STATISTICAL ANALYSIS PLAN FOR THE TOBOGM STUDY

Trial registration number ACTRN12616000924459

Version: 1

Date: 02/09/2022

1 TOBOGM TRIAL

This is an RCT with open label among RCT cases, double blinding of controls and a parallel cohort of women with normal OGTT. Trial registration number: ACTRN12616000924459

2 DATA SOURCE

Data will be collected from 17 hospital sites in Australia, Austria, India, and Sweden.

2.1 Participating sites

Australian sites:
Campbelltown/Camden Hospitals, South Western Sydney, NSW
Liverpool Hospital, South Western Sydney, NSW
Bankstown-Lidcombe Hospital, South Western Sydney, NSW
Fairfield Hospital, South Western Sydney, NSW
The Canberra Hospital, ACT
Westmead Hospital, Western Sydney, NSW
Blacktown Hospital, Western Sydney, NSW
Nepean Hospital, Western Sydney, NSW
Royal Prince Alfred Hospital, Sydney, NSW
Monash Health, Clayton, VIC
Women's & Children's Hospital, Adelaide, SA
Flinders Medical Centre, Bedford Park, SA
Lyell McEwin Hospital, Elizabeth Vale, SA
Fiona Stanley Hospital, WA

International sites:
Vienna Hospital, Austria
University Hospital Örebro, Sweden
Dr.Mohan's Diabetes Specialities Centre, Chennai, India

3 ANALYSIS OBJECTIVES

Hypothesis 1: That diagnosis and treatment of pregnant women prior to 20 weeks gestation with hyperglycaemia consistent with GDM diagnosed at 24-28 weeks gestation but less than DIP (1), reduces adverse pregnancy outcomes (i.e. benefits women and/or their offspring)

Hypothesis 2: That diagnosis and treatment of pregnant women prior to 20 weeks gestation with hyperglycaemia consistent with GDM diagnosed at 24-28 weeks gestation but less than DIP, is associated with reduced fetal nutrition as reflected by reduced fetal lean body mass.

Overall aims:

1. [Observational: To define the prevalence and natural history of 'Booking GDM' and the clinical relevance of current diagnostic criteria.
2. [Clinical]: To test whether treatment of 'booking GDM' will reduce the sequelae of maternal 'hyperglycaemia' without increasing the risk of fetal under-nutrition.
3. [Translational/population health]: TOBOGM will address major evidence gaps in a key public health priority area and will directly inform clinical practice.

4 RANDOMISATION

The TOBOGM RCT tests whether GDM, diagnosed using criteria established for 24-28 weeks gestation and present at antenatal booking ("Booking GDM"), should be treated or whether the diagnosis decision can be deferred to 24-28 weeks. Women with GDM for early treatment (n=400) are unblinded but women in the deferred treatment arm (n=400) have their status masked, which is maintained by also including women without GDM (n=800) as "decoys" in the trial cohort. Decoys are randomly selected from the predicted 2400 women without "Booking GDM". The remaining 1600 women (approximately) are included for chart review but are not in the trial. Decoys undergo exactly the same procedures as women in the deferred treatment arm. Those with a fasting glucose ≥6.1 mmol/l and/or 2 hour glucose ≥11.1 mmol/l have been excluded from the RCT and referred directly for treatment for safety reasons.

Two criteria for GDM are able to be compared: one with an odds ratio of 1.75 for adverse outcomes (low) and one with an odds ratio of 2.0 for adverse outcomes (high) (2). The actual glucose values for the oral glucose tolerance test (OGTT) are kept blinded for both the antenatal booking and 24-28 week OGTTs. The TOBOGM RCT involves 2 randomisation processes, one of which is stratified by site (Decoy selection) and one of which is stratified by site and glycaemic band (High vs Low) (RCT).

Randomisation will be undertaken by minimisation within each site and stratum using a bespoke electronic randomiser (Techtonic, UK). The randomisation will be undertaken by staff separate from the data collection and analysis (ie blinded).

5 SAMPLE SIZE AND POWER

Power calculations were 2 tailed, alpha 0.05, power 0.8, 1:1 case:control. TOBOGM is powered to detect a 6% difference between the treated and untreated 'Booking GDM' groups. This is because the trial is not a test of treatment vs no treatment of GDM, but of treatment <20 weeks vs deferral to 24-28 weeks gestation when women will be treated if GDM is then found. It is therefore proposed to use the difference between the composite for treated GDM women compared to normal control: 10% vs 4%.

6 OUTCOMES AND ENDPOINTS

Outcome	Definition	Timepoint
Primary outcome [PO1]	**Primary pregnancy outcomes**: Composite of, respiratory distress (defined as need for at least 4 hours of respiratory support with supplemental oxygen, continuous positive airway pressure, or intermittent positive-pressure ventilation during the first 24 hours after delivery), need for phototherapy, birth trauma using IADPSG criteria, pre-term birth (<37 weeks gestation), stillbirth/death, shoulder dystocia (vaginal cephalic delivery that requires additional obstetric manoeuvres to deliver the fetus after the head has delivered and gentle traction has failed) and birthweight greater than or equal to 4.5kgs.	Birth
Primary outcome [PO2]	**Primary maternal outcome:** Incidence of hypertension in pregnancy (composite of pre-eclampsia/eclampsia/gestational hypertension). This will be collected from the notes	Collected any time between 34 weeks gestation and birth
Primary outcome [PO3]	**Primary neonatal outcome**: Fetal *lean body mass* measured by the Catalano equation (3) Derived from neonatal anthropometric measurements: ie birthweight-fat mass, where fat mass=0.54657+ 0.39055 (birthweight kg) + 0.0453 (flank skinfold mm)– 0.03237 (length cm).	Birth
Secondary outcome [SO1]	Neonatal *fat mass* from Catalano equation, Derived from neonatal anthropometric measurements: ie birthweight-fat mass, where fat mass= 0.54657+0.39055 (birthweight kg) + 0.0453 (flank skinfold mm)– 0.03237 (length cm)	Birth
Secondary outcome [SO2]	Ethnicity adjusted customised centile for birth weight (large and small), derived from data from notes (Gardosi) (4).	Birth
Secondary outcome [SO3]	Neonate mean upper arm circumference, measured within 72 hours postnatal.	Within 72 hours postnatal
Secondary outcome [SO4]	Neonate severe hypoglycaemia (heelprick glucose <1.6mmol/l neonate), derived from notes.	At any point up to 72 hours after birth
Secondary outcome [SO5]	Neonatal intensive care unit bed days, derived from notes	Collected at any point up to 28 days after birth
SO6	Sum of neonatal callipers, measured using callipers	Birth-72 hours postnatal
SO7	Maternal gestational weight gain, measured using seca scales.	End point is 36-38 weeks gestation. Baseline measure is from initial weight measurement on entry into the study
SO8	Caesarean section, from notes.	Birth
SO9	Induction of labour, from notes	Birth
SO10	Maternal hypoglycaemia, data collected from various sources including patient notes and meter downloads	Birth
SO11	Perineal trauma, from notes	Birth
SO12	Never breastfed, determined by bespoke study questionnaire	Up to 12 weeks postnatal, completed on a single occasion within this timeframe
SO13	Quality of Life (EQ5D) (calculated Index)	28 weeks gestation, 6-12 weeks postnatal
SO14	Birthweight, from notes	Birth
SO15	1-2 hour heelprick glucose less than or equal to 2.2mmol/l.	Birth

7 COVARIATES/CONFOUNDERS

We will adjust for potential confounders such as mother's age, BMI, smoking status. Covariates that are known to affect the outcome will be investigated for inclusion into the multivariable regression models.

Maternal covariates to be included are:

- M*aternal age (years)*
- *Ethnicity*
  - East Asian/SE Asian
  - Aboriginal/African/South American/Other
  - Middle eastern
  - South Asian
  - White European
  - Māori/Pacific Island descent
- *Primigravida Yes/No*
- *Maternal pre-pregnancy BMI (or from first antenatal weight if not known) (kg/m^2)*
- *Maternal smoking status (Current vs not)*
- *Maternal highest qualification* (University degree or higher vs not/not specified)

8 HANDLING OF MISSING VALUES

Exploring missing data mechanisms: The missing data mechanism process that is believed to have generated the missing values will be explored.

An appropriate analysis would be conducted to depending on the mechanism of missingness i.e. Missing Completely at Random (MCAR), Missing at Random (MAR) or Missing Not at Random (MNAR). Multiple imputation methods would be used to compare the results derived from the Intention to Treat (ITT) method.

8 STATISTICAL METHODOLOGY

Analyses will be by intention to treat with a per protocol sensitivity analysis. Descriptive analyses will summarise demographic characteristics of the included women in the data set and other key data. Initial analyses will involve computations of Fisher's exact tests (for dichotomous outcome measures), Chi-square tests (for categorical data with greater than or equal to 2 levels) and generalised linear mixed models (for continuous measures). Missing data will be examined and replaced with multiple imputation when indicated using the most suitable applicable method. Data will be described using 95% confidence intervals.

A series of univariate regression analyses will first examine associations between the outcome and covariates specified as fixed effects, and will also include (cluster/site) as a random effect to adjust for clustering. Continuous outcomes will be analysed using linear regression and binary outcomes using logistic regression. Non-normally distributed data will be analysed by either log

transformation of the outcome or non-parametric methods. Outcomes with multiple measurements during predefined time points would be analysed using generalised linear mixed models for repeated outcomes with the participant specified as a random effect.
A final multivariable model for each outcome will include covariates that have p-values <0.2 in univariate analysis and/or are highly relevant to the research question. Multivariable linear regression analysis is undertaken to remove confounding effects such as maternal smoking and BMI. All analyses will be adjusted for potential clustering by site by specifying site as a random effect in all models. Other factors entered into the models will be fixed effects. Robustness of the final model will be examined with bootstrapped samples of the same size with replacement.

9 MEASURES TO ADJUST FOR MULTIPLICITY

As stated above and in the protocol paper, there are three primary outcomes (PO1, PO2 and PO3). The study *apriori* significance level (alpha) is set at 0.05; and for the primary outcome analyses we apply a gatekeeping testing strategy to assist avoiding Type 1 errors (5). Specifically, if PO1 has p<0.05, then PO2 will be examined; next, if PO2 has p<0.05, then PO3 will be examined. This gateway approach to the primary outcome was deemed as the most suitable method as it enables a clear focus on major adverse pregnancy outcomes. This gatekeeper strategy was adopted after the publication of the protocol paper, after discussion and consensus reached among the CI team. This updated to the trial methods was registered with the ANZCTR record prior to the final data collection.

10. PUBLICATION PLAN

All secondary outcomes will be examined in exploratory analyses after the analyses with the primary outcomes are done. This means that the first study publication will report the primary outcome findings, and a separate second publication will report the exploratory analyses for the secondary outcomes.

11 PER PROTOCOL, SENSITIVITY ANALYSIS

A per-protocol analysis would be conducted to compare the results obtained by intension to treat analysis (6). Robustness of the final models will be examined with bootstrapped samples of the same size with replacement.

12 FURTHER EXPLORATORY SECONDARY ANALYSIS

Exploratory secondary analyses:

1. Comparison within 2 glycaemia bands:
    a. High Risk Band (HB): fasting 5.3–6.0 mmol/L and/or 1 hour ≥10.6 mmol/L and/or 2 hour 9.0–11.0 mmol/L
    b. Low Risk Band (LB): Not HB and fasting 5.1–5.2 mmol/L and/or 1 hour 10.0–10.5 mmol/L and/or 2 hour 8.5–8.9 mmol/L
2. Comparison within 2 gestational week/Trimester bands:
    a. with an OGTT <14 weeks gestation
    b. with an OGTT >14 weeks gestation.

10 STATISTICAL PACKAGES

All analysis would be conducted using Stata v16 (Stata corp.) (7) and R statistical packages.

REFERENCES:

1. Rani PR, Begum J. Screening and Diagnosis of Gestational Diabetes Mellitus, Where Do We Stand. J Clin Diagn Res 2016;10(4): QE01–QE04.
2. International Association of Diabetes and Pregnancy Study Groups Consensus Panel. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes care 2010;33(3):676–82.
3. Catalano PM, Thomas AJ, Avallone DA, Amini SB. Anthropometric estimation of neonatal body composition. Am J Obstet Gynecol 1995;173: 1176–81.
4. GROW; Gestation Network; Birmingham, UK, www.gestation.net.
5. US Food and Drug Administration. Multiple endpoints in clinical trials: guidance for industry 2017. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm536750.pdf. Accessed September 4, 2022.
6. Gilbert PB, Shepherd BE, Hudgens MG. Sensitivity analysis of per-protocol time-to-event treatment efficacy in randomized clinical trials. J American Stat Assoc 2013;108(503):789–800.
7. StataCorp. 2017. Stata 15 Base Reference Manual. College Station, TX: Stata Press.