Statistical Analysis Plan

TRIAL FULL TITLE	Real Time Continuous Glucose Monitoring in Neonatal Intensive Care
EUDRACT NUMBER	
SAP VERSION	2.0
ISRCTN NUMBER	12793535
SAP VERSION DATE	24/07/2019
TRIAL STATISTICIAN	Dr Simon Bond
TRIAL CHIEF INVESTIGATOR	Dr Kathryn Beardsall
SAP AUTHOR	Annabel Allison

1 SAP Signatures

I give my approval for the attached SAP entitled REACT dated 24/07/2019.

Chief Investigator

Name: Dr Kathryn Beardsall

Signature: _____

Date: _____

Trial Statistician

Name: Dr Simon Bond

Signature: _____

Date: _____

SAP Author

Name: Annabel Allison

Signature: _____

Date: _____

2 Table of Contents

3 Abbreviations and Definitions

AE	Adverse Event
ACF	Autocorrelation Function
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
AVD	Assisted Vaginal Delivery
BG	Blood Glucose
BPD	Bronchopulmonary dysplasia
CGM	Real Time Continuous Glucose Monitoring
CRF	Case Report Form
CRIB II	Clinical Risk Index for Babies Score
DBM	Donor Breast Milk
DMC	Data Monitoring Committee
EBM	Expressed Breast Milk
MA	Moving Average
MCAR	Missing Completely at Random
NEC	Necrotising Enterocolitis
NHS	National Health Service
NICU	Neonatal Intensive Care Unit
PACF	Partial-Autocorrelation Function
PDA	Patent Ductus Arteriosus
RCT	Randomised Controlled Trial
ROP	Retinopathy of Prematurity
SAP	Statistical Analysis Plan
SD	Standard Deviation
SDS	Standard Deviation Score
SG	Sensor Glucose
SVD	Spontaneous Vaginal Delivery
TMG	Trial Management Group
TSC	Trial Steering Committee

4 Introduction

4.1 Preface

Increasing numbers of infants are being born preterm. These infants require intensive care and have a high risk of mortality and morbidity [1]. Surviving infants have a high incidence of long term health problems, including learning difficulties with significant long term costs to the National Health Service (NHS) and society [2]. Treatable neonatal causes of long term health problems have been difficult to establish. National Priorities for Research have highlighted investigation of the management of babies born too early or too small, and evaluation of the reasons for variations in outcome of "high risk" neonates. Early postnatal glucose control may be an important modifiable risk factor for clinical outcomes. In utero, glucose levels are normally maintained between 4-6mmol/l [3], but infants born preterm are at risk of both hyperglycaemia (>8mmol/l, 20-86%) and hypoglycaemia (<2.6mmol/l, 17%) [4].

4.2 Purpose of the analyses

These analyses will evaluate the efficacy, safety, and utility of real time continuous glucose monitoring (CGM) in Neonatal Intensive Care (NICU).

5 Study Objectives and Endpoints

5.1 Study Objectives

(ICH E3; 8)

5.1.1 Primary Objectives

- To evaluate the efficacy of CGM in helping control levels of glucose in the preterm infant
- To evaluate clinical acceptability in the preterm infant
- To assess safety in terms of risk for hypoglycaemia in the preterm infant

5.1.2 Secondary Objective

- To evaluate the cost-effectiveness and NHS importance of such an intervention

5.2 Endpoints

(ICH E9; 2.2.2)

5.2.1 Primary Outcome

- Percentage of time sensor glucose (SG) in target range of 2.6-10mmol/l within the first 6 days of life in preterm infants

5.2.2 Secondary Outcomes

Efficacy

- Mean SG in the first 6 days
- Percentage of time SG in target of 4-8mmol/l within the first 6 days of life
- SG variability within individuals as assessed by within-patient standard deviation
- Percentage of time glucose levels in hyperglycaemic range (SG>15mmol/l)

Acceptability

- Clinical staff rating score of impact on clinical care
- Frequency of blood glucose (BG) monitoring

Safety

- Incidence of hypoglycaemia defined as any episode of BG >2.2mmol/l and <2.6mmol/l
- Incidence of hypoglycaemia defined as a continuous episode of SG <2.6mmol/l for >1 hour
- Incidence of severe hypoglycaemia defined as any episode of BG ≤2.2mmol/l

Health Economics

- Cost-effectiveness expressed in terms of incremental cost per additional case of adequate glucose control between 2.6-10mmol/l

5.2.3 Exploratory Outcomes

- Mortality before 36 weeks corrected gestational age
- Maximum severity of Retinopathy of Prematurity (ROP) across all retinal examinations [5]

- Bronchopulmonary dysplasia (BPD): need for supplemental oxygen or respiratory support at 36 weeks corrected gestational age
- Microbiologically confirmed or clinically suspected late onset invasive infection
- Necrotising enterocolitis (NEC) requiring surgical intervention (including peritoneal drainage) or causing death
- Patent ductus arteriosus (PDA) requiring medical or surgical treatment
- Maximum grade of intracranial haemorrhage before discharge (Papile and Burstein grading)
- Growth (weight, length, and head circumference at the end of week 1 and at 36 weeks corrected gestation)
- Nutritional intake during the first week of life (carbohydrate, protein, and lipid)
- Use of insulin during the first and second week of life

6 Study Methods

6.1 General Study Design and Plan

(ICH E3; 9)

This is an open-label, multi-centre, parallel group, randomised controlled trial (RCT) comparing CGM with paper based algorithm to standard clinical management (control). The control group will have their glucose control monitored and managed according to standard clinical practice using point of care BG monitoring. These patients will have a sensor inserted and the device will collect glucose data continuously but the clinical team will be blinded to that data.

Treatment will be assigned using stratified block randomisation. The flow chart in Figure 1 shows the timing of randomisation, and the sequence and duration of all study periods.

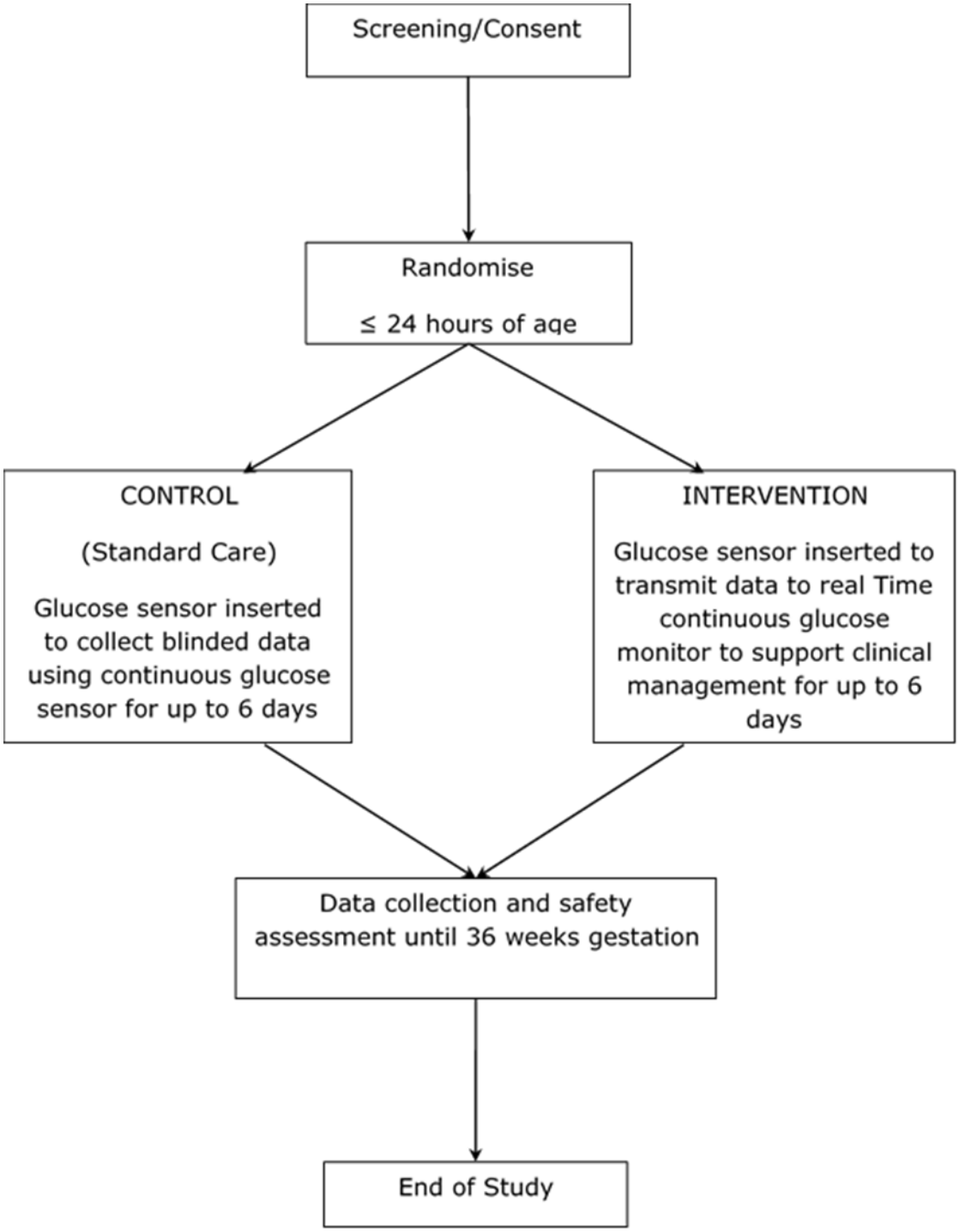

Figure 1: Study design flowchart

6.2 Inclusion-Exclusion Criteria and General Study Population

(ICH E3; 9.3. ICH E9; 2.2.1)

6.2.1 Inclusion Criteria

To be included in the trial the patient needs:

- Parental informed consent
- To be ≤33+6 weeks gestation
- To be ≤24 hours of age
- Birth weight ≤1200g

6.2.2 Exclusion Criteria

The presence of any of the following will preclude patient inclusion:

- A lethal congenital abnormality known at trial entry
- Any congenital metabolic disorder known at trial entry
- Neonates who, in the opinion of the treating clinician at trial entry, have no realistic prospect of survival

6.3 Randomisation and Blinding

(ICH E3; 9.4.3, 9.4.6. ICH E9; 2.3.1, 2.3.2)

6.3.1 Randomisation

Following parental consent and confirmation of eligibility, patients will be randomised on a 1:1 ratio into control and intervention arms using a web based system, Tenalea. The randomisation will use stratified block randomisation with varying block sizes. The stratification factors will be recruiting centres and gestation (<26 weeks gestation, ≥26 weeks gestation).

6.3.2 Blinding

This is an open study in which the clinical staff, research team, and parents will be aware of the study arm and intervention. In the control group arm, the display screens of the sensors will be obscured by a cover fastened by a tamper proof seal so that the clinical team are blinded to the data being collected.

6.4 Study Variables

(ICH E3; 9.5.1. ICH E9; 2.2.2)

Table 1 shows the timing of assessments. All participants will already be hospitalised in a level 3 NICU and it is expected that the majority of participants will still be hospitalised at the 36 weeks corrected gestation assessment. In the event where participants have been discharged from the recruiting centre, a time window of ±7 days will apply to both the 14 day and 36 weeks corrected gestation follow-up assessments.

- SG will be recorded in either mmol/L or mg/dl. SG measured in mg/dl will be converted to mmol/L by dividing the measurements by 18 [6]

- Clinical staff rating score – the questionnaire consists of five Likert scale questions and two free text. The Likert scale responses will be combined to give an overall score ranging from 5 to 25 with 5 = positive response and 25 = negative response

- Ethnicity is measured on an unordered categorical scale for which 1 = White, 2 = Black, 3 = Asian, 4 = Mixed, and 5 = Other

- Gestational age at birth is measured in weeks and days

- Weight is measured in grams

- Body length and head circumference are measured in cm

- Temperature immediately after birth is recorded in °C

- Base excess is recorded in mEq/L

- Maternal use of antenatal steroids, maternal chorioamnionitis, and maternal diabetes are all binary variables with 1 = Yes and 0 = No

Table 1: Time and Events Table

	≤ 24 Hours of Age	**Trial Assessments**			**Final Assessment**
Day	Study Day 1	Study Days 2 - 6	Study Day 7	Study Day 14	36 weeks corrected gestation
Assessments					
Informed Consent	×				
Inclusion/Exclusion	×				
Demography - NICU admission date - Name of referring Unit, if applicable - Date & time of birth - Sex - Gestational age at time of birth - Ethnicity	×				
Randomisation - Date of randomisation/person randomising - Case or control	x				
Maternal Pregnancy & Delivery History - Date of LMP - Date of EDD - Antenatal history - Singleton/Multiple birth – ranking - Method of delivery - Apgar scores - Resuscitation history - Temperature on Admission - Base excess on Admission - Birth weight/Length/Head Circumference	×				

	≤ 24 Hours of Age	**Trial Assessments**			**Final Assessment**
Day	Study Day 1	Study Days 2 - 6	Study Day 7	Study Day 14	36 weeks corrected gestation
AE/SAE/ADE/SADE	**x**	**x**	**x**	**x**	**x**
Specific Concomitant Medications	**X**	**x**	**X**		
Clinical care - Medication - Oral/IV intake – protein, lipid and dextrose - Total daily dose of Insulin - Record hypoglycaemia BG >2.2 - <2.6mmol/l - Record hypoglycaemia BG <2.2mmol/l - Record time and result of BG, POC, lactate & ketones - Record hypoglycaemia SG <2.6mmol/l - Record hypoglycaemia SG <2.2mmol/l - Insulin requirement from Day 7	**X** **x** **x** **x** **x** **x** **x** **x**	**X** **x** **x** **x** **x** **x** **x** **x**		**x**	
Baby's clinical condition/Care Record - Cardiovascular - PDA - Respiratory support - Gastrointestinal - NEC - Intraventricular Haemorrhage - Sepsis					**x**
Growth Assessment - Weight/Length/Head			**x**	**x**	**x**

	≤ 24 Hours of Age	**Trial Assessments**			**Final Assessment**
Day	Study Day 1	Study Days 2 - 6	Study Day 7	Study Day 14	36 weeks corrected gestation
Circumference					
Insertion of Glucose Sensor - Date & time of insertion - Insertion site - Sensor Lot no. - Sensor expiry date - Monitor details	**x**				
Sensor site check	**x**	**x**	**x**	**x**	**x**
Removal of Glucose Sensor Date of removal			**x**		
Data Monitoring with CGM	**x**	**x**			
Recording use of paper based algorithm (Appendix 1) regarding administration of Insulin & 20% Dextrose	**x**	**x**			
Parent & Clinician Questionnaires[1]		**X**	**x**		
Assessment of Resource Use - Mode of transport for admission/transfer - LOS (in days) at each level of neonatal care - Record number(s) of USS, EEG, MRI, ECHO, ECG, CT Scan - Record number of any other diagnostic tests/ procedures/ reviews carried out - Number of surgical procedures carried out					**x**
ROP – Maximum grade of ROP and confirmation of vascularised/not					**x**

	≤ 24 Hours of Age	**Trial Assessments**			**Final Assessment**
Day	Study Day 1	Study Days 2 - 6	Study Day 7	Study Day 14	36 weeks corrected gestation
vascularised					
Record number of days of each level of care required using BAPM definition					**x**
Discharge date/transfer date (record destination)/Details of Death if applicable	**x**				

7 Sample Size

(ICH E3; 9.7.2. ICH E9; 3.5)

Based on data from the REACT feasibility study (IRAS ID: 150193) and historical control data, we conservatively assume that the standard deviation (SD) of the primary endpoint (percentage time SG in target range of 2.6-10mml/l within the first six days of life) is 22%. A sample size of 200 participants will enable a treatment effect of a 10% increase in the mean value of the primary endpoint to be detected with 90% power using a two-sided 5% significance test in the primary analysis. Based on a consensus of expert opinion drawn from the Trial Steering Committee (TSC), Data Monitoring Committee (DMC) and the Trial Management Group (TMG), a difference of 10% is believed to be of minimal clinical relevance. It is expected that a small number of patients will be withdrawn from the study. Reasons for withdrawal from the intervention include transfer to the patient's local NICU, withdrawal of parental consent, or death.

8 General Considerations

8.1 Timing of Analyses

The final analysis will be performed after 200 subjects have completed the final assessment or have dropped out prior to the final assessment. The statistical analysis plan (SAP) will be approved according to the requirements of CCTU/SOP023 before database lock. Permission to lock the database will be requested from the chief investigator using CCTU/FRM094, and the data programmers will deposit the data in a folder to which the trial statistician is given restricted access (following CCTU/SOP057). Logical and graphical checks will be carried out to check that no out of range values are present. Data queries will be referred to the data management team for resolution (using CCTU/TPL064).

8.2 Analysis Populations

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 5.2)

8.2.1 Full Analysis Population

- All randomised subjects. This population will be used to analyse the primary and secondary efficacy data.

8.2.2 Per Protocol Population

Subjects will be excluded from the per protocol analysis for the following reasons:

- Intervention group: those patients with SG recorded on the paper CRF less than 80% of the time, where the percentage of time SG is recorded on the paper CRF is as defined in Section 9.2.5 a)
- Control group: those patients that have at least one occurrence where more than 12 CGM readings are available whilst the tamper tag is open

The DMC chair will also look at those subjects with <24hours worth of data to decide whether or not they should be included in the analysis.

The primary analysis will also be performed on this population.

8.2.3 Safety Population

- All consented subjects. This population will be used when analysing safety data.

8.3 Covariates and Subgroups

(ICH E3; 9.7.1, 11.4.2.1. ICH E9; 5.7)

Recruiting centre and gestation (<26 weeks, ≥26 weeks) will be used as stratification factors in the randomisation and therefore will be adjusted for in the final analysis.

The analysis will look for a treatment interaction effect with the following baseline variables: centre, sex, corrected gestational age, birth weight standard deviation score (SDS), use of antenatal steroids, maternal chorioamnionitis, and maternal diabetes using the regression framework in an exploratory, non-confirmatory manner.

8.4 Missing Data

(ICH E3; 9.7.1, 11.4.2.2. ICH E9; 5.3. EMA Guideline on Missing Data in Confirmatory Clinical Trials)

We anticipate a very low rate of drop-out in the clinical setting; however, the CONSORT diagram will illustrate any such drop-outs. Missing data at the individual variable level will be reported within the summary statistics by reference to the number of complete cases.

The primary outcome is based on CGM data. There are two types of missing CGM data:

1. Completely missing (i.e. due to failed sensor insertion or calibration issues)
2. Missing period of SG data whilst in the study (i.e. due to calibration issues/sensor replacement/death/withdrawal)

SG profile plots for each subject will be produced as a visual aid for the amount of missing CGM data. Summary statistics for the number of SG measurements recorded by the CGM will also be produced. The primary and secondary efficacy analyses will use all available SG measurements for each subject.

Analyses of the exploratory outcomes will use complete cases, which implicitly assumes any missing data is missing completely at random (MCAR).

8.5 Interim Analyses and Data Monitoring

(ICH E3; 9.7.1, 11.4.2.3. ICH E9; 4.1, FDA Feb 2010 "Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics")

8.5.1 Purpose of Interim Analyses

There are no planned interim analyses but the DMC will meet to review the trial's progress including updated figures on recruitment, data quality, and main endpoints including safety data.

8.5.2 Planned Schedule for DMC

The DMC will meet after 50 patients have been recruited and again after 125 patients have been recruited.

8.5.3 Scope of Adaptations

The DMC may recommend:

1. The trial continues as planned
2. Early stopping for clear evidence of harm due to a treatment in terms of its safety profile, strong evidence in favour of either treatment, or external evidence
3. Extending recruitment or follow-up
4. Sanctioning and/or proposing protocol changes

8.5.4 Practical Measures to Minimise Bias

The study statistician will produce the reports for the DMC. The members of the DMC and the study statistician will see data by treatment group at the interim. This includes efficacy and safety data by treatment group.

The chief investigator and other members of the TMG, as well as members of the DMC, will be able to see recruitment data, data quality, safety event details, and compliance to the paper-based algorithm. Representatives of the sponsor, funder or regulator can also see this information if required.

8.6 Multi-centre Studies

(ICH E3; 9.7.1, 11.4.2.4. ICH E9; 3.2)

The data from each centre will be combined for the analysis. The primary analysis will adjust for recruiting centre as it is a stratification factor of the randomisation.

8.7 Multiple Testing

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 2.2.5)

Multiplicity of data is potentially an issue in this study due to the fact that there are multiple secondary endpoints to be tested. Therefore, methods will be used to reduce the likelihood of a type I error. All of the efficacy endpoints will be ranked with one being the most important (the primary endpoint), two being the next most important, and so on.

The rankings of the efficacy endpoints in order of importance are:

Primary:

1. Percentage of time SG in target range of 2.6-10mmol/l within the first 6 days of life in preterm infants

Secondary:

2. Mean SG in the first 6 days
3. Percentage of time SG in target of 4-8mmol/l within the first 6 days of life
4. SG variability within individuals as assessed by within-patient SD
5. Percentage of time glucose levels in hyperglycaemic range (SG>15mmol/l)

The primary endpoint and the two top-ranked secondary endpoints are seen as the most important and so will be tested. The Benjamini-Hochberg procedure [7] will

then be applied to these three endpoints. The P-values for the three tests will be calculated and ranked in order of significance, $p_{(1)} < p_{(2)} < p_{(3)}$. The highest ranked P-value, $p_{(k)}$, such that $p_{(i)} < 0.05i/3$ for all $i \leq k$, is identified, and $0.05k/3$ is used as the nominal significance level, with any P-values lower than this value considered as significant.

For example, suppose the three hypotheses had P-values of $0.01, 0.04$ and 0.08. As $0.01 < 0.05 \times \frac{1}{3} = (0.017)$ and $0.04 > 0.05 \times \frac{2}{3} = 0.033$, the only hypothesis in this example to be rejected would be the one producing the P-value of 0.01. By doing this, if only one endpoint is significant at the 5% level, there needs to be stronger evidence to claim a significant result, but if all three are significant at the 5% level, a 5% level remains for each individual hypothesis.

The remaining two secondary endpoints will then be tested using a gatekeeping procedure. The next secondary endpoint down the list will be tested. If the P-value is > 0.05, none of the other endpoints are tested. However, if the P-value is < 0.05, the next hypothesis will then be tested. The same logic is then used and the procedure continues until a P-value > 0.05 is produced or all of the predefined endpoints are tested. By employing this method of only testing some of the endpoints if all of the previous P-values are < 0.05 it decreases the probability of obtaining a false positive result.

Using both of these rules ensures that the risk of obtaining a type I error is kept to a minimum. The results from all five efficacy endpoints will be reported. However, the discussion of the results will focus on the first three outcomes rather than the two endpoints further down the list. Any of the endpoints that are significant using the Benjamini-Hochberg method will be seen as important but consideration will be given to the fact that the study was designed to test the primary outcome and not the two most important secondary outcomes.

The safety analyses will all be tested regardless of the results of the primary outcome and any secondary efficacy outcomes tested.

9 Summary of Study Data

All continuous variables will be summarised using the following descriptive statistics: n (non-missing sample size), mean, standard deviation, median, maximum, and minimum. The frequency and percentages (based on the non-missing sample size) of observed levels will be reported for all categorical measures. All summary tables will be structured with a column for each treatment and overall in the order (Standard, CGM, Total) and will be annotated with the total population size relevant to that table/treatment, including any missing observations. Tables and figures will be annotated with the analysis population upon which they are based.

9.1 Subject Disposition

A CONSORT diagram (Figure 2) will be presented to show patient disposition within the study. The number eligible will be determined from the screening logs. The question "Was the baby randomised?" on page 4 of the CRF will be used to determine the number randomised. The number of patients allocated to each arm will be determined using the "Randomisation arm" question on page 4 of the CRF. The date and time of sensor insertion will be used to determine whether the patient reached study day 1 (i.e. received treatment). The number of patients with CGM data will be taken as the number of patients with any SG measurement recorded by the CGM. The date and time sensor removed will be used to determine whether the patient reached study day 7. The assessment date at day 14 will be used to determine if the patient reached day 14 of the study. The 36 weeks corrected gestation date will be used to determine if the patient reached 36 weeks corrected gestation. The withdrawal form of the CRF will be used to determine if a patient is withdrawn before each of the study time-points.

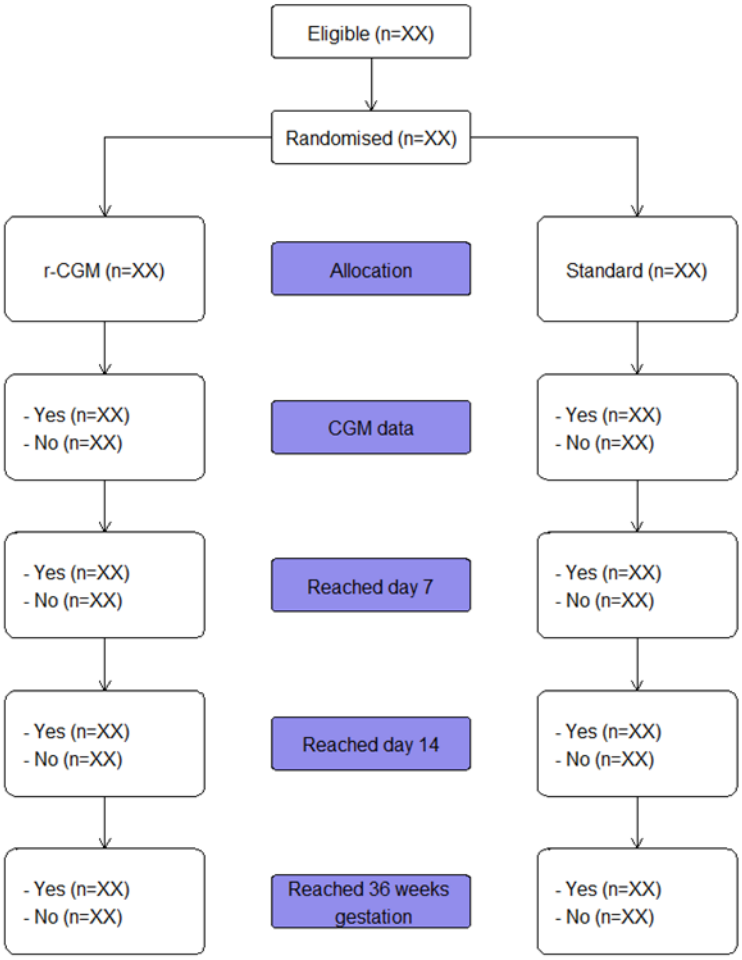

Figure 2: CONSORT diagram

9.2 Derived variables

The following variables will be derived for each subject, $i = 1,..,n$. The subscript i has been removed from equations to aid readability. The primary and secondary outcomes will be calculated using data from the first 6 days of life.

9.2.1 Primary outcome

- Percentage of time SG in target of 2.6-10mmol/l will be calculated from the SG readings taken every 5 minutes as follows:

$$\frac{100 \times \text{Number of readings within target}}{n}$$

where n is the total number of readings

9.2.2 Secondary outcomes

Efficacy

- Mean SG will be calculated from the SG readings taken every 5 minutes as follows:

$$\frac{\sum SG}{n}$$

where n is the total number of readings

- Percentage of time SG in target of 4-8mmol/l and percentage of time SG in hyperglycaemic range will be calculated in the same way as for the primary outcome (Section 9.2.1)

- SG variability will be calculated in two ways:
    1. From the SG readings taken every 5 minutes

$$\frac{\sum(SG - \text{mean}(SG))^2}{n-1}$$

where n is the total number of readings. This gives the unconditional variance (i.e. it does not take into account the ordering of the observations)

    2. From fitting an Autoregressive Integrated Moving Average (ARIMA) model to each subject's time series data. The initial model will include an autoregressive (AR) term of order one and no moving average (MA) terms. To check whether the model is appropriate, a plot of each time series will be produced to assess whether there is any trend in the data or seasonality. The Autocorrelation Function (ACF) and Partial-Autocorrelation Function (PACF) will be plotted to check that the correct model has been used. A plot of residuals against time will be produced to check that the residuals have mean zero and constant variance. If the plots suggest the model fit is poor, models with first order differences, more or less AR terms, and MA terms may be considered. The estimated conditional variance from the model will be calculated.

Safety

- The safety endpoints are binary outcomes derived from the BG and SG readings

9.2.3 Demographic and Baseline Variables

- Sensor duration will be calculated as time of sensor removal – time of sensor insertion (hours)

- The amount of CGM data (hours) will be calculated as

$$\frac{\text{Number of CGM readings}}{12}$$

- CRIB II score will be calculated according to [8]

- SDS for weight, body length, and head circumference will be calculated using the LMS method of Cole [9] and the growth reference from the birth section of the UK-WHO growth reference [10]. Age will be adjusted for gestation as follows:

$$\frac{\{\text{actual age (weeks)} - (40 - \text{gestational age (weeks)})\} \times 7}{365.25} \text{ years}$$

9.2.4 Medications

- Total number of times medications (inotropes, antibiotics, caffeine, morphine, corticosteroids, insulin) are given across study days 1 to 7 will be calculated as

$$\sum_{d}^{7} I_d$$

where d is the study day ($d = 1,..,7$) and I_d is an indicator taking the value one if a subject received the medication on day d and zero otherwise

9.2.5 Treatment Compliance

- Percentage of time SG is recorded on the paper CRF (CGM group only) will be calculated in two ways:

a)

$$\frac{100 \times \text{Number of times SG recorded on paper CRF}}{\text{Sensor duration (hours)}}$$

b)

$$\frac{100 \times \text{Number of times SG recorded on paper CRF}}{\text{Number of CGM readings}/12}$$

- Percentage of time insulin recorded when SG >8mmol/l (CGM group only) will be calculated as

$$\frac{100 \times \text{Number of insulin rate values} \mid \text{SG} > 8}{\text{Number of SG values} > 8}$$

for SG recorded on the paper CRF

9.2.6 Exploratory Outcomes

- Amount of insulin given (insulin units per kg per day) over days 1 to 7 will be calculated as

$$\frac{\sum_{d=1}^{D} 1000 \times \text{insulin strength (IU/ml)}_d \times \text{volume (ml)}_d}{\text{weight (g)} \times D}$$

where d is the study day ($d = 1,..,D$ for $D \leq 7$ dependent on whether or not subject i is still in the study at day 7). Amount of insulin given between days 8 and 14 will be calculated similarly

- Amount of dextrose given (mg per kg per day) over days 1 to 7 will be calculated as

$$\frac{\sum_{d=1}^{D} 1000 \times \text{concentration (g/l)}_d \times \text{volume (ml)}_d}{\text{weight (g)} \times D}$$

where d is the study day ($d = 1,..,D$ for $D \leq 7$ dependent on whether or not subject i is still in the study at day 7)

- Amount of lipid given (g per kg per day) over days 1 to 7 will be calculated as

$$\frac{\sum_{d=1}^{D} \text{concentration (g/l)}_d \times \text{volume (ml)}_d}{\text{weight (g)} \times D}$$

where d is the study day ($d = 1,..,D$ for $D \leq 7$ dependent on whether or not subject i is still in the study at day 7)

- Amount of parenteral glucose given (g per kg per day) over days 1 to 7 will be calculated as

$$\frac{\sum_{d=1}^{D} \text{concentration (g/l)}_d \times \text{volume (ml)}_d}{\text{weight (g)} \times D}$$

where d is the study day ($d = 1,..,D$ for $D \leq 7$ dependent on whether or not subject i is still in the study at day 7)

- Amount of amino acid given (g per kg per day) over days 1 to 7 will be calculated as

$$\frac{\sum_{d=1}^{D} 10 \times \text{concentration (g/100ml)}_d \times \text{volume (ml)}_d}{\text{weight (g)} \times D}$$

where d is the study day ($d = 1,..,D$ for $D \leq 7$ dependent on whether or not subject i is still in the study at day 7)

- Amount of oral nutrition given (ml per kg per day) over days 1 to 7 will be calculated as

$$\frac{\sum_{d=1}^{D} 1000 \times (\text{EBM (ml)}_d + \text{DBM (ml)}_d + \text{Formula (ml)}_d)}{\text{weight (g)} \times D}$$

where d is the study day ($d = 1,..,D$ for $D \leq 7$ dependent on whether or not subject i is still in the study at day 7)

9.3 Protocol Deviations

Protocol deviations will be defined for the control and intervention groups as follows:

- Control group: at least one occurrence where more than 12 CGM readings are available whilst the tamper tag is open

- Intervention group: those patients with SG recorded on the paper CRF less than 80% of the time as defined in Section 9.2.5 a)

Patients that deviate from the protocol based on the above definitions will be excluded from the per protocol analysis.

9.4 Demographic and Baseline Variables

The following demographic and baseline variables will be summarised according to Section 9 using the full analysis population.

9.4.1 Demographics

- Gestational age (weeks)
- Gender
- Ethnicity (White/Black/Asian/Mixed/Other)
- Number of babies delivered (singleton/multiple)
- Weight (g) – reported as actual and SDS
- Body length (cm) – reported as actual and SDS
- Head circumference (cm) – reported as actual and SDS
- CRIB II score
- Sensor duration (hours)
- Amount of CGM data (hours)
- Number of sensors used

9.4.2 Pregnancy events

- Use of antenatal steroids >24hrs prior to delivery (yes/no)
- Use of antibiotics in 24 hours prior to delivery (yes/no)
- Maternal diabetes (yes/no)
- Maternal chorioamnionitis (yes/no/unknown)

9.4.3 Delivery details

- Delivery mode (spontaneous vaginal delivery (SVD)/caesarean section pre onset of labour/caesarean section post onset of labour/assisted vaginal delivery (AVD))
- Resuscitation required post-delivery (yes/no)
- Resuscitation category (intubated/oxygen required/cardiac massage given/intravenous drugs given)

9.5 Medications

Use of the following medications in study days 1 to 7 will be summarised according to Section 9 using the full analysis population.

- Inotropes (total number of times given across days 1 to 7)
- Antibiotics (total number of times given across days 1 to 7)
- Caffeine (total number of times given across days 1 to 7)

- Morphine (total number of times given across days 1 to 7
- Corticosteroids (total number of times given across days 1 to 7)
- Insulin (total number of times given across days 1 to 7)
- Insulin (insulin units per kg per day)
- Dextrose (mg per kg per day)
- Lipid (g per kg per day)
- Amino acids (g per kg per day)
- Parenteral glucose (g per kg per day)
- Oral nutrition (ml per kg per day)

9.6 Study Day 14

The following variables measured at day 14 will be summarised according to Section 9 using the full analysis population:

- Weight (g) – reported as actual and SDS
- Body length (cm) – reported as actual and SDS
- Head circumference (cm) – reported as actual and SDS
- Insulin (total between day 8 and 14, insulin strength per kg per day)

9.7 Treatment Compliance

To assess compliance with the paper based algorithm in the intervention group, plots of the SG recorded on the paper CRFs and insulin rate against time from sensor insertion will be produced. If SG is above 8mmol/l the algorithm suggests either starting insulin or reducing insulin rate depending on whether SG is falling, stable, or rising. Summary statistics for the percentage of time insulin rate is recorded when SG is above 8mmol/l will be reported. Summary statistics for the percentage of time SG is recorded on the paper CRF will also be reported using the two derivations provided in Section 9.2.5.

To assess compliance with blinding in the control group, the following will be reported:

- Summary statistics for the continuous variable length of time open
- Frequency and percentage of patients who have at least one occurrence where more than 12 CGM readings are available whilst the tamper tag is open

10 Efficacy Analyses

All efficacy variables will be summarised according to Section 9. The primary efficacy outcome will be analysed using both the full analysis population and the per protocol population. The secondary efficacy outcomes will be analysed using the full analysis population only.

10.1 Primary Efficacy Analysis

The primary efficacy outcome is

- Percentage time SG in target range of 2.6-10mmol/l within the first 6 days of life

A plot of the cumulative distribution function of the primary outcome will be produced. The primary analysis will be a linear regression model adjusted for the randomisation strata (gestation, centre). The model makes the assumption that missing data is MCAR, that is, there is no relationship between whether CGM data is missing and values in the dataset (missing or observed). The following confirmatory hypotheses will be tested using a 2-sided significance test at the 5% level:

- Null hypothesis: there is no difference in the primary outcome between the two treatment arms
- Alternative hypothesis: there is a difference in the primary outcome between the two treatment arms

Histograms and QQ-plots will be used to assess the normality of the residuals and a plot of the residuals versus fitted values will be used to assess homogeneity of variance.

The results from the linear regression will be presented in a table showing the treatment effect (CGM compared to Standard), standard error, 95% confidence interval, and the P-value.

As a sensitivity analysis, a linear regression model adjusted for start time of the CGM (time from birth) and baseline SG value will be fit to the data. The adjustment for start time will be incorporated as there is an expectation that a delayed start could worsen the performance of the treatment. The adjustment for baseline SG will be incorporated as it may be a predictor of the performance of treatment. The

randomisation strata will also be adjusted for in the model and the model checks used for the primary analysis will be repeated.

10.2 Secondary Efficacy Analyses

The following secondary efficacy analyses will be tested:

- The effect of CGM compared to standard care on the mean SG within the first 6 days
- The effect of CGM compared to standard care on the percentage time SG in target of 4-8mmol/l within the first 6 days
- The effect of CGM compared to standard care on the SG variability within individuals, as measured by within-patient SD
- The effect of CGM compared to standard care on the percentage time in hyperglycaemic range (SG >15mmol/l)

As the outcomes are all continuous, they will be analysed using the primary analysis method described in Section 10.1. The logarithm of within-patient SD will be taken before fitting the regression model.

11 Safety Analyses

When calculating the incidence of adverse events (AEs) and adverse device effects (ADEs), the numerator will be the number of patients with an event and the denominator will be the total population size for the relevant group. All summary tables will be structured with a column for each treatment and overall in the order (Standard, CGM, Total). The analysis will be based on the safety population.

11.1 Safety Outcomes

The three safety analyses to be performed are:

- The effect of CGM compared to standard care on the incidence of hypoglycaemia defined as any episode of BG >2.2mmol/l and <2.6mmol/l
- The effect of CGM compared to standard care on the incidence of hypoglycaemia defined as a continuous episode of SG <2.6mmol/l for >1 hour
- The effect of CGM compared to standard care on the incidence of hypoglycaemia defined as any episode of BG ≤2.2mmol/l

Summary statistics for the three safety outcomes will be produced according to Section 9. Summary statistics will also be reported for the length of time SG <2.6mmol/l. Logistic regression will be performed on each outcome adjusting for the randomisation strata.

11.2 Adverse Events

AEs will be assessed continuously during each patient's stay in the NICU. They will be listed and summarised according to Section 11.

11.3 Adverse Device Effects

ADEs are events related to the use of the investigational medical device. This includes any AE resulting from insufficiencies or inadequacies in the instruction for use, the deployment of the implantation, the installation, the operation, or any malfunction of the investigational medical device. It also includes any event that is a result of a use error or intentional misuse. They will be listed and summarised according to Section 11.

12 Exploratory Analyses

The exploratory outcomes will be summarised according to Section 9 using the full analysis population. Growth variables will be summarised using actual and SDS values but SDS will be used for analysis only.

The exploratory analyses are the effect of CGM compared to standard care on:

- mortality before 36 weeks corrected gestational age
- maximum severity of ROP across all retinal examinations
- BPD (need for supplemental oxygen or respiratory support at 36 weeks corrected gestational age)
- microbiologically confirmed or clinically suspected late onset invasive infection
- NEC requiring medical or surgical treatment
- PDA requiring medical or surgical treatment
- intracerebral pathology

which are measured from trial entry until either hospital discharge or 36 weeks corrected gestation, and

- growth (weight, length and head circumference at day 7 and 36 weeks corrected gestation)
- nutritional intake during the first week of life
- total insulin use during the first and second week of life

Continuous outcomes will be analysed using linear regression, binary outcomes will be analysed using logistic regression, and ordinal outcomes will be analysed using ordinal logistic regression. All models will adjust for the randomisation strata (gestation, centre).

13 Acceptability Analyses

The acceptability analyses will be based on the full analysis population.

The clinical staff questionnaires consist of five Likert scale questions and two free text responses. The responses to the Likert questions will be summarised using a stacked bar chart. An overall score will be calculated according to Section 6.4 and a histogram produced to show the distribution.

A histogram of the time between BG measurements by treatment group will be produced in order to look at the frequency of BG monitoring. Summary statistics will also be calculated according to Section 9 for length of time between BG measurements.

14 Figures

A CONSORT diagram showing patient disposition will be produced. Profile plots of SG will be produced for each patient and will be colour coded by treatment group. Plots of SG recorded on the paper CRFs against insulin rate will be produced for each patient in the CGM group. A stacked bar chart showing responses to the clinical staff questionnaire will be produced and a histogram will be used to show the distribution of the overall score. A histogram will be used to show the distribution of the timing between BG measurements by treatment group.

15 Reporting Conventions

P-values ≥ 0.001 will be reported to 3 decimal places in this report; P-values less than 0.001 will be reported as "<0.001". For manuscripts that use the values

produced in this report, conventions detailed in [11] will be implemented. The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

16 Technical Details

At the time of writing, the SAP is based on Version 4.0 of the protocol. The software package R version 3.5 will be used on a Windows computer and copies of the code written will be stored in the following folder:

- V:\STATISTICS\STUDY FOLDER\Paediatrics\REACT\3 Analysis\Final

To provide high quality code that is understandable and allows reproduction of the analysis, the following points will be followed. Each report and individual table or graph will have:

- Date and time stamp
- Name of the code file
- Author
- Population used

The version control system Git will be used and the individual code files will have comments that convey:

- Author
- Date and time of writing
- Description of any revisions
- Description of the inputs and outputs
- Reference to any parent code file that runs the code

The population to be used in a table or figure will be explicitly set at the start of a block of code that computed the output by looking up the population from the table of tables.

A reviewing statistician will independently reproduce the primary analyses. The reviewing statistician will have an overview of the entire analyses and will explicitly

check the code by producing a random selection of tables, as well as any other pieces of code as desired.

17 Summary of Changes to the Protocol

The endpoint "clinical use of algorithm" has been moved from acceptability analysis to treatment compliance, as it fits better into this category.

The original sample size calculation stated that recruitment would continue until 200 babies had a minimum of 5 days of CGM data. As it is not feasible to recruit until 200 babies have a minimum of 5 days of CGM data, recruitment will continue until either 200 babies have been recruited or until 31 December 2018.

18 Changes from SAP Version 1.0

- Recruitment continued until 31 January 2019.
- The CONSORT diagram has been revised (Figure 3, Section 20). The number screened is taken from the screening logs. The number consented is the number of unique subject IDs in the MACRO database. The number excluded was calculated using the question "Was the baby randomised?" on page 4 of the CRF. One subject was excluded post-randomisation as they were found to be ineligible. The number of subjects that reached day 7 was taken to be those with either a non-missing study day 7 assessment date or a non-missing assessment date at either day 14 or 36 weeks corrected gestation. The number of subjects that reached day 14 was taken to be those with either a non-missing study day 14 assessment date or a non-missing 36 weeks corrected gestation assessment date. The number of patients that died during the study was calculated using information recorded on the withdrawals form of the CRF.
- The full analysis and safety population definitions have been edited to exclude ineligible subjects (including those found to be retrospectively ineligible).
- A scatter plot showing mean insulin given (units per kg) and mean SG against study day by treatment group has been produced.
- An AR2 model with first order differencing was fitted to each subject's CGM data to calculate within-patient SD.

- The plot of SG recorded on paper CRFs and insulin rate (Section 9.7) was plotted against time from birth instead of time from sensor insertion.
- A histogram showing the amount of CGM data recorded by treatment group has been produced.
- Mean and error bar plots showing the change in weight SDS, body length SDS, and head circumference SDS over time by treatment group have been produced.
- The following additional sensitivity analyses of the primary outcome were performed:
    - Linear regression weighted by number of CGM measurements recorded per patient (adjusted for randomisation strata)
    - Linear regression adjusted for time to first SG measurement and first SG value recorded (adjusted for randomisation strata)
    - Linear regression with an interaction between treatment and time to first SG measurement (adjusted for randomisation strata and first SG)
    - Linear regression with an interaction between treatment and first SG measurement (adjusted for randomisation strata and time to first SG measurement)
    - Generalised least squares model allowing for the variance to differ between treatment groups (adjusted for randomisation strata)
    - Random effects model with a random intercept for site (adjusted for gestation)
    - Linear regression model adjusting for gestation only
- The following post-hoc analyses were performed:
    - Linear regression model for the primary outcome with a three-way interaction between first SG, gender, and treatment group (adjusted for randomisation strata)
    - Linear regression model for the percentage of time SG <2.6mmol/l adjusted for randomisation strata
- A listing showing details of SG hypoglycaemic episodes (SG <2.6mmol/l) has been produced. This includes the following information:
    - Subject ID
    - Treatment group
    - Study day when SG <2.6mmol/l
    - Start time of hypoglycaemic episode
    - End time of hypoglycaemic episode

  - Length of episode (hours)
  - Time BG (blood gas or point of care) was taken (if any taken during the SG hypoglycaemic episode)
  - BG (blood gas and/or point of care) reading
- Analysis of the parent questionnaires was performed. The questionnaire is completed by parents of the CGM group only and consists of three Likert scale questions and two free text questions. The Likert scale responses were converted to an overall score ranging from 3 to 15 with 3 = positive response and 15 = negative response. The analyses were as for the staff questionnaires in Section 13.
- Listings of withdrawals from treatment only and withdrawals from the trial have been produced. These include the following information:
  - Subject ID
  - Treatment group
  - Date of randomisation
  - Date of withdrawal
  - Reason for withdrawal/description of withdrawal
  - Number of intervention days completed –withdrawal from treatment only
  - Later withdrawn from the trial (yes/no) – withdrawal from treatment only
  - Study timing (days 1 to 7/days 8 to 14/post day 14) – withdrawal from trial only
- A listing of deaths has been produced and includes the following information:
  - Subject ID
  - Treatment group
  - Date of randomisation
  - Date of death
  - Study timing (day 1 to 7/days 8 to 14/post day 14)
- The meta table has been updated:

meta_table_v0.5.csv

19 References

[1] K. Costeloe, A. Hennessy, T. Gibson, N. Marlow and A. Wilkinson, "The EPICure study: outcomes to discharge from hospital for infants born at the threshold of viability," *Pediatrics,* vol. 106, no. 4, pp. 659–71, 2000.

[2] S. Petrou, "The economic consequences of preterm birth during the first 10 years of life," *BJOG,* vol. 112, no. 1, pp. 10–15, 2005.

[3] D. Economides, A. Proudler and K. Nicolaides, "Plasma insulin in appropriate- and small-for-gestational-age fetuses," *American Journal of Obstetrics & Gynecology,* vol. 160, no. 5, pp. 1091–4, 1989.

[4] K. Beardsall, S. Vanhaesebrouck, A. Ogilvy-Stuart, C. Vanhole, C. Palmer, K. Ong and et.al, "Prevalence and determinants of hyperglycemia in very low birth weight infants: cohort analyses of the NIRTURE study," *Pediatrics,* vol. 157, no. 5, pp. 715–9, 2010.

[5] G. Gole, A. Ells, X. Katz, G. Holmstrom, A. Fielder, A. Capone Jr, J. Flynn, W. Good, J. Holmes, J. McNamara and E. Palmer, "The international classification of retinopathy of prematurity revisited," *JAMA Opthalmology,* vol. 123, no. 7, pp. 991–99, 2005.

[6] J. Center, "Conversion Table for Blood Glucose Monitoring," Joslin Diabetes Center, 2018. [Online]. Available: http://www.joslin.org/info/conversion_table_for_blood_glucose_monitoring.html. [Accessed 26 January 2018].

[7] Y. Benjamini and Y. Hochberg, "Controlling the False Delivery Rate: a Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society,* vol. 57, no. 1, pp. 289–300, 1995.

[8] G. Parry, J. Tucker and W. Tarnow-Mordi, "Crib II: an update of the clinical risk index for babies score," *The Lancet,* vol. 361, no. 9371, pp. 1789–1791, 2003.

[9] T. J. Cole and P. J. Green, "Smoothing reference centile curves: the LMS method and penalized likelihood," *Statistics in Medicine,* vol. 11, pp. 1305–19, 1992.

[10] C. M. Wright, A. F. Williams and D. Elliman, "Using the new UK-WHO growth charts," *BMJ,* vol. 340, p. 1140, 2010.

[11] T. Cole, "Too many digits: the presentation of numerical data," *Archives of Disease in Childhood,* vol. 100, pp. 608-609, 2015.

20 Listing of Tables, Listings and Figures

A listing of tables and figures can be found in the following document:

meta_table_v0.3.csv

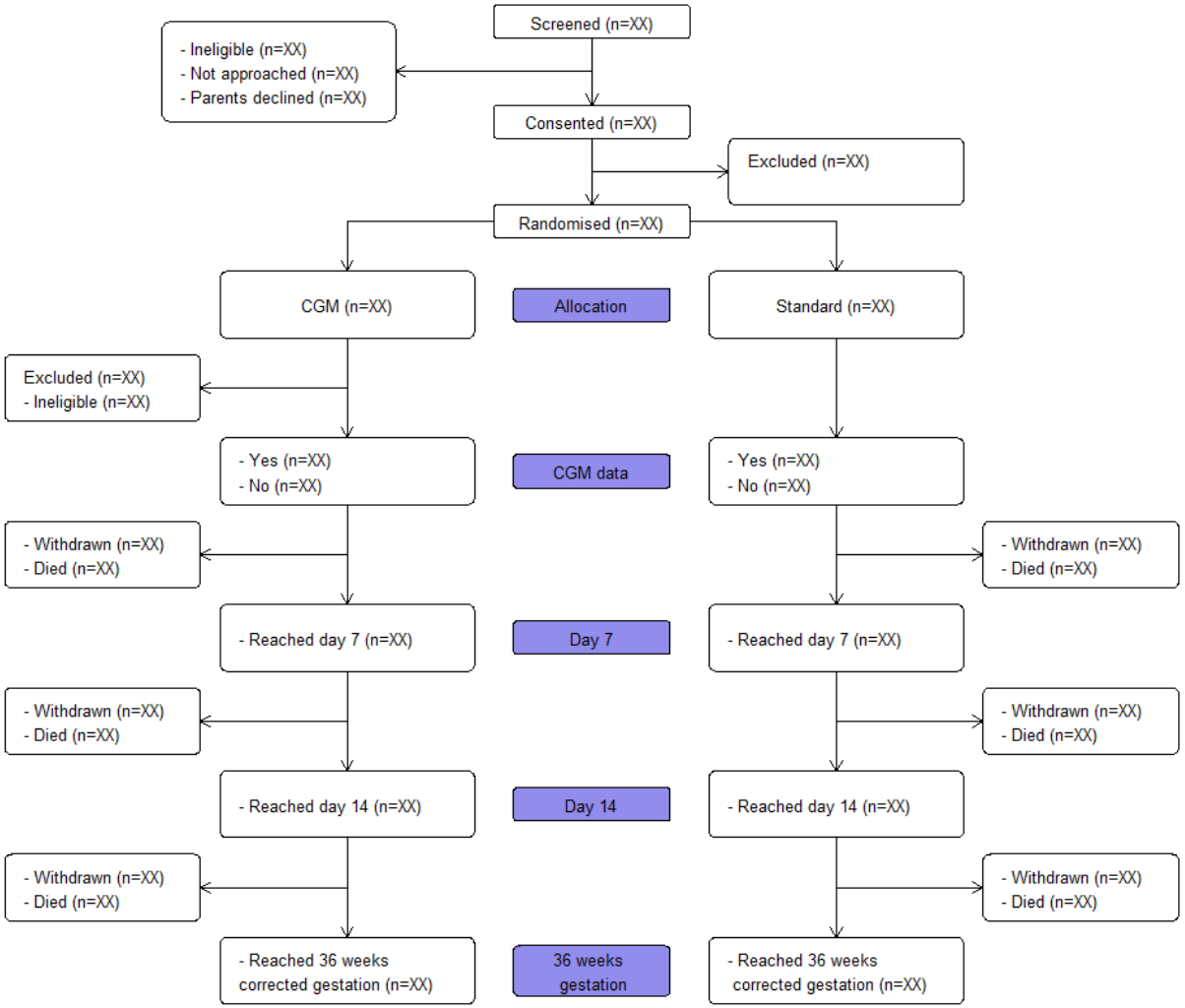

Figure 3: Revised CONSORT diagram