorical variables will be summarised by frequencies and percentage per category. Number of missing values will be reported too. Statistical software will be used for handling and analysis the data for this trial (Stata, SAS).

3.2 Analysis Populations

The intention to treat (ITT) population will be all participants successfully and appropriately randomised, irrespective of treatment received. All analyses will be by ITT at the time of allocation and will also include those incorrectly randomised, if any, as this is a pragmatic trial. Not all women will have a vaginal delivery as planned. We will therefore collect numbers having a caesarean section broken down by type/indication as defined using the Robson score. The compliance population will be women in the intervention group who are induced only between 38^{+0} and 38^{+4}.

Table A: Analysis plan

Intention-To-Treat (ITT) analysis	All those women allocated to either early induction or standard care arm
Per protocol analysis	All those women randomised to early induction and induced at 38+0 to 38+4 days compared to women in the standard care group who had not started labour by 38+0, and were not induced or delivered by elective caesarean section between 38+0 to 38+4 days

3.3 Withdrawals

All withdrawals will be summarised by group using frequencies and percentages. Such summaries will include the level of withdrawal (i.e. complete or partial), the timing (i.e. after randomisation or after time point 2) and the reason for withdrawal (i.e. participant's or clinician's decision).

3.4 Follow-up

Follow-up rates will be calculated at each of the follow-up time points as the number of participants assessed at time *T* out of the total number that should have been assessed at time *T*.

3.5 Adherence to protocol

A protocol non-compliance is defined as a failure to adhere to the protocol. All known protocol non-compliances (deviations and violations) will be listed in the final report, by intervention arm.

4 STATISTICAL ANALYSES AND SUMMARIES

4.1 Participation data

The flow of participants through the trial will be illustrated using a CONSORT diagram (Appendix 2). This will include:

- Number screened.
- Of those screened, how many women were ineligible or declined or missed due to non-availability of research staff.
- Number of women randomised and allocated to each intervention.
- Number of women/babies withdrawn, died, lost to follow-up at 2 and 6 months.
- Number included in the final analyses through the flow of the study - randomisation to end of follow-up.
- Number excluded from analysis - withdrawn, died and lost to follow-up.

4.2 Screening data

Screening information will be presented as frequencies and percentages. The screening summary will include information for both randomised and non-randomised (cohort) participants.

4.3 Baseline Comparisons

Baseline characteristics will be summarised and presented for each intervention arm at randomisation and baseline, with no formal testing (see Appendix 4). For continuous data, the number of participants (n), mean, standard deviation (sd), median and interquartile range (IQR) will be used appropriately to summarise characteristics by treatment allocation. The number (%) of participants will be used to summarise categorical outcome measures.

4.4 Efficacy Comparisons - Primary Outcome

4.4.1 *Primary analysis for the primary outcome – Estimand E1*

The primary analysis will be based on the assessment of the event of interest between intervention and control, here incidence of shoulder dystocia, measured as presence or absence (binary outcome). This is the primary estimand (E1) and is going to include only adjudicated dystocias (appendix 5).

(Estimand Strategy) The primary estimand (E1) is intended to provide a reliable estimate of the treatment effect, here the early induction, within an intention-to-treat context.

(Analysis set) The ITT set (see Table A, Section 3.2) will contain all those women randomised and successfully allocated to a trial arm, including all protocol violations (i.e. wrongly randomised, induced not per protocol).

(Analysis Methodology) The difference in shoulder dystocia between the two arms will be provided as risk ratio/odds ratios for the intervention by using a GLM (binomial/logistic) model to adjust for covariates, including the minimisation variables (site, maternal age [≤35 years, >35 years] and estimated fetal weight centile [≤95th EFW centile. >95th EFW centile]) and clinical factors (i.e., diet controlled gestational diabetes), in case of an obvious imbalance. The model will provide risk ratios/odds ratio and the corresponding 95% confidence intervals[3], to ease (clinical) interpretation. A fixed effect regression model (based on binary data) will be fitted to the primary outcome using robust standard errors (Huber - White), to account for site-effect. Such a flexible model can be used consistently across the (binary) study outcomes (primary and secondary). The model that is going to be used will be of this form:

ln[p/1-p] = β0 + β1[treatment] + β2[site] + β3 [maternal age] + β4 [estimated fetal weight centile] + β5[presence of diabetes]

(Intercurrent Events) Three types of intercurrent events would be expected to a minimal extent: (1) very few non-eligible randomised women due to incorrect readings of the estimated fetal weight centile, and using the ultrasound scan report prior to 35+0 days (2) change of the planned induction timing for both arms due to emergencies (e.g., no fetal movement) and (3) few cases of missing data due to early dropouts for patient reasons (e.g. patient moved to another area/hospital). As this is a pragmatic trial, the above-mentioned protocol violations will be included in the primary analysis (ITT).

(Population Level Summary) The estimate will be summarised as a risk ratio/odds ratio (unadjusted and adjusted) for the early induction group with its 95% confidence interval.

4.4.2 *Secondary analysis for the primary outcome – Estimand E2*

The above-mentioned analysis will be repeated for deriving an estimate for the second estimand (E2). In this case the analysis set will be different as the protocol violations will be

excluded. This is an alternative per protocol analysis approach (see Table A, Section 3.2), contributing as a sensitivity analysis for the primary endpoint only. Apart from the analysis sets, all remaining elements of this estimand will be the same as for estimand E1.

4.4.3 *Sensitivity analysis for the primary outcome - Estimand E3*

As a final sensitivity analysis, and only in case of more than 10% missingness in the model covariates, we are going to repeat the primary analysis having imputed the missing values. Multiple imputation methods will be based on the MICE algorithm[4 5 6]. In addition to that, we will explore departures from missing at random for the incomplete outcome data using the mean score method[9].

4.5 Efficacy Comparison -Secondary Outcomes

Secondary binary outcomes will be analysed in a similar way to the primary outcome and risk ratios with corresponding confidence intervals, adjusted and unadjusted, will be reported (Table 11, Appendix 4). Continuous outcomes will be analysed using linear regression models and mean differences will be reported, using the same adjusting covariates as in the primary analysis. If required, sensitivity analyses will be computed (for example, assessment of missing data using multiple imputation).

4.6 Exploratory Comparison: Pre-specified subgroup analyses

We will conduct a pre-planned conventional subgroup analysis using an interaction term for two key variables: maternal body mass index and estimated fetal weight centile. At this stage it is unknown to us whether enough events at the end of this trial will permit analysis using pre-specified groups (i.e, BMI normal, overweight, obese). We will adapt the methodology in line with the fraction of data available. Significance level of 0.10 will be used for interpreting.

4.7 Safety Comparison

We plan to compare, by arm, the number of safety events (SAEs and AEs) recorded per participant in total, then as clinical groups and as severity above. P-values from the statistical tests for counts will be provided. Mother safety and the baby safety data will be analysed separately. The definitions of SAE were changed after the start of the trial, however all SAEs that have been reported, regardless of definition change, will be reported.

4.8 Cohort group – non-randomised volunteers

The trial team is aware of the fact that some interested and eligible women may refuse to be randomised as due to treatment preferences. Our plan is to encourage them to consent to participation as non-randomised participants after having met the same enrolment criteria. The data collection process will be identical for those participants, however as they are not randomised, only limited analyses will be undertaken with them, which will be cautiously approached, and interpreted. Thus, the study adopts the Parallel Cohort design (sometimes known as Comprehensive Cohort Design (CCD)) and allows participants with certain treatment preferences to still provide data for the study (i.e. planned caesarean section). It is essential for those volunteers to provide full consent and be followed up the same as the randomised participants. For maximising the benefits of this parallel cohort design, we separated the cohort to those women requesting an elective caesarean section and those otherwise. A summary table (Table 21, Appendix 4) will provide comparisons of the baseline characteristics between the (all) randomised and the separated cohorts, highlighting statistically significant differences, if any.

By running parallel to the randomised arms, the cohort will allow a useful comparison of the baseline characteristics between those eligible and randomised and those eligible but not randomised, providing an opportunity to detect selection bias, if any. The reasoning behind this approach is to check for generalisability (external validity). For this particular baseline comparison, p-values will be displayed. However, we are aware that the comparison is based on an untenable and untested assumption that the non-randomised sample is representative of the true population. Also, no formal power calculations had been developed for determining the size of this (cohort) group prior to recruitment. For those reason, we will not interpret/comment on the results, nor will we undertake any outcome comparisons.

Furthermore, the trial team has a specific interest on (limited) number of outcomes in the planned caesarean section cohort group as compared against the intervention. Briefly, this will include cost-effectiveness, quality of life and safety comparisons (i.e. number of adverse events).

Finally, no other analyses are going to include the parallel cohort.

4.9 Significance levels and adjustment of p-values for multiplicity

All statistical tests will be two-sided at the 5% significance level with the exception of subgroup analyses where 10% significance will be used. The estimate with the corresponding 95% confidence interval and the p-value will be reported for each test undertaken. No adjustments for multiplicity will be made

4.10 Trial success and the role of the PPI

The success of the trial will be judged primarily on the basis of statistical significance at alpha 5% that was used for the sample size calculations at the design stage. However, in the case positive but not strong enough effect size failing to reach statistical significance, we will explore clinical significance by combining the risks and the benefits of the intervention in a meaningful way[7]. The PPI members for this study will have an important role in "weighting" the importance of the risks and benefits from the intervention. Various scenarios will be discussed and when possible clinical significance will be assessed. In this way, the chance of a misleading conclusion based on "noisy" p-values will be avoided.

4.11 Statistical Software

All analyses and reporting will be conducted using validated commercial statistical software available within the unit (either Stata or SAS or R, depending on the preference of the trial statistician).

5 REFERENCES

1. DeMets, D.L., T.D. Cook, and K.A. Buhr, Guidelines for Statistical Analysis Plans. JAMA, 2017. 318(23): p. 2301-2303.

2. Gamble, C., et al., Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA, 2017. 318(23): p. 2337-2343.

3. McCullagh, P. and J.A. Nelder, Generalized linear models. Monographs on Statistics and Applied Probability 37. 1989.

4. Carpenter, J.R. and M.G. Kenward, Missing data in randomised controlled trials: a practical guide. 2007.

5. Wood, A.M., I.R. White, and P. Royston, How should variable selection be performed with multiply imputed data? Stat Med, 2008. 27(17): p. 3227-46.

6. White, I.R., P. Royston, and A.M. Wood, Multiple imputation using chained equations: Issues and guidance for practice. Stat Med, 2011. 30(4): p. 377-99.

7. Shakespeare, T.P., et al., Improving interpretation of clinical studies by use of confidence levels, clinical significance curves, and risk-benefit contours. Lancet, 2001. 357(9265): p. 1349-53.

7. Calis, K.A., et al., Recommendations for data monitoring committees from the Clinical Trials Transformation Initiative. Clin Trials, 2017. 14(4): p. 342-348.

8. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol 2004; 159(7):702-6.

9. White, I., Carpenter, J. and Horton, N. (2018). A mean score method for sensitivity analysis to departures from the missing at random assumption in randomised trials. *Statistica Sinica*.

6 APPENDIX

Appendix 1: Outcome Definitions

Outcome	Time point	Derivation of outcome
Primary outcome		
Shoulder dystocia	2	The incidence of shoulder dystocia is a categorical outcome of 'Yes' or 'No', defined by the Royal College of Obstetricians and Gynaecologists (RCOG) as '*a vaginal cephalic delivery that requires additional obstetric manoeuvres to deliver the fetus after the head has delivered and gentle traction has failed'.* [1,2] Shoulder dystocia will be confirmed by a notes review, undertaken by an independent expert panel.
Secondary outcomes – Fetal		
Time recorded between delivery of the head and delivery of the body	2	Calculated by; 'Date and time of head delivered' - 'Date and time of body delivered', which are both variables collected in the Delivery section of the CRF. This variable only applies to woman who delivered vaginally.
Time in labour ward	2	Calculated by; 'Date and time woman left the labour ward' – 'Date and time admitted to labour ward', which are both variables collected in the Delivery section of the CRF. This variable applies to all women who delivered in hospital.
Time from commencement of active second stage of labour until fetal expulsion	2	Calculated by; 'Date and time body delivered' – 'Date and time active second stage of labour commenced', for vaginal deliveries. Or; 'Date and time body delivered' – 'Date and time of delivery', for Caesarean sections. These variables are collected in the Delivery section of the CRF and apply to all women.
Stillbirths	2	
Neonatal death	4	Has the baby died since delivery (not including stillbirths)
Birth weight (g)	4	
Gestation at birth (Weeks + Days)	4	
Apgar score at five minutes	4	
Fractures	4	Humeral or clavicular fracture
Brachial plexus injuries	4	
Admission to the neonatal unit/duration of stay	4	Admission to neonatal unit: Whether baby received any additional care, either at the same or a different hospital following delivery Duration of stay at neonatal unit: Calculated by; 'Date of discharge – 'Date of admission', for both same and different hospital transfers.
Hypoxic-ischaemic encephalopathy (HIE)	4	

Use of phototherapy	4	
Respiratory morbidity	4	Including the baby receiving the following; Supplemental O_2, Mechanical ventilation, Non-invasive respiratory support, Extracorporeal Membrane Oxygenation (ECMO), or Nitric oxide (iNO) therapy.
Hypoglycaemia	4	A single value of <2.6 mmol/L
Proportion under specialist medical care at two months for a problem related to intra-partum experience	5	
Maternal report of infant health concerns at six months	6	Including: - Has baby been diagnosed with a palsy - Noticed baby having difficulties with arm or finger movement - Does baby's face droop on one side - Is baby receiving ongoing specialist medical care (details) - Other concerns about baby's health and development
Hospital readmission within 30 days post-natal inpatient discharge	4	
Secondary outcomes – Maternal		
Duration of hospital stay prior to delivery (days)	2	Calculated by; 'Date and time body delivered' – 'Date and time admitted to hospital' if vaginal delivery Or; 'Date and time of delivery – 'Date and time admitted to hospital' if Caesarean section
Mode of delivery	2	Including; Vaginal delivery, Instrumental delivery (Ventouse), Instrumental delivery (Forceps), Instrumental delivery (Rotational forceps), Planned caesarean section and Emergency caesarean section.
Perineal tear	2	Episiotomy or spontaneous 1st to 4th degree perineal tear
Cervical laceration or tear	2	
Primary postpartum haemorrhage	2	≥500ml
Retained placenta	2	
Death	2	Death of woman during complications of labour
Sepsis	2, 3	During labour or within 24 hours post-partum
Fever	2, 3	Fever >38.0°c in labour or within 24 hours post-partum
Duration of hospital stay after delivery (days)	3	Calculated by; 'Date and time discharged from hospital' – 'Date and time body delivered', if vaginal delivery Or; 'Date and time discharged from hospital' – 'Date and time of delivery', if Caesarean section
Uptake of breastfeeding	3	Including; - Did woman and baby have skin-to-skin contact immediately following birth - Did woman start breastfeeding

		- Was woman still breastfeeding at discharge from the ward (if gave birth in hospital) - Is breastfeeding exclusive or partial (is breastfeeding at discharge)
Hospital readmission within 30 days of postnatal inpatient discharge	3	
Secondary outcomes – Longer term; Women' physical and psychological health and satisfaction with delivery		
Experience; Six simple questions (SSQ) at two months [5]	5	The SSQ includes 6 statements relating to the care received during pregnancy, where each statement can be scored from one (Strongly disagree) to seven (Strongly agree), depending on whether you agree or disagree with the statements. Each of the six statements will be summarized individually.
Duration of exclusive and partial breast feeding as assessed at two and six months	5, 6	Exclusive breastfeeding: Only giving the baby breast milk. Partial breastfeeding: Giving baby breast milk and formula. This is captured at post-partum, two and six months post-delivery.
Health-related quality of life (EQ-5D-5L) at baseline, two and six months [6]	1, 5, 6	The EQ-5D-5L score ranges from <0-1, where a higher score reflects a better quality of life.
Edinburgh Post-natal Depression Scale (EPDS) at baseline, two and six months [7]	1, 5, 6	The EPDS is a 10-item questionnaire, with each item ranging from 0 to 3, with higher score indicating more severe symptoms. Each of the 10 item scores are combined to give an overall score, ranging from 0 to 30, with higher scores indicating more severe symptoms.
Impact of Events scale (IES) two months [8]	5	The IES consists of 15 questions, to measure the amount of distress that is associated with a specific event. It has possible answers of: Not at all (0), Rarely (1), Sometimes (3), and Often (5). The 15 items are totalled to give a score ranging from 0-75. Scores ranging from 0-8 indicate no meaningful impact, 9-25 indicates an impact event, 26-43 indicates a powerful impact event and a score range of 44-75 indicates a severe impact event.
Post-partum bonding questionnaire at two months [9]	5	The post-partum bonding questionnaire consists if 8 questions, ranging from 0-3. Scores range from 0-24, with a higher score indicating a worse level of bonding.
Maternal report of infant health at two and six months	5, 6	Report of infant health includes; - Diagnosis of a palsy - Difficulties with arm or finger movement - Face drooping on one side - Ongoing specialist medical care - Other concerns about baby's health and development
Urinary incontinence ICIQ-UI short form assessed at baseline, two and six months [10]	1, 5, 6	The ICIQ-UI form consists of 4 questions, with questions 1-3 combining to form a score ranging from 0-21, with higher score indicating worse incontinence.
Faecal incontinence assessed at baseline, two and six months [11]	1, 5, 6	The faecal incontinence outcome consists of 5 questions relating to bowel function of the mother. These questions do not give a score.
Sexual function at baseline and six months [11]	1, 6	Sexual intercourse score ranges from 0 (not applicable) to 4 (Always) with higher scores indicating a higher frequency of pain.
Maternal and infant death at six months from HES-ONS linked mortality data	6	

1 - Baseline, 2 - Intra Partum, 3 - Post Partum Mother, 4 - Post Partum baby, 5 - Month 2 Follow-up, 6 - Month 6 Follow-up

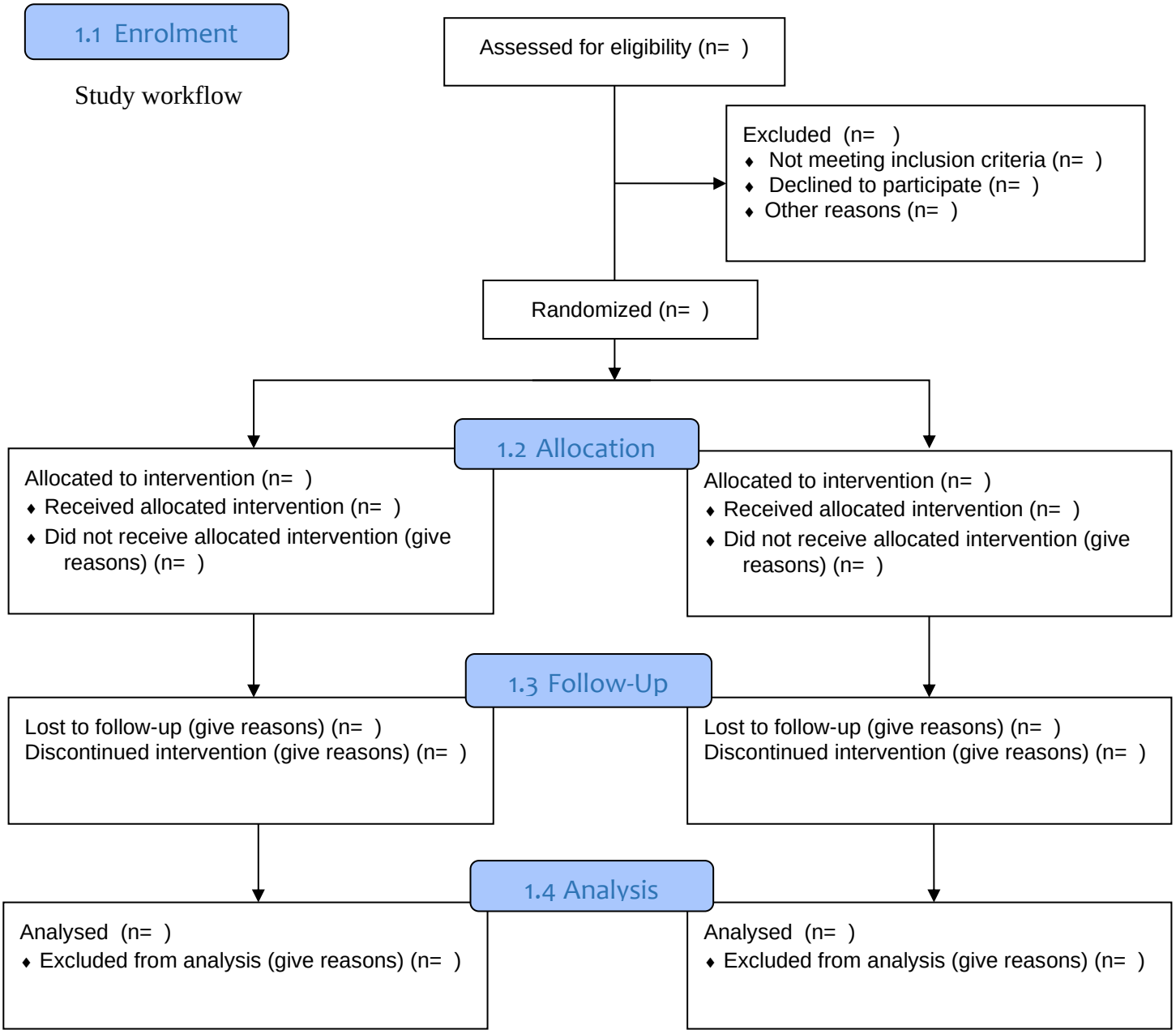
Appendix 2: Diagrams

CONSORT diagram

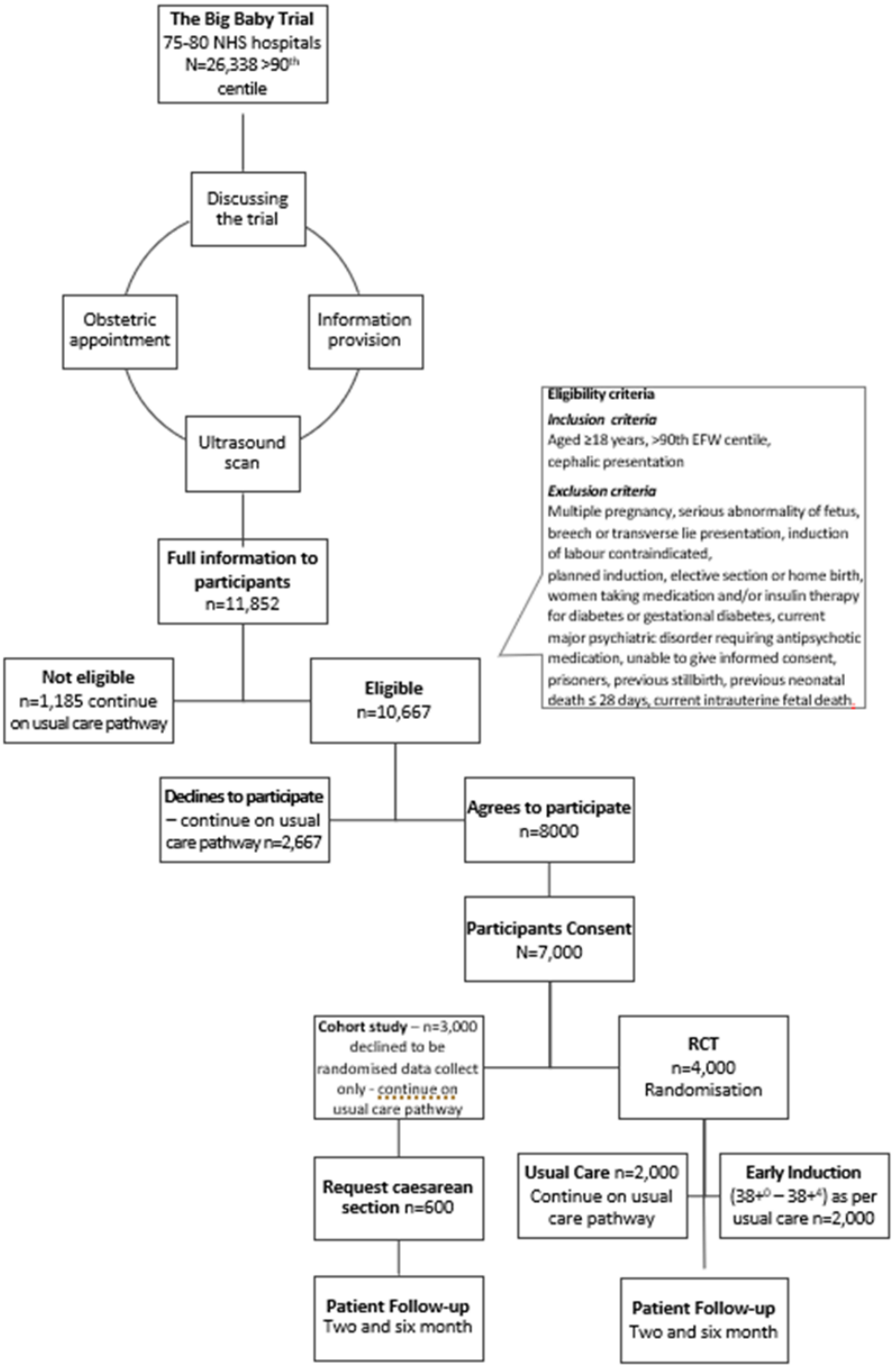
CONSORT
TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram

Study workflow



Study Flow Diagram



Appendix 3: Summary of outcome measures and delivery time points

Type of Data	Outcome measures	Time points					
		1	2	3	4	5	6
Demographic	Date of Birth, ethnic group, height, weight	X					
GROW Chart	Date of dating scan, estimated delivery date, gestation at dating scan, parity, date of GROW scan (35+0-38+0), gestation at GROW scan, estimated fetal weight, GROW centile	X					
Medical History	The current obstetric status and medical status will be collected, along with data on tobacco use, alcohol use and corticosteroid use. Previous pregnancy data will be collected.	X					
Delivery	Date of delivery, onset of labour, and details of the induction, delivery and shoulder dystocia. Details include: Time recorded between delivery of the head and delivery of the body, time in labour ward, time from commencement of active second stage of labour until fetal expulsion, duration of hospital stay prior to delivery, mode of delivery, primary postpartum haemorrhage (≥500ml) Medication used for pain relief and antibiotics during delivery will be documented		x				
Post-delivery Maternal	Details on maternal perineal injury, complications of labour, breast feeding, immediate post-delivery hospitalisation and discharge home will be collected. Details include: perineal tear (episiotomy or spontaneous 1st to 4th degree perineal tear), cervical laceration or tear, duration of hospital stay prior to delivery, retained placenta, death,			x			

	sepsis, fever (>38.0°c), duration of hospital stay after delivery, uptake of breastfeeding						
Post-delivery Baby	Details on the baby's immediate birth outcome, birth injuries, immediate post-delivery hospitalisation, neonatal procedures and conditions, respiratory support and discharge will be collected. Details include: Stillbirths, neonatal death, birth weight, gestation at birth, Apgar score at five minutes, fractures, brachial plexus injuries, admission to the neonatal unit / duration of stay, hypoxic-ischaemic encephalopathy, use of phototherapy, respiratory morbidity, hypoglycaemia.						
Adverse Events	Event, start date, stop date, severity and outcome will be collected	x	x	x	x		
Medication	Medication name, start date, stop date, dose, unit, route and indication will be collected	x	x	x	x		
Unscheduled visits	Details of any known, unscheduled hospital visits by the mother	x					
Unscheduled visits	Details of any known, unscheduled hospital visits by the mother or baby after discharge (after discharge up to 30 days post discharge. This will not be collected for cohort participant not requesting an elective C-Section)					x	
Questionnaire	Mother's report of infant health, Mother's report of her health, breastfeeding. Details include: proportion of babies under specialist medical care at two months for a problem related to intra-partum experience maternal report of infant health concerns at six months in hospital health care costs					x	x

Questionnaire	EQ-5D-5L, Edinburgh Post-Natal Depression Score, urinary incontinence - ICIQ-UI, faecal incontinence,	x				x	x
Questionnaire	Birth Experience; six simple questions, Post-Partum Bonding questionnaire, Impact of Events Scale					x	

1 . Baseline
2 . Intra Partum
3. Post Partum Mother
4 . Post Partum baby
5 . Month 2 Follow-up
6. Month 6 Follow-up

Appendix 4: Analysis Tables (by arm)

Figure 1: CONSORT Diagram for Big Baby trial

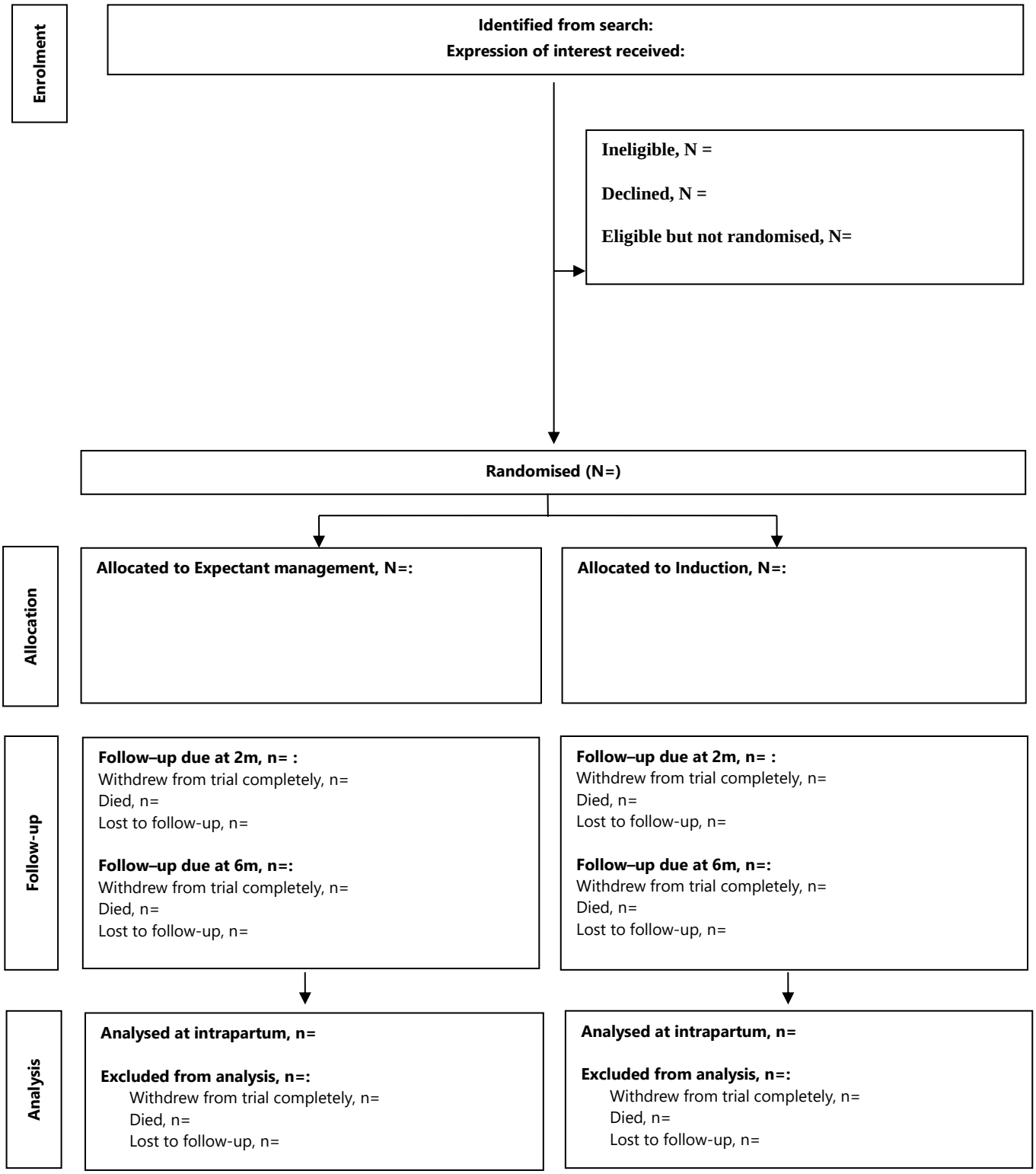


Table 1: Pre-randomisation (screening log) summary

	Total N (%)
Eligible	N
Consent sought at 35^{+0} to 38^{+0} weeks	N (%)
Ineligible (of which):	N
Woman under 18 years old	N (%)
Fetus <90th EFW customised centile on ultrasound at 35^{+0}-38^{+0}	N (%)
Not cephalic	N (%)
Multiple pregnancy	N (%)
Induction of labour contra-indicated	N (%)
Fetus with known serious abnormality	N (%)
Home birth or elective caesarean section already planned	N (%)
Caesarean section or induction already indicated for other reasons	N (%)
Taking oral medication and/or insulin for diabetes and gestational diabetes	N (%)
Major psychiatric disorder requiring antipsychotic medication	N (%)
Unable to give informed consent	N (%)
Prisoner	N (%)
Previous stillbirth	N (%)
Previous neonatal death	N (%)
Current intrauterine fetal death	N (%)
Delivered prior to randomisation	N (%)
No obstetric consent obtained	N (%)
Other	N (%)

% in the eligible categories are of the total eligibles;
% in the ineligible categories are of the total ineligibles

Table 2: Screening and Randomisation

Screening	**Total N=**
Number screened (N)	
Eligible (N)	
Enrolled (N)	
Non-Randomised (N)	
Randomised (N)	
Randomisations per quarter	
Recruitment Rate	

Table 3: Demographic characteristics and completeness of randomised participants

Recruitment demographics		**Expectant management N=**	**Induction N=**	**Total N=**
Maternal age at recruitment	≤35 years			
	>35 years			
	Missing n (%)			
Maternal age at recruitment (years)	Mean (s.d)			
	Median (IQR)			
	minimum-maximum			
	Missing n (%)			
Eligibility Check Demographics		**Expectant management N=**	**Induction N=**	**Total N=**
Parity (≥24 weeks)	0 - n (%)			
	1 - n (%)			
	2 - n (%)			
	3+ - n (%)			
	Missing – n (%)			
GROW Centile	90.1 to 91.0 – n (%)			
	91.1 to 92.0 – n (%)			
	92.1 to 93.0 – n (%)			
	93.1 to 94.0 – n (%)			
	94.1 to 95.0 – n (%)			
	95.1 to 96.0 – n (%)			
	96.1 to 97.0 – n (%)			
	97.1 to 98.0 – n (%)			
	98.1 to 99.0 – n (%)			
	99.1 to 100.0 – n (%)			
	Missing – n (%)			

Table 4: Baseline data and completeness of randomised participants

Baseline variables		**Expectant management N=**	**Induction N=**	**Total N=**
Obstetric Status:				
Gestational diabetes mellitus (diet controlled only)	No n (%)			
	Yes n (%)			
	Missing n (%)			
Other obstetric status	No n (%)			
	Yes n (%)			
	Missing n (%)			
Tobacco use:				
Was woman a smoker at booking visit	No n (%)			
	Yes n (%)			
	Missing n (%)			

Corticosteroids:				
Received antenatal Corticosteroid course for fetal lung maturation during pregnancy	No n (%)			
	Yes n (%)			
	Missing n (%)			

Table 5: Previous pregnancy data and completeness of randomised participants

Previous pregnancy information		**Expectant management N=**	**Induction N=**	**Total N=**
Number of women who had previous pregnancies				
Number of previous pregnancies				
Multiple births[1]	No n (%)			
	Yes n (%)			
	Missing n (%)			
Number of previous births[2]				
Mode of delivery [3]	Vaginal n (%)			
	Assisted vaginal n (%)			
	Caesarean n (%)			
	Missing n (%)			
Shoulder dystocia	No n (%)			
	Yes n (%)			
	Missing n (%)			
Brachial plexus injury	No n (%)			
	Yes n (%)			
	Missing n (%)			
Gestation at delivery (wk+d)	Mean (s.d)			
	Median (IQR)			
	Missing n (%)			
Birthweight (g)	Mean (s.d)			
	Median (IQR)			
	Missing n (%)			

[1] Percentages for Multiple births are calculated using the total number of previous pregnancies in the respective arm.
[2] One pregnancy of twins would count as two births.
[3] Percentages for Mode of delivery, Shoulder Dystocia and Brachial plexus injury are all calculated using the total number of previous births in the respective arm.

Table 6: Reason for withdrawal, by withdrawal type, treatment arm and overall

Reason for withdrawal	Expectant management N=			Induction N=			Total (N=X)
	Withdrawn from randomisation to first follow-up	Withdrawn from follow-up	Withdrawn from long-term future follow-up	Withdrawn from intervention	Withdrawn from follow-up	Withdrawn from long-term future follow-up	
Woman's decision	N	N	N	N	N	N	N
Protocol Violation	N	N	N	N	N	N	N
Lost to follow-up	N	N	N	N	N	N	N
Trial stopped	N	N	N	N	N	N	N
Adverse event	N	N	N	N	N	N	N
Other	N	N	N	N	N	N	N
Total	N	N	N	N	N	N	N

Table 7: Level of withdrawal by treatment arm and overall[4]

Reason	Expectant management N=	Induction N=	Total
Withdrawn from Randomisation to first follow-up	N	N	N
Withdrawn from follow-up	N	N	N
Withdrawn from long-term possible future follow-up (within 25 years)	N	N	N
Total	N	N	N

Table 8: Timing of withdrawals throughout the trial (Woman)

	Expectant management N=	Induction N=	TOTAL
RANDOMISATION			
N			
Withdrawal	n (%)	n (%)	N
DELIVERY (RANDOMISATION TO DISCHARGE)			
N			

[4] Woman can choose more than one option of withdrawal, therefore there may be overlap in the table

Withdrawal (cumulative)	n (%)	n (%)	N
POST DELIVERY (DISCHARGE TO 2 MONTHS FOLLOW-UP)			
N			
Withdrawal (cumulative)	n (%)	n (%)	N
Non-responders	n (%)	n (%)	N
Died	n (%)	n (%)	N
POST DELIVERY (2 MONTHS TO 6 MONTHS FOLLOW-UP)			
N			
Withdrawal (cumulative)	n (%)	n (%)	N
Non-responders	n (%)	n (%)	N
Died	n (%)	n (%)	N

Table 9: Timing of withdrawals throughout the trial (Baby)

	Expectant management N=	**Induction N=**	**TOTAL**
RANDOMISATION			
N			
Withdrawal	n (%)	n (%)	N
DELIVERY (RANDOMISATION TO DISCHARGE)			
N			
Withdrawal (cumulative)	n (%)	n (%)	N
POST DELIVERY (DISCHARGE TO 2 MONTHS FOLLOW-UP)			
N			
Withdrawal (cumulative)	n (%)	n (%)	N
Non-responders	n (%)	n (%)	N
Died	n (%)	n (%)	N
POST DELIVERY (2 MONTHS TO 6 MONTHS FOLLOW-UP)			
N			
Withdrawal (cumulative)	n (%)	n (%)	N
Non-responders	n (%)	n (%)	N
Died	n (%)	n (%)	N

Table 10: Primary outcome measures and completeness of randomised participants, split by treatment arm

Primary Outcome		Expectant management N=	Induction N=	Unadjusted estimate (95% CI); p-value ł	Adjusted estimate (95% CI)**; p-value ł
(Estimand 1) Incidence of Shoulder Dystocia (Adjudication panel)*	Yes - n (%)				
	No - n (%)				
	Missing - n (%)				
	Undeterminable - n (%)				
(Estimand 2) Incidence of Shoulder Dystocia (Adjudication panel)	Yes - n (%)				
	No - n (%)				
	Missing - n (%)				
	Undeterminable - n (%)				
Incidence of shoulder dystocia (CRF)	Yes - n (%)				
	No - n (%)				
	Missing - n (%)				

*Primary analysis

Adjusted using variables: Site, estimated fetal weight centile and Maternal Age; **ł – statistical analysis based on complete data

Table 11: Secondary outcomes, split by RCT arms, with unadjusted and adjusted estimates

Secondary outcomes		Expectant care N=	Induction N=	Unadjusted estimate (95% CI); p-value ŧ	Adjusted estimate (95% CI)*; p-value ŧ
Time between delivery of head and delivery of body (minutes) – (Vaginal delivery only)[5]	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				
Time from commencement of active second stage of labour until fetal expulsion (minutes) – Vaginal delivery only	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				
	Not available/not applicable n (%)				
Intrapartum fetal outcomes:					
Time in labour ward (hours) – Vaginal & C-section[6]	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				
	Not applicable n (%)				
Post-partum Fetal outcomes:					
Stillbirth	Stillbirth n (%)				
	Missing n (%)				

[5] Calculated by Date and time head delivered – Date and time body delivered.
[6] Calculated by Date and time discharged from labour ward – Date and time admitted to labour ward.

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ŧ**	**Adjusted estimate (95% CI)*; p-value ŧ**
Neonatal death	Deaths n (%)				
	Missing n (%)				
Birthweight (g)	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				
Birthweight difference between treatment arms (g)	Mean difference (Expectant-Intervention)				
	Median difference (Expectant-Intervention)				
Gestation at birth (days)	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				
Gestation difference between treatment arms (days)	Mean difference (Expectant-Intervention)				
	Median difference (Expectant-Intervention)				
Birthweight customised centile	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				
Apgar score at five minutes	Score 7-10 (Good) n (%)				
	Score 0-6 (Poor) n (%)				
	Missing n (%)				

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ŧ**	**Adjusted estimate (95% CI)*; p-value ŧ**
Humeral fracture	No n (%)				
	Yes n (%)				
	Missing n (%)				
Clavicular fracture	No n (%)				
	Yes n (%)				
	Missing n (%)				
Brachial plexus palsy	No n (%)				
	Yes n (%)				
	Missing n (%)				
Admission to neonatal unit/ Receive additional care[7]	No n (%)				
	Yes n (%)				
	Missing n (%)				
Duration of stay at neonatal unit (days)[8]	N (Only if received additional care)				
	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				

[7] Counted if admitted to Intensive care, High dependency care, Special care or Transitional care.
[8] Calculated by Date of discharge-Date of admission, for both same and different hospital transfers.

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ŧ**	**Adjusted estimate (95% CI)*; p-value ŧ**
Hypoxic-ischaemic encephalopathy diagnosed	No n (%)				
	Yes n (%)				
	Missing n (%)				
Use of phototherapy	No n (%)				
	Yes n (%)				
	Missing n (%)				
Supplemental O_2	No n (%)				
	Yes n (%)				
	Missing n (%)				
Mechanical ventilation	No n (%)				
	Yes n (%)				
	Missing n (%)				
Non-invasive respiratory support	No n (%)				
	Yes n (%)				
	Missing n (%)				
Extracorporeal Membrane oxygenation	No n (%)				
	Yes n (%)				
	Missing n (%)				

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ŧ**	**Adjusted estimate (95% CI)*; p-value ŧ**
Nitric oxide (iNO) therapy	No n (%)				
	Yes n (%)				
	Missing n (%)				
Hypoglycaemia	No n (%)				
	Yes n (%)				
	Missing n (%)				
Neonatal readmissions					
Hospital readmission within 30 days of postnatal inpatient discharge	No n (%)				
	Yes n (%)				
	Unknown n (%)				
	Missing n (%)				
Intrapartum maternal outcomes:					
Duration of hospital stay prior to delivery (days)[9]	Mean (s.d)				
	Median (IQR)				
	Missing n (%)				
	Not applicable n (%)				

[9] Calculated by Date and time of third stage of labour – Date and time admitted to hospital.

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ŧ**	**Adjusted estimate (95% CI)*; p-value ŧ**
Labour type onset	Spontaneous n (%)				
	Induced n (%)				
	Caesarean section n (%)				
Mode of delivery	Vaginal delivery n (%)				
	Instrumental delivery - Ventouse n (%)				
	Instrumental delivery – Forceps n (%)				
	Instrumental delivery – Rotational forceps n (%)				
	Elective caesarean section n (%)				
	Emergency caesarean section n (%)				
	Missing n (%)				
Presentation at birth	Cephalic n (%)				
	Breech n (%)				
	Transverse lie n (%)				
Estimated blood loss at delivery (mls)	N				
	Mean (sd)				
	Median (IQR)				
	Minimum-Maximum				
	Missing – n (%)				

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ŧ**	**Adjusted estimate (95% CI)*; p-value ŧ**
Primary postpartum haemorrhage (≥500ml)	No n (%)				
	Yes n (%)				
	Missing n (%)				
Duration of hospital stay after delivery (days)[10]	Mean (s.d.)				
	Median (IQR)				
	Missing n (%)				
Intrapartum maternal trauma:					
Did the woman have an episiotomy	No n (%)				
	Yes n (%)				
	Missing n (%)				
Did the woman sustain a perineal injury?	No n (%)				
	Yes n (%)				
	Missing n (%)				
If perineal injury, please indicate the degree type **	N (only if had perineal tear)				
	First degree n (%)				

[10] Calculated by Date and time discharged from hospital - Date and time of third stage of labour.

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ł**	**Adjusted estimate (95% CI)*; p-value ł**
	Second degree n (%)				
	Third degree – 3a n (%)				
	Third degree – 3b n (%)				
	Third degree – 3c n (%)				
	Fourth degree n (%)				
	Missing n (%)				
Cervical laceration or tear	No n (%)				
	Yes n (%)				
	Missing n (%)				
Did the woman have retained placenta requiring manual removal?	No n (%)				
	Yes n (%)				
	Missing n (%)				
Did the woman die?	No n (%)				
	Yes n (%)				
	Missing n (%)				
Did the woman have sepsis in labour or within the 24 hours post-partum?	No n (%)				
	Yes n (%)				
	Missing n (%)				
	No n (%)				

Secondary outcomes		**Expectant care N=**	**Induction N=**	**Unadjusted estimate (95% CI); p-value ł**	**Adjusted estimate (95% CI)*; p-value ł**
Did the woman have a fever>38°C in labour or within 24 hours postpartum?	Yes n (%)				
	Missing n (%)				
Maternal readmissions					
Hospital readmission within 30 days of postnatal inpatient discharge	No n (%)				
	Yes n (%)				
	Missing n (%)				
Post-partum composite outcomes:					
Intra-partum birth injury: one or both of fractures or brachial plexus injury	No n (%)				
	Yes n (%)				
	Missing n (%)				
Prematurity associated problems: one or both of use of phototherapy or respiratory support	No n (%)				
	Yes n (%)				
	Missing n (%)				
Intra-partum composite outcomes:					
Maternal intra-partum complications: one of more of 3rd or 4th degree perineal tear, cervical laceration or tear, or primary postpartum haemorrhage	No n (%)				
	Yes n (%)				
	Missing n (%)				

Secondary outcomes		Expectant care N=	Induction N=	Unadjusted estimate (95% CI); p-value ł	Adjusted estimate (95% CI)*; p-value ł
Adverse Events					
Did the woman or her baby experience any adverse events up to the point of discharge from hospital following delivery?	Yes – n (%)				
	No – n (%)				
	Missing – n (%)				

*Adjusted using variables: Site, estimated Fetal weight centile and Maternal Age; **ł – statistical analysis based on complete data**

Table 12: Hospitalisations – Pre and Post Delivery

Hospitalisations		Expectant care N=	Induction N=	Total N=
Unscheduled hospital visit: Post recruitment prior to delivery				
Unscheduled	Yes n(%)			
Number of hospitalisations	N			
Admission	Yes n (%)			
Stay in days	Mean (SD)			
Medication	Yes n (%)			
Number of adverse events prior to delivery	N			

Unscheduled hospital visit: ≤30 days Post-delivery (Neonatal)				
Baby died since discharge	Yes n (%)			
Number of hospitalisations	N			
Admission	Yes n (%)			
Admission related to shoulder dystocia	Yes n (%)			
Stay in days	Mean (SD)			
Unscheduled hospital visit: ≤30 days Post-delivery (Maternal)				
Number of hospitalisations	N			
Admission	Yes n (%)			
Stay in days	Mean (SD)			

Table 13: Subgroup analyses of incidence of shoulder dystocia

Subgroups	**Expectant management N (%)**	**Intervention N(%)**	**Effect estimate (95% CI)**	**Interaction effect (95% CI); p-value**
Maternal body mass index *				
Subgroup 1				
Subgroup 2				
Estimated Fetal weight centile *				
Subgroup 1				
Subgroup 2				

*Cut point for subgroups to be determined when full data available

Table 14: Summary of RCT & Cohort patients (Requesting and not requesting planned C-section) split by Robson scores

Robson Score – n (%)	RCT N=	Requesting planned C-section N=	Not requesting planned C-section N=
1- Nulliparous women with single cephalic pregnancy, at greater than or equal to 37 weeks gestation in spontaneous labour			
2- Nulliparous women with single cephalic pregnancy, at greater than or equal to 37 weeks gestation who either had labour induced or were delivered by a caesarean section before labour			
3- Multiparous women, without a previous uterine scar, with a single cephalic pregnancy at greater than or equal to 37 weeks in spontaneous labour			
4- Multiparous women, without a previous uterine scar, with a single cephalic pregnancy at greater than or equal to 37 weeks who either had labour induced or were delivered by a caesarean section			

5- All multiparous women, with at least one previous uterine scar and a single cephalic pregnancy at great than or equal to 37 weeks gestation			

Table 15: Summary of Adverse events (AEs) for baby, split by treatment arm

Study number	Event details	Onset Date	Resolved date	Duration	Relatedness	Severity	Study arm

Table 16: Summary of Adverse events (AEs) for woman, split by treatment arm

Study number	Event details	Onset Date	Resolved date	Duration	Relatedness	Severity	Study arm

Table 17: Summary of Serious Adverse events (SAEs) for baby, split by treatment arm

Study number	Event details	Onset Date	Resolved date	Duration	Relatedness	Severity	Study arm

Table 18: Summary of Serious Adverse events (SAEs) for woman, split by treatment arm

Study number	Event details	Onset Date	Resolved date	Duration	Relatedness	Severity	Study arm

Table 19: Summary of protocol violations, deviations etc.

CAPA Number	TNO	Issue	Date aware	Date resolved/actions implemented	File note/ Deviation/ Violation/ Breach	Treatment Group

1						
…						

Table 20: Safety –Adverse Events

Outcome	**Expectant management N (%) - AE**	**Expectant management N (%) - SAE**	**Intervention N (%) - AE**	**Intervention N (%) - SAE**	**Total N (%) - AE**	**Total N (%) - SAE**
System Organ Class (SOC)						
Blood and lymphatic system disorders n (%)						
Cardiac disorders n (%)						
Congenital, familial and genetic disorders n (%)						
Endocrine disorders n (%)						
Gastrointestinal disorders n (%)						
General disorders and administration site conditions n (%)						
Hepatobiliary disorders n (%)						
Infections and infestations n (%)						
Injury, poisoning and procedural complications n (%)						

Outcome	**Expectant management N (%) - AE**	**Expectant management N (%) - SAE**	**Intervention N (%) - AE**	**Intervention N (%) - SAE**	**Total N (%) - AE**	**Total N (%) - SAE**
Investigations n (%)						
Metabolism and nutrition disorders n (%)						
Pregnancy, puerperium and perinatal conditions n (%)						
Renal and urinary disorders n (%)						
Respiratory, thoracic and mediastinal disorders n (%)						
Skin and subcutaneous tissue disorders n (%)						
Surgical and medical procedures n (%)						
Vascular disorders n (%)						
Not coded n (%)						
Immune system disorders n (%)						
Infections and infestations n (%)						
Injury, poisoning and procedural complications n (%)						

Outcome	Expectant management N (%) - AE	Expectant management N (%) - SAE	Intervention N (%) - AE	Intervention N (%) - SAE	Total N (%) - AE	Total N (%) - SAE
Immune system disorders n (%)						
At least one SAE reported						
At least one AE reported						
No AE/SAE reported						
Only 1 event reported						
Only 2 events reported						
Only 3 events reported						
4 and more events reported						

Table 21: Baseline characteristics and delivery outcomes for randomised and non-randomised (cohort) participants

Variables		RCT - Total N=	Cohort - Requesting planned C-section N=	Cohort – Not requesting planned C-section N=
Recruitment				
Maternal age at recruitment	≤35 years			
	>35 years			
Maternal age at recruitment	N			
	Mean (sd)			
	Median (IQR)			
	Minimum - Maximum			

	Missing – n (%)			
Parity (≥24 weeks)	0 – n (%)			
	1 – n (%)			
	2 – n (%)			
	3+ – n (%)			
	Missing – n (%)			
GROW Centile	90.1-91.0 – n (%)			
	91.1-92.0 – n (%)			
	92.1-93.0 – n (%)			
	93.1-94.0 – n (%)			
	94.1-95.0 – n (%)			
	95.1-96.0 – n (%)			
	96.1-97.0 – n (%)			
	97.1-98.0 – n (%)			
	98.1-99.0 – n (%)			
	99.1-100.0 – n (%)			
	Missing – n (%)			
Baseline				
Gestational diabetes mellitus (diet controlled)	Yes – n (%)			
	No – n (%)			
	Missing – n (%)			
Was there a previous C-section	Yes – n (%)			
	No – n (%)			
	Missing – n (%)			
Delivery outcomes				
Onset of labour type	Spontaneous – n (%)			
	Induced – n (%)			
	No labour (Caesarean section) – n (%)			
	Missing – n (%)			
	Vaginal delivery – n (%)			
	Instrumental delivery – n (%)			

Final mode of delivery (If onset vaginal)	Emergency caesarean section – n (%)			
	Missing – n (%)			
Was there shoulder dystocia?	Yes – n (%)			
	No – n (%)			
	Missing – n (%)			
Baby outcomes				
Outcome at delivery	Live – n (%)			
	Stillbirth – n (%)			
	Missing – n (%)			
Birth weight (g)	N			
	Mean (sd)			
	Median (IQR)			
	Minimum - Maximum			
	Missing – n (%)			
Gestation at birth (Weeks + Days)	N			
	Mean (sd)			
	Median (IQR)			
	Missing – n (%)			
Birthweight customised centile	<70.0 – n (%)			
	70.0-75.0 – n (%)			
	75.1-80.0 – n (%)			
	80.1-85.0 – n (%)			
	85.1-90.0 – n (%)			
	90.1-95.0 – n (%)			
	95.1-100.0 – n (%)			
	Missing – n (%)			

Table 22: Long-term outcomes (6-months)

Variables	**Expectant management N=**	**Intervention N=**	**Mean Difference (95% CI)**
EPDS Score			
0-11 - n (%)			
≥12 (Highly possible depression) – n (%)			
Mean (SD)			
ICIQ-UI			
Mean (SD)			
Never – urine does not leak – n (%)			
Leaks before you can get to toilet – n (%)			
Leaks when you cough or sneeze – n (%)			
Leaks when you are asleep – n (%)			
Leaks when you are physically active/exercising – n (%)			
Leaks when you have finished urination and are dressed – n (%)			
Leaks for no obvious reason – n (%)			
Leaks all the time – n (%)			
Bowel Function – n (%)			
Did you sometimes experience loss of bowel control?			

Did you sometimes have soiling from back passage on your underwear?			
Did you sometimes feel the need to go and have to go immediately?			
Have you taken treatment for constipation?			
Have you had haemorrhoids (piles)?			
Sexual Intercourse – Pain during sexual intercourse – n (%)			
Not applicable			
Never			
Sometimes			
Most times			
Always			
EQ-5D-5L – Index score			
Mean (SD)*			
Breastfeeding			
Breast milk only – n(%)			
Breast milk and formula – n(%)			
Infant health concerns			
Has baby been diagnosed with a palsy? (Yes) – n (%)			
Having difficulty with arm or finger movement (Yes) – n (%)			

Does face droop on one side? (Yes) – n (%)			
Receiving ongoing specialist medical care? (Yes) – n (%)			
Have concerns about baby's health and development? (Yes) – n (%)			

*Health Economic outcome

Appendix 5: The Estimands Framework

Primary Estimand attributes (E1)

Estimand attribute	Description
Population	All randomised patients who were successfully randomised to either early induction or standard care arm, including all protocol violations
Treatment condition(s)	Early induction of labour compared to standard care
Variable (outcome)	Presence of shoulder dystocia (yes/no) binary outcome, as determined by an independent adjudication committee.
Strategies used to handle intercurrent events	1) Randomised in error (non-eligible from EFW centile) 2) Change of planned induction timing for all randomised due to emergencies 3) Missing data due to dropouts

	All above protocol violations will be included in the ITT analysis as this is a pragmatic trial
Population-level summary measure	Risk/Odds ratio of intervention unadjusted and adjusted for stratification variables (site, maternal age [≤35 years, >35 years] and estimated fetal weight centile [≤95th EFW centile. >95th EFW centile]) and clinical factors (i.e. diet controlled gestational diabetes) with 95% confidence intervals

ITT: Intention to treat analysis.

Secondary Estimand attributes (E2)

Estimand attribute	Description
Population	All randomised patients who were successfully randomised to either early induction **and induced at 38+0 to 38+4 days** compared to women in the standard care arm **who had not started labour by 38+0, and were not induced or delivered by elective caesarean section between 38+0 and 38+4 days**
Treatment condition(s)	Early induction of labour compared to standard care
Variable (outcome)	Presence of shoulder dystocia (yes/no) binary outcome, as determined by an independent adjudication committee.
Strategies used to handle intercurrent events	1) Randomised in error (non-eligible from EFW centile) 2) Change of planned induction timing for all randomised due to emergencies 3) Missing data due to dropouts

	All above protocol violations will be excluded from the analysis using a per-protocol approach
Population-level summary measure	Risk/Odds ratio of intervention unadjusted and adjusted for stratification variables (site, maternal age [≤35 years, >35 years] and estimated fetal weight centile [≤95th EFW centile. >95th EFW centile]) and clinical factors (i.e. diet controlled gestational diabetes) with 95% confidence intervals

Sensitivity Estimand attributes (E3) (Only implemented if more than 10% missingness in model covariates)

Estimand attribute	Description
Population	All randomised patients who were successfully randomised to either early induction or standard care arm, including all protocol violations
Treatment condition(s)	Early induction of labour compared to standard care
Variable (outcome)	Presence of shoulder dystocia (yes/no) binary outcome, as determined by an independent adjudication committee.
Strategies used to handle intercurrent events	1) Randomised in error (non-eligible from EFW centile) 2) Change of planned induction timing for all randomised due to emergencies 3) Missing data due to dropouts All above protocol violations will be included in the ITT analysis as this is a

	pragmatic trial, with missing data imputed using multiple imputation methods
Population-level summary measure	Risk/Odds ratio of intervention unadjusted and adjusted for stratification variables (site, maternal age [≤35 years, >35 years] and estimated fetal weight centile [≤95th EFW centile. >95th EFW centile]) and clinical factors (i.e. diet controlled gestational diabetes) with 95% confidence intervals and missing values imputed using multiple imputation methods.