DiGest Pre-specified Statistical Analysis Plan **Version 2; 18 Nov 2023.**

This is a randomised controlled trial of a nutritional intervention to identify if control of late gestational weight gain (GWG; 28 weeks to term) using a reduced energy diet can improve maternal and neonatal outcomes in gestational diabetes. The initial analysis plan for this study was assessed by Prof Vern Farewell, medical statistician in the MRC Biostatistics Unit (now retired). It was also externally peer reviewed with independent statistical input at the Diabetes UK research committee. This pre-specified analysis plan has been completed with input from Mr Stephen Sharp, medical statistician at the MRC Epidemiology Unit, Cambridge.

Planned analyses

A fundamental requirement for this study is to demonstrate a difference between total energy intake corrected for energy expenditure between control and intervention groups. It is anticipated that this will translate into a difference in late GWG between groups.

Women were recruited following an abnormal oral glucose tolerance test (OGTT) and those who consented were randomly assigned to receive either a 1200kcal or 2000kcal dietary intervention. Randomisation was performed independently of the study team. Randomisation was implemented using the library 'blockrand' in the statistical package R. The randomisation was done in permuted blocks of size 6, stratified by centre by Prof Farewell.

The following outcomes will be assessed:

Neonatal primary outcome:

- Standardised birth weight (standardised for infant sex and gestational age at delivery)

Maternal primary outcome:

- Maternal weight change (28-36 weeks)

Secondary outcomes: Secondary outcomes will be reported in line with the core outcome set for diabetes in pregnancy (Egan et al., 2020). Some secondary outcomes are no longer achievable due to changes during covid-19 (when sampling had to change to dried blood spots rather than plasma, and there was reduced access to cord blood, placental tissue and amniotic fluid).

For secondary outcomes related to pregnancy events, the baseline data will be that collected at enrolment, ~28 weeks. For secondary outcomes related to postpartum health, the baseline data will also be that collected at enrolment, ~28 weeks.

The most important secondary outcomes which are still achievable and will be included in the main paper are included below:

Key secondary outcomes:

- Maternal time in range (TIR) on continuous glucose monitoring (CGM) at 36 weeks (range 3.5-6.7 mmol/l)
- Maternal average glucose on CGM at 36 weeks.

Other maternal secondary outcomes:

- Caesarean section
- Requirement for medication (insulin/metformin) treatment by 36 weeks' gestation.
- Maternal weight/BMI at 3 months postpartum.
- Maternal glycaemia (HbA1c and CGM metrics) at 3 months postpartum.
- Maternal blood pressure at 36 weeks' gestation and 3 months postpartum

Neonatal secondary outcomes:

- Large for gestational age (using local, national and international centiles and customised centiles)
- Cord blood C-peptide
- Neonatal admission to the neonatal intensive care unit (NICU)
- Estimated gestational age at birth

Additional safety-related analyses:

- Small for gestational age (SGA) and intrauterine growth restriction (IUGR)
- Adverse pregnancy outcomes
    - Stillbirth
    - Neonatal death
    - Major congenital anomaly

Analysis methods – Primary Outcomes

The analysis is based upon the intention-to-treat principle, whereby all participants with available data will be included in the group to which they were randomised, regardless of their level of compliance.

The data analysis plan was designed in consultation with Mr Stephen Sharp and the Trial Steering Committee and follows advice from the EMA (https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-adjustment-baseline-covariates-clinical-trials_en.pdf). "No more than a few covariates should be included in the primary analysis. Even though methods of adjustment, such as analysis of covariance, can theoretically adjust for a large number of covariates it is safer to pre-specify a simple model. Results based on such a model are more likely to be numerically stable, the assumptions underpinning the statistical model are easier to validate and generalisability of the results may be improved." (Section 6).

In line with recommendations, one or two pre-specified adjustments can be useful. We considered that for analysis of each variable, adjustment for study centre and baseline measurement of that variable (at enrolment ~28 weeks; where available) were the most appropriate adjustments.

For the primary neonatal outcome, we will estimate the difference in standardised birthweight between the intervention and control group using linear regression with adjustment for the stratification variable of study centre.

Example Stata code: regress [stdbirthweight] i.trialarm i.studycentre

For the primary maternal outcome, weight change from enrolment to 36 weeks, we will use linear regression adjusted for study centre and baseline weight.

Stata code: regress [maternalweightchange] i.trialarm baselineweight i.studycentre

The primary analysis reported in the main DiGest publication will include all participants with data available for the relevant outcome at 36 weeks. For the change in maternal weight outcome, any participants with missing data at baseline will be included using the missing indicator method (White & Thompson, Stat Med 2005; 24: 993-1007). Any participants with missing outcome data at 36 weeks (e.g. due to withdrawals, deliveries before 36 weeks, and occasional unavoidable missed visits during Covid-19) will be excluded.

Example Stata code for missing indicator method:

gen baselineweight_star=baselineweight

summ baselineweight

replace baselineweight_star=r(mean) if baselineweight==.

gen baselineweight_miss=(baselineweight==.)

regress [maternalweightchange] i.trialarm baselineweight_star i.baselineweight_miss i.studycentre

If the data show >10% of women included in the primary analysis have had <4 weeks' exposure to the intervention, a sensitivity analysis for the primary outcomes will also be performed as follows:

- Including women with early deliveries, with adjustment for gestational age at delivery
- Including all recruited participants with multiple imputation.

Missing data analyses will be focussed on the primary outcomes (i.e. birthweight and change in maternal weight).

1. We will report the % of missing data for each primary outcome within each randomised group.
2. We will summarise key baseline characteristics separately in those with and without missing data for each primary outcome (included as an appendix)
3. Analyses of all primary and secondary outcomes will be complete case analyses, i.e. excluding individuals with missing data.
4. For each primary outcome, if there is >10% missing data then a sensitivity analysis using multiple imputation will be performed.

Analysis methods – Secondary Outcomes

Secondary outcomes will be analysed using linear/logistic regression adjusted for study centre and baseline variable, where relevant. For example, maternal systolic blood pressure at 36 weeks will be adjusted for study centre and baseline maternal systolic blood pressure:

Code: regress [outcome] i.trialarm baselinemeasure i.studycentre

Code: logistic [outcome] i.trialarm baselinemeasure i.studycentre

Since gestational age at birth is a confounder to some neonatal outcomes, and on the causal pathway for others, a further sensitivity analysis will be performed for neonatal outcomes using only pregnancies which delivered at term (at $\geq$37 weeks' gestation).

A number of safety analyses will also be performed to compare rates of SGA, stillbirth, maternal death and neonatal death between groups. Data will be presented as n(%). Consent was taken to collect and analyse limited data for women who withdraw from the study, for example, to allow assessment of reasons for withdrawal and to highlight any safety issues.

References

Egan AM, Bogdanet D, Griffin TP, Kgosidialwa O, Cervar-Zivkovic M, Dempsey E, Allotey J, Alvarado F, Clarson C, Cooray SD, de Valk HW, Galjaard S, Loeken MR, Maresh MJA, Napoli A, O'Shea PM, Wender-Ozegowska E, van Poppel MNM, Thangaratinam S, Crowther C, Biesty LM, Devane D, Dunne FP; INSPIRED research group. A core outcome set for studies of gestational diabetes mellitus prevention and treatment. Diabetologia. 2020 Jun;63(6):1120-1127. doi: 10.1007/s00125-020-05123-6. PMID: 32193573.

Hernan MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. N Engl J Med 2017;377(14):1391-1398. DOI: 10.1056/NEJMsm1605385