Statistical Analysis Plan
Randomized Trial of Maternal Progesterone Therapy to Improve Neurodevelopmental Outcomes in Infants with Congenital Heart Disease.

J. William Gaynor, MD, Principal Investigator
Mary Putt, PhD, PhD, ScD, Statistician

ClinicalTrials.gov identifier (NCT number): NCT02133573

Sponsor: Children's Hospital of Philadelphia

Version Date: 3/7/2022

Design: Randomized phase II 2-arm trial design
Control: Vaginal lubricant, twice a day
Intervention: Vaginal gel, 90mg twice a day

Consort Diagram: Per recommendations. http://www.consort-statement.org/consort-statement/flow-diagram

Primary Endpoint

The primary endpoint is the Motor Scale of the Bayley-III at 18 months of age.

Secondary Endpoints (Pre-specified in the protocol)[1]

Secondary endpoints pre-specified in the protocol are:

1. Developmental (Bayley-III at 18 months of age)
    a. Cognitive Scales
    b. Language Scales
    c. Fine motor score
    d. Gross motor score
    e. Receptive Communication
    f. Expressive Communication
2. Adherence to study protocol
    a. Progesterone regime (Average adherence by medication return and by diary)
    b. Complete pre- and post-natal MRI's
    c. Complete 18-month neuro-developmental evaluation
3. MRI-based
    a. Total maturation score (TMS) on Pre-operative MRI
    b. Brain volumes on Preoperative MRI:
        i. Whole Brain
        ii. Cerebellum
    c. White Matter Injury
        i. Volume on early post-natal brain MRI
        ii. Volume on early post-operative brain MRI
        iii. Volume change from pre-operative to post-operative brain MRI.

'Secondary' endpoints (Post Protocol) As clinical practice and research findings regarding progesterone use during pregnancy evolved over the course of accrual to the trial, these new secondary endpoints were identified. They were prespecified prior to closure of the database, and while the principal investigator and research team, other than Dr. Putt and Ms. Choi (statistical analyst), were blinded. They include;

1. Fetal Variables: Fetal echocardiogram and OB ultrasound Doppler flows (umbilical artery, uterine artery, middle cerebral artery, oligohydramnios, fetal growth restriction (BW < 10^{th} percentile))
2. Maternal Complications: (gestational hypertension, preeclampsia, abruption, , gestational DM, maternal tobacco use, maternal hypothyroidism)

[1] We anticipate the trial resulting in at least two publications related to endpoints. In advance of seeing the unblinded data we anticipate a publication describing in depth results related to the primary endpoint and adherence. A brief description of key MRI-based endpoints may be included. A second publication will focus on MRI-based endpoints.

3. Impaired maternal-fetal environment (presence of 1 or more of gestational hypertension, preeclampsia, gestational DM, maternal tobacco use, maternal hypothyroidism)
4. Delivery Complications (abruption, post-partum bleeding, DVT/PE, perinatal mood/anxiety disorder, wound complications (infection, breakdown))
5. Postmenstrual age of neonate at time of birth,
    a. Categorical: born earlier than 37 weeks postmenstrual age (PMA)
    b. Continuous: PMA
6. Type of Delivery (Vaginal, Elective C-section, Emergent C-section)
7. Birthweight (BW), absolute and as Z-score
8. Placental descriptors: (placental weight (PW), PW < 10^{th} percentile, ratio of BW to PW, presence of infarction, fetal vascular malperfusion, maternal vascular malperfusion, fetal vascular thrombi).
9. Brain growth and development. Trajectories of total and regional brain volumes and TMS from fetal to first post-natal MRI,
10. Post-operative seizures
11. Microhemorrhage and infarction
12. Age at post-operative MRI
13. Operative management variables (Cardiac operation during first admission y/n, age at first operation, cardiopulmonary bypass (CPB) y/n, weight at first operation, Clancy class, , deep hypothermia circulatory arrest (DHCA) time, Total support time (TST), antegrade cerebral perfusion time, need for extracorporeal membrane oxygenation (ECMO), additional operations with CPB, additional TST, additional DHCA)
14. Mortality (in-hospital, on-study and post-study)
15. Length of stay and selected complications for first surgical admission
16. Weight, length and HC at 18 months, absolute and as WHO percentile and Z-score

Intention to Treat Principle Unless otherwise specified, all analyses are conducted on an 'as-randomized' basis, irrespective of adherence to treatment. This analysis follows the intention to treat principle pre-specified in the protocol.

Descriptive Statistics (Intention to treat cohort)

1. Analyses will be done for the overall cohort as well as by primary cardiac diagnosis (HLHS, TGA, Other).

2. Baseline characteristics of mother: Using descriptive statistics (as appropriate, means median, number of subjects, proportion, standard deviation (SD), interquartile range (IQR)) to summarize demographic (Age of mother, race, ethnicity, education level of parents, clinical characteristics of mother) and clinical characteristics (Primary cardiac diagnosis, time to surgery, whether surgery was carried out) of the study cohort overall and by arm. Assess rates of missing data in baseline characteristics.

3. Determine rates of completion of primary endpoint of the study by arm. Describe causes of attrition including death (by arm). Kaplan-Meier plots of time-to-death by arm while on-study.

4. Characteristics of neonate. Type of birth (vaginal, emergency cesarean, planned cesarean), complications of pregnancy, progesterone level in cord blood, postmenstrual age of neonate, birthweight, sex, genetic anomaly, apo-lipoprotein E (*APOE*) genotype)

5. Primary & secondary outcomes. These outcomes will be described using summary statistics, and for continuous variables, graphically to visualize the distribution, overall and by arm.

6. Genetic anomalies. The protocol pre-specified inclusion of 22q11 individuals. Additional anomalies as genomic information has evolved will be described. Subjects will be classified as Abnormal, Suspect, or Normal based on examination by a dysmorphologist and genetic testing (WES, microarray, and other clinical testing). Subjects with 22q11 deletions will be classified as Abnormal.

Efficacy Analysis

1. *Primary Endpoint (Motor Scale, Bayley III).* The analysis will use a two-sided two-sample T-interval (estimation) or T-test (hypothesis testing). The confidence level will be set at 90% and the Type I error at 10%.
2. Normality will be assessed. A transformation, or alternatively a non-parametric test, will be used as appropriate.
3. Sensitivity of results to diagnosis group will be assessed by including diagnosis group as a stratification variable in a linear regression model.
4. *Continuous Secondary Endpoints* using the same analysis as the primary endpoint:
    a. Cognitive and Language Scales (Bayley III)
    b. TMS
    c. Brain volumes
    d. WMI volumes
    e. PVL volumes
    f. Adherence to study drug
5. *Binary Secondary Endpoints* will use an exact binomial confidence interval to estimate differences between groups and will be tested using Fisher's Exact test. The confidence level will be set at 90% and the Type I error at 10%. Binary outcomes include
    a. Presence of WMI pre- and post-operatively
    b. Completion of MRIs
    c. Completion of Bayley III at 18 months
6. *Per Protocol and As Treated*
    a. The analyses of the primary and secondary endpoints described above will be repeated with alternatives to the intention to treat cohort.
    b. A 'strict' *per protocol* analysis will be used, considering only those subjects who were eligible for the study and who received the treatment as specified in the protocol. This analysis will consider only those subjects who completed at least 90% of their doses at the time of study-based discontinuation (pre-planned at XX weeks or due to earlier labor and delivery). Adherence will be measured in two ways (1) number of medication doses returned and (2) mother's diaries of medication use. Differences in the two measures of adherence will be described.
    c. As a sensitivity analysis, a 'relaxed' *per protocol* analysis will follow (a) but use completion of 80% of doses instead of 90% of doses, as the cut-off.
    d. An *as treated* analysis will consider subjects based on their 'dose of progesterone' based on their assignment to control or intervention arm, as well as adherence to study drug in the progesterone arm. In this case we will include dose as a continuous covariate in a regression model, using linear regression for continuous outcomes or logistic regression for binary outcomes. We will assess interactions between dose and treatment arm. This analysis will estimate the effect of the treatment intervention conditional on different levels of adherence to study medication.

7. *Modeling to Adjust for Covariate Imbalance*. Linear or logistic regression analyses of the primary and secondary outcomes will be carried out with adjustment for baseline covariates of fetal brain maturation and development. The covariates will include type of CHD (primary cardiac diagnosis), race, ethnicity, maternal education and socioeconomic status, presence of 22q11 deletions, other identified genetic anomalies and (*APOE*) genotype.

8. Missing data and truncation by death
    a. In sensitivity analyses, multiple imputation will be used to impute missing outcome data using baseline covariates. Results will be compared to the primary analysis using the complete data.
    b. Some subjects will have missing data for the 18-month endpoint due to truncation by death. In sensitivity analyses, we will assess the impact of this source of missingness using principal stratification.

9. Subgroup analyses of Primary & Pre-specified Secondary Endpoints:
10. Primary cardiac diagnosis
11. With versus without 22q deletions
12. With versus without any genetic anomaly
13. Males versus Female infants
14. Subjects undergoing surgical repair or not
15. With/without impaired maternal-fetal environment as defined above

Safety Analysis

All subjects the study at will be included in the safety analysis. The frequencies of adverse event (AE) by type, body system, severity and relationship to study drug will be summarized for both the mother and the neonate. Serious adverse events (SAE), if any, will be described in detail.AE incidence will be summarized along with the corresponding exact binomial 95% two-sided confidence intervals.

Further analyses anticipated for additional paper(s)

Using linear regression, we will assess the association between the neurodevelopmental outcomes with:

1) MRI findings (microhemorrhage, gray and white matter volumes, white matter injury (WMI) presence and volume, infarction, volume and growth trajectory) to neurodevelopmental (ND) outcomes
2) post-operative seizures
3) maternal complications and baseline clinical variables including diabetes, hypertension, smoking status, pre-eclampsia, hypothyroidism
4) placental findings
5) impaired maternal-fetal environment

Using regression techniques, including logistic regression for white matter injury (WMI) and competing risk time-to-event analyses (length of stay) with death as a competing risk), we will assess the association between the outcome (WMI, length of stay) and

1) maternal complications and baseline clinical variables
2) placental findings
3) trajectories of brain growth