**Pre-Analysis Plan: Impact of Nurse-Family Partnership on Maternal and Early Child Outcomes**

*Note: This is version 2 of the pre-analysis plan reflecting plans for study design as of January 20, 2021. The first unblinding is planned for January 21, 2021 when interim analyses will be prepared for Pay-for-Success payments. The previous version of this pre-analysis plan (version 1) is available on our study website.*

A full study protocol describing study motivation, design, and plans for analysis has been published in McConnell et al. (2020).[1] We do not repeat those materials here. In this pre-analysis plan, we provide more detail about our planned analytical choices including analytic details for defining primary and secondary outcomes for the three domains of analysis defined in the study protocol in Addendums 1-3. This analysis plan will be updated as analytical plans for additional longitudinal outcomes are developed.

## 1. Methods

While other methods are outlined in the study protocol, in this pre-analysis plan we provide further details on methods related to sample construction and analytical plans.

### A. Sample construction

#### A.1. Tracking study moms and babies

All study outcomes will be measured using administrative data sources, which require matching to state administrative records. Because we plan to analyze outcomes for both the mother and child, we define several relevant samples for analysis of outcomes for the mother-child dyad. For most analyses, we will want to consider the sample of mother-baby dyads where a live birth occurred from the pregnancy in gestation at the time of enrollment. We call this the index birth. Mother-baby dyads who experience an index birth will be the primary sample for longitudinal analyses examining the impact of the program. As the mother is the unit of randomization and we expect multiple births to be rare, we plan to aggregate outcomes pertaining to children resulting from the index birth to the mother level. To identify the index birth, we first match the mother to a birth certificate in vital records. If the date of birth on the birth certificate is within 120 days before or after the estimated delivery date as reported on the baseline survey, we will consider the birth to be related to the pregnancy that was in gestation at the time of study enrollment. Births that occur outside of this window will not be included in the sample of index births.

We may also analyze outcomes for the broader sample of mothers with at least one matched administrative record related to the pregnancy reported at the time of enrollment. We will call this the index pregnancy. We define the index pregnancy as follows: when there is a matched

index birth or fetal death in vital records[i], we define the pregnancy period as the period covering the weeks of gestation provided in the obstetric estimate on the vital record until the date of birth or fetal death. In the absence of an index birth or fetal death, we define the pregnancy period as the period covering 42 weeks prior to the expected delivery date provided on the baseline survey. We will consider a mother to be matched to the index pregnancy if she has an any matched administrative records pertaining to the index pregnancy (including antenatal care, probable pregnancy loss or birth related records from Medicaid claims or hospital discharge or evidence of a birth or fetal death from vital records). Table 1 summarizes the definitions and data sources used to identify index pregnancies and live index births.

When tracking children in the study, we will define two key periods. We anticipate that the children of most mothers who enroll in the trial will be eligible for Medicaid for their entire first year of life, as the income requirements for infants are similar to those of pregnant women. However, as summarized in Table 2, in some years of the study, income requirements became more restrictive after the first year of the child's life. Furthermore, families whose children remain income-eligible into their child's second year of life must submit an annual review form to retain coverage. Therefore, we plan to measure outcomes for children resulting from index births in the study to two critical time points: through their first 12 months of life and through their first 24 months of life.

Table 1. Sample Tracking Definitions

| Sample tracking measure | Data Source(s) | Definition |
|---|---|---|
| Index pregnancy | Birth Certificates, fetal death records, Medicaid claims and eligibility, hospital discharge | Having matched administrative records for one or more of the outcomes listed below during the period of the index pregnancy. This period is defined as covering the weeks of gestation provided in the obstetric estimate on the vital record until the date of birth or fetal death, when a mother has a matched index birth or fetal death record. In the absence of an index birth or fetal death, we define the pregnancy period as the time covering 42 weeks prior to the |

[i] Specified in McConnell et al. (2020) as a death of the fetus occurring at or after 20 weeks of gestation.

| | | |
|---|---|---|
| | | expected delivery date provided on the baseline survey.<br>● A live birth certificate for the index birth,<br>● A Kuklina identified live birth in Medicaid claims or hospital discharge records occurring within 120 days of the expected delivery date<br>● A probable pregnancy loss in Medicaid claims or hospital discharge occurring during the index pregnancy<br>● A fetal death record in Vital Records occurring during the index pregnancy<br>● An antenatal care visit in Medicaid claims occurring during the index pregnancy<br>● Enrollment in Medicaid for pregnant women overlapping with the index pregnancy period |
| Index birth | Baseline survey, birth certificates | A birth identified by a matched birth certificate in Vital Records with a date of birth within 120 days before or after the estimated delivery date reported on the baseline survey. |

Table 2. South Carolina Medicaid income eligibility criteria as a percent of the Federal Poverty Line over the study period

| | Jan 2016 | Jan 2017 | Jan 2018 | Jan 2019 | Jan 2020 | Jan 2021 |
|---|---|---|---|---|---|---|
| Pregnant women[2,3] | 199% | 199% | 199% | 199% | 199% | 194% |
| Parents[4,5] | 67% | 67% | 67% | 67% | 67% | 95% |
| Children 0-1[3,6] | 213% | 213% | 213% | 213% | 194% | 208% |
| Children 1-5[3,7] | 213% | 213% | 213% | 213% | 143% | 208% |

| | | | | | | |
|---|---|---|---|---|---|---|
| Children 6-18[3],[8] | 213% | 213% | 213% | 213% | 133% | 208% |

### A.2. Characterizing missing data

We may be unable to track mother and child outcomes for pregnant women enrolled in our trial over the entire period spanning pregnancy and the child's first two years of life for various reasons. First, we may be unable to match the identifying information provided by women at study enrollment to administrative records because of poor data quality or because women enrolled in the trial moved out of state. Second, we may be unable to match women to their index birth because of problems linking mother and child records within South Carolina's administrative record system or because their pregnancy ended in pregnancy loss. We may be unable to observe child outcomes through 24 months if a child dies before reaching 24 months of age. For outcomes that can only be observed in Medicaid claims data, such as child preventative health outcomes and outpatient care utilization for mothers, we will not observe outcomes for mothers or children not enrolled in Medicaid. Whenever possible, we will report missing outcome data and characterize missing outcome data across each of these groups. More details are provided in Addendums 1-3 regarding plans to characterize missing outcome data related to each of the three domains of potential program impact and plans for sensitivity analyses to account for potential patterns in missing outcome data.

Because administrative records update continuously over time, we anticipate that mothers and babies classified as falling within our study sample definitions may change as administrative data is updated. For example, a mother enrolled in the trial whom we cannot initially match to any records may later match to state Medicaid records. Furthermore, administrative data may experience delayed reporting or data errors that are corrected later. We may revisit analyses several years after completion to assess the robustness of results to updated administrative records.

## B. Analytical considerations

### B.1. Control variables

In addition to reporting unadjusted models of program impact, we plan to report models that adjust for characteristics at baseline that may be associated with our primary academic outcomes. These include participant age, race and ethnicity, gestational age at study enrollment, relationship with the father of the child, education, employment, use of social services, housing stability, health care utilization, health behaviors and physical and mental health status all measured during the baseline survey. Specific variable definitions that we plan to include as control variables are specified below. Binary variables will be coded as “1”, “0”, or missing; for coding categorical variables, we will consider distributions across the control and

treatment groups including missing values. We plan to use the same set of control variables in all analyses.

- Implementing Agency
    - Indicators for each implementing agency enrolling study participants
- Demographics:
    - Indicator for age equal to 15, 16, or 17
    - Indicator for age equal to or greater than 28
    - Race – indicators for non-Hispanic Black, non-Hispanic white, other
    - Ethnicity – indicator for Hispanic/Latina
- Gestational age at time of study enrollment
    - A continuous variable of weeks to delivery as calculated by difference in estimated due date and the survey date
- Relationship with father of the child
    - Indicator for daily interaction with father of the child
- Education
    - Indicator for high school diploma or GED with no higher education
    - Indicator for less than high school diploma
- Employment, Income, and Financial Resources
    - Indicator for whether the mother is working for pay at the time of the survey
- Social services
    - Indicator for receiving one or more social service (i.e. TANF, SNAP, SSI, unemployment benefits, and WIC)
- Housing stability
    - Indicator for moving two or more times in the previous twelve months
    - Indicator for living with parents
- Access and utilization of health care, including mental health and maternal health
    - Indicator for receiving at least one antenatal care visit before time of survey
    - Indicator for having obtained care at a hospital ER in past six months
    - Indicator for receiving mental health treatment in the previous year
- Health behavior (e.g. drinking and smoking)
    - Indicator for reporting having consumed alcohol in the three months before pregnancy
    - Indicator for reporting having smoked cigarettes in the three months before pregnancy
- Psychological state/resources (measured on a PHQ-2 scale)
    - Indicator for PHQ-2 score of 3 or higher.
- Baseline measure of self-reported health

    - Indicator for self-reported health described as fair or poor.
- Maternal perceived stress level
    - Indicator for PSS-4 score of 4 or higher.
- Self-reported pregnancy risk factors
    - Indicator for pre-pregnancy weight and height yielding BMI of normal.
- Family planning
    - Indicator for responding yes to having access to a place for family planning or birth control
    - Indicator for responding yes to wanting more children one day

### B.2. Missing values of control variables

We will treat baseline responses as missing if sample members responded to questions with "Don't know" or "Refused to Answer." We will use the dummy-variable adjustment method to account for missing baseline covariates in our analysis.[9] Specifically, we will set missing covariate values to a constant value and add indicators (i.e., dummy variables) for missing values to the impact analysis model. We will also assess our results' robustness to a model specification without baseline covariates. Finally, based on the prevalence and patterns of missing covariate data, we may specify a multiple imputation model to impute missing values.

### B.3. Alternate Empirical Specifications

Per McConnell et al. (2020), we will estimate intent-to-treat (ITT) effects as our primary empirical specification.[1] We may also consider secondary specifications that incorporate data on actual program participation, taking advantage of the randomization as an instrument for participation to estimate local average treatment effects and to examine average characteristics of those participating in the program. We anticipate that the primary alternative definitions of interest for program participation will be based on reaching the milestone of participating in home visiting through the period surrounding the expected delivery date, the child's first birthday and the child's second birthday. We would operationalize participation through the expected delivery date and the child's first birthday as mothers who receive a nurse visit within 14 days of the milestone. Participation through the child's second birthday would be operationalized as a visit within 1 month (31 days) of the milestone. In accordance with the NFP program model, mothers should receive a visit approximately once every two weeks in the period leading up to delivery through the first 21 months of the child's life and should receive a visit once every month in months 21-24 of the child's life.

# Nurse-Family Partnership Evaluation Analysis Plan Addendum 1: Pregnancy, Birth and Maternal Health Outcomes.

*Note: This version of Addendum 1 reflects plans for study design as of January 20, 2021. The first unblinding is planned for January 21, 2021 when interim outcomes will be analyzed for Pay-for-Success payments.*

The goal of this analysis plan addendum is to enumerate analysis specific to the pregnancy, birth, and maternal health outcomes domain.

## 1. Methods

A full treatment of our empirical approach, primary study outcomes, planned subgroup analyses, and statistical power is given in sections 2.4-2.10 of McConnell et al. (2020).[1] The following sections provide additional context and details on planned analyses for the pregnancy, birth, and maternal health outcomes domain of the NFP evaluation.

### 1.1. Defining the Sample for Analysis

For analysis of our primary adverse birth outcome, we will consider our primary sample to be mothers who experience either a fetal death or index birth (defined in section A.1 of the pre-analysis plan). For analysis of other outcomes, our primary sample will be mothers with an index birth, as many outcomes will not be observable in the case of fetal death.

### 1.2. Construction of Study Outcomes

We use a combination of South Carolina vital records, Medicaid claims, and hospital discharge records to construct study outcomes. In cases where outcomes differ between records, (e.g. gestational age at delivery, birthweight) we will use estimates from South Carolina vital records.

#### 1.2.1 Primary Outcome

Our primary outcome for this domain is an adverse birth outcome, which we define as having a preterm birth (less than 37 weeks' gestation), a newborn being small for gestational age (less than 10th percentile of US births conditional on gestational age), having low-birth weight (less than 2500 grams), or experiencing perinatal mortality (fetal death occurring at or after 20 weeks' gestation or mortality in the first 7 days of life). Data for the adverse birth outcome will come from South Carolina birth records, death records, and fetal death records.

For mothers with multiple births, we define the outcome based on having an adverse birth outcome for any child from the index pregnancy. While preterm birth and other adverse birth outcomes may be more common among multiple births, we anticipate that rates of multiple births will be equal across treatment and control arms. We may also explore alternative specifications of this outcome that include only singleton births.

### 1.2.1 Secondary Outcomes

We will examine several secondary infant and maternal outcomes. For infant outcomes, we will examine each adverse birth outcome individually (SGA, preterm, low birth weight, perinatal mortality). We will also examine large-for-gestational age (>90$^{th}$ percentile of US births conditional on gestational age), very low birth weight (<1500g), a continuous measure of birth weight, and extremely preterm (<28 weeks gestation) using birth certificate records. We will also examine neonatal morbidity, defined as assisted ventilation immediately after delivery, assisted ventilation for more than six hours, seizure, receipt of surfactant replacement therapy, or receipt of antibiotics for suspected sepsis using data from the birth certificate.[ii,10,11] Finally, we will examine NICU admission of at least overnight, defined as a claim with a procedure code of 99468, 99469, 99477, 99478, 99479, or 99480 on the day of delivery and also on the following day.

We will also examine several maternal health related secondary outcomes at birth, including cesarean delivery (from the birth certificate) and severe acute maternal morbidity from hospital discharge and Medicaid data (as defined by the Centers for Disease Control and Prevention).[12] We will examine maternal mortality up to one year after birth using vital records. We will examine receipt of any postpartum visit within the first 12 weeks postpartum using Medicaid claims data, defined as a claim with a diagnosis code of Z39 or a HCPCS code of 59430.

Next, we will examine the outcome of neonatal abstinence disorder and/or maternal drug/substance abuse both during the index pregnancy period and over the period spanning the index pregnancy and the first 24 months following the index birth. This outcome is defined as a composite outcome. While the outcome is only specified over the longer time-period in our protocol and trial registry, exploratory analysis of this outcome during the index pregnancy will be valuable because of the strong relationship between drug use and birth outcomes.[13] We will define the composite outcome related to substance abuse as any record in the Medicaid claims or hospital discharge data with a code listed in Table 1.1. We will also examine robustness to a definition in which we exclude the tobacco codes which we expect to make up the majority of claims we observe matching to maternal substance use.

Next, we will examine maternal experience of violence or homicide during the index pregnancy and during the first 24 months after the index birth using a combination of death records and Medicaid and hospital discharge data. This is defined as any record matching a code listed in

[ii] In our study protocol (McConnell et al 2020) we indicated that we planned to derive neonatal morbidity from discharge records. However, in order to make our analyses parallel with other analyses looking at neonatal morbidity (Han et al 2020) and because of concerns about the ability to translate these concepts directly into claims data (Ford et al 2007), we plan to rely instead on birth certificate data.

Table 1.2. As above, experience of violence was originally specified over 24 months, but we will explore analysis of this outcome during the index pregnancy because of the negative associations between exposure to maternal violence and birth outcomes.[14] Our current definitions of substance abuse and maternal experience of violence are limited to the codes defined here; however, we will explore incorporating broader indicators of challenges with substance abuse derived from pharmacy claims, child claims, and criminal justice records. We will update these outcome specifications prior to any analysis of comparisons across treatment and control groups for these outcomes.

We will also examine use of social services during pregnancy and during the first 24 months after delivery. During pregnancy, we will examine any receipt of WIC benefits (derived from the birth certificate) and any receipt of Supplemental Nutrition Assistance Program (SNAP) benefits (from South Carolina Department of Social Services data). During the first 24 months after delivery, we will examine total months of receiving SNAP benefits and Temporary Assistance for Needy Families (TANF) benefits, as well as a measure of benefit churn, defined as receiving SNAP or TANF benefits at any time during a given year and having experienced at least one break in participation of four months or less that started and/or ended during the year.[15]

Table 1.1. Codes for composite outcomes related to neonatal abstinence disorder and maternal drug/substance abuse

| Code Description | ICD-10 Code(s) |
|---|---|
| Neonatal abstinence disorder | P96.1, P04* |
| Opioids | F11.* |
| Tobacco | F17.2*, O99.33*, Z72.0 |
| Alcohol | V11.3*, F10.* |
| Sedative | F13.* |
| Cocaine | F14.* |
| Amphetamines | F15.* |
| Cannabis | F12.* |

Source: Jarlenski et al. 2020[16]

Table 1.2. Codes for maternal experience of violence

| Code Description | ICD-10 Code(s) |
|---|---|
| Adult neglect | T74.01; T76.01 |
| Adult physical abuse | T74.11; T76.11 |
| Adult sexual abuse | T74.21; T76.21 |
| Adult psychological abuse | T74.31; T76.31 |
| Unspecified adult maltreatment | T74.91; T76.91 |

| | |
|---|---|
| Husband, perpetrator of maltreatment | Y0701 |
| Assault by unarmed brawl | Y040 |
| Observation after rape | Z0441 |
| Assault by other bodily force | Y048 |
| Unspecified perpetrator of maltreatment<br>Other family member perpetrator of maltreatment | Y079<br>Y07499 |
| Encounter for mental health services for victim of spousal or partner abuse | Z691x |
| Encounter for mental health services for victim or perpetrator of other abuse | Z698x |
| Encounter for observation following alleged adult physical abuse | Z0471 |

Sources: Crosswalk of ICD-9 to ICD-10 codes from Davidov et al. 2017 and Schafer et al. 2008.[17,18]

Antenatal Care Utilization & Quality

Analyses of outcomes related to antenatal care utilization and quality is planned as a stand-alone analysis. We will examine three groups of outcomes related to antenatal care utilization and quality: utilization of health care services during pregnancy, guideline-recommended antenatal care, and indicators related to antenatal health. We list the outcomes that fall under each area, and their associated data sources, below. We also include the CPT/HCPCS, NDC (National Drug Code) and ICD-10 codes used to identify the outcomes that use Medicaid claims or hospital discharge data in Table 1.3.

The primary sample for this analysis will be mothers with an index birth. We will use vital records, hospital discharge, and Medicaid claims data for the analyses. Mothers who do not match to one of these data sources will not be included in the sample for the corresponding outcomes (e.g., mothers who do not match to Medicaid claims will not be included in the analyses of guideline-recommended antenatal care).

Utilization of health care services during pregnancy

- Adequacy of Prenatal Care Utilization Index (APNCU) [vital records data]
- Number of emergency department visits during pregnancy [all-payer hospital discharge data]
    - To better understand the effect of NFP on emergency department utilization, we may also examine treatment vs. control differences in emergency department use by trimester of the emergency department visit and by the enrollee's overall adequacy of prenatal care (APNCU).
- Number of ultrasounds during pregnancy [Medicaid claims data]

- Consultation with maternal fetal medicine specialist [all-payer hospital discharge data]
- Dental visit during pregnancy [Medicaid claims data]

Guideline recommended antenatal care

- Anatomy scan (between 18-22 weeks' gestation) [Medicaid claims data]
- Gestational diabetes test (between 24-28 weeks' gestation) [Medicaid claims data]
- TDAP vaccine (between 27-36 weeks' gestation) [Medicaid claims data]
- Group B strep test (between 35-38 weeks' gestation) [Medicaid claims data]

Antenatal health

- Smoking cessation during pregnancy [vital records data]
    - We will consider an enrollee to have ceased smoking during pregnancy if their vital record indicates that they smoked pre-pregnancy and that they did not smoke during pregnancy.

- Recommended gestational weight gain [vital records data]
    - Gestational weight gain will be classified as "recommended" if it is within the guidelines published by the National Academy of Medicine (formerly the Institute of Medicine).[19]

The guideline-recommended antenatal care services should be received at the appropriate gestational age during the pregnancy. The gestational age at the time of the test will be calculated as the date of the test minus the approximate date of last menstrual period (in days). We will approximate the enrollee's date of last menstrual period using the child's date of birth and obstetrician's estimate of gestational age at delivery from the vital records. If these fields are missing, we will define the approximate date of last menstrual period as the beginning date of the index pregnancy (defined above).

Each of these outcomes can only be measured among pregnancies that reach the gestational age at which the care is recommended. The main analyses will only include pregnancies that last at least as long as the gestational age recommended for each outcome. However, it is possible that NFP has a direct effect on gestational age, which would mean that the sample of mothers observed could differ across treatment and control arms. If that is the case (i.e., if gestational age at birth between treatment and control group is statistically significantly different), then we will conduct a bounding exercise as a robustness check for the guideline-concordant care outcomes.[20] Using Group B Strep as an example, the main analysis will compute the average treatment effect (ATE) among people whose pregnancies last at least 35 weeks. If gestational age is statistically significantly different between control and treatment arms, we will estimate two bounds on this ATE: (1) a recalculated ATE assuming *all* people who

did not reach 35 weeks' gestation would have received the test, and (2) a recalculated ATE assuming *none* of the people who did not reach 35 weeks' gestation would have received the test.

Table 1.3. Codes for analyses of antenatal care utilization and quality

| | CPT/HCPCS | NDC | ICD-10 |
|---|---|---|---|
| Dental visit | D1110, D0120, D0140, D0150, D0160, D0170, D0191, D1206, D1208, D02*, D0330, D0340, D0350, D2391, D2392, D2393 D1351, D71*, D7210, D7220, D7230, D7240, D7241, D7250 | | |
| Ultrasound | 76801, 76805, 76811, 76813, 76815, 76816, 76817, 76818, 76819, 76810, 76812 | | |
| Anatomy scan ultrasound (18-22 weeks) | 76805, 76811, 76815, 76816, 76817, 76810, 76812 | | |
| Gestational diabetes test (24-28 weeks) | 82950, 82951, 82947 | | |
| TDAP vaccine (27-36 weeks) | 90696, 90697, 90698, 90700, 90701, 90714, 90715, 90471, 90472, 90460, 90461 | 49281040010, 49281040015, 49281040020, 58160084211, 58160084252 | |
| Group B streptococcus test (35-38 weeks) | 87150, 3294F, 87802, 87653, 87801, 87081, 87084, 87070, 87077, 87147 | | Z36.85 |

Maternal utilization of mental health services

Analyses of outcomes related to maternal utilization of mental health services is also planned as a stand-alone analysis. NFP guidelines require nurses to screen for depression and anxiety at pre-specified intervals throughout the program, and to follow their agencies' protocols for referral and care coordination for women who screen positive. While directly measuring the

prevalence of depression and anxiety in our study sample would help us assess the impact of NFP on mental health, our study design relies exclusively on administrative data sources for outcome measurement. Therefore, our study design will seek to analyze how NFP changed the utilization of mental health services during the perinatal period.

We will include all women with an index birth in our analytic sample, regardless of whether they had a prior mental health diagnosis before pregnancy. Because few women in our study are enrolled in Medicaid prior to pregnancy, we cannot observe prior diagnoses or identify treatment initiation. However, we expect rates of prior mental health diagnoses before pregnancy to be equal across treatment and control arms. Because the outpatient mental health treatment outcomes rely exclusively on Medicaid claims, we will restrict the analytic sample for outpatient outcomes to those women who retain full Medicaid coverage through 60 days postpartum. We will not restrict the sample for inpatient mental health outcomes as these rely on utilization that will be observable in all-payer hospital discharge data, regardless of payer.

While the medical definitions for depression and anxiety are distinct, we examine them together because they are comorbid conditions[21,22] and NFP screens for them simultaneously. We define outpatient mental health treatment as a composite outcome of either a diagnosis for depression/anxiety/stress reaction or a filled prescription for antidepressants/anxiolytics or an outpatient psychotherapy visit during pregnancy or the first 60 days postpartum. We list the ICD-10 diagnosis codes, therapeutic class codes, and Current Procedure Terminology (CPT) codes used to define these outcomes in Tables 1.4, 1.5, and 1.6. Consistent with prior research[23], we consider all outpatient psychotherapy visits, including individual, family or group therapy to be mental health visits.

Table 1.4. Diagnosis codes for depression/anxiety/stress reaction

| Code description | ICD-10 Code(s) |
|---|---|
| Depression | F53; O906; F99; O9934*; F32*; F33* |
| Anxiety | F41.* |
| Stress-reaction | F43.* |

Table 1.5. Therapeutic class codes for filled antidepressants or anxiolytics

| Code description | Therapeutic class codes |
|---|---|
| Antidepressants | 281604 |

| | |
|---|---|
| Anxiolytics | 2824* |

Table 1.6. Outpatient psychotherapy codes

| Code description | CPT codes |
|---|---|
| Psychotherapy (individual, family or group) | 90804-90815; 90832-90834; 90836- 90840; 90845-90847; 90849; 90853; 90857; 90862; 90875; 90876 |

We will also look at these outpatient measures individually (i.e., diagnosis, filled prescription, and psychotherapy) and in combination to enhance our understanding of treatment patterns. For example, we may be interested in the proportion of mothers who receive a diagnosis but do not fill a prescription or receive psychotherapy as this could signal structural or individual-level barriers to treatment.[23] We may also be interested in the proportion of women who receive medication only, as psychotherapy is often recommended as a first-line treatment for mild or moderate depression and in combination with medication when the depression is severe.[24] We may also explore the timing of outpatient mental services (measured in days). NFP could lead to earlier treatment of perinatal depression/anxiety through regular screening, coordinating care, providing a warm hand-off, or by reducing the stigma surrounding mental health and treatment.

For most of the outcomes related to outpatient mental health treatment, we restrict the measurement period to pregnancy through 60 days postpartum because these outcomes are derived from Medicaid claims data and maintenance of coverage past 60 days may be affected by participation in NFP. However, depression and anxiety can take longer than 60 days to manifest, and even longer for the mother to access healthcare. In future analyses, we will consider measuring outpatient mental health treatment patterns through 6, 12 and 24 months postpartum if treatment and control group members are equally likely to be enrolled in *full-coverage* Medicaid plans through these time points.

In our clinical trials registry, we defined a measure of mental health treatment follow-up care as: a second antidepressant prescription or outpatient mental health visit within 120 days of the initial treatment ("acute phase"). However, as noted above, differential Medicaid coverage rates between the treatment arms after 60 days postpartum may preclude us from measuring the treatment effect for this outcome reliably. For example, if treatment group members are more likely to be enrolled in Medicaid after 60 days postpartum than control group members, observed differences in mental health treatment patterns between the treatment and control groups could be due to group differences in the composition of who remains on Medicaid.

To assess whether attrition from full-coverage Medicaid plans affects the comparability of our treatment and control groups, we will assess overall attrition and differences in attrition between the treatment and control groups at the following time points: 60 days (for our main outpatient mental health treatment outcomes), 6 months, 12 months and 24 months postpartum (for other potential measurement periods of interest). We do not consider enrollment in partial-coverage Medicaid plans (such as Family Planning) since these plans cover mental health screenings, but do not reimburse for medications or psychotherapy. Specifically, we will compare partial and continuous enrollment rates in full coverage plans between treatment and control groups at each time point. Continuous coverage will be measured using Medicaid enrollment data and will be defined as enrollment during all months, starting with the month of childbirth through 60 days, 6 months, 12 months and 24 months postpartum. It's possible that while average Medicaid enrollment will not differ between the two groups, the type of women retaining coverage could be affected by NFP. Therefore, we will also test whether the treatment and control group members who remain in the analytic sample at each time point are similar on important baseline characteristics using a joint F-test.

For inpatient mental health outcomes, we will use hospital discharge data to measure mental health related visits on both the extensive and intensive margins. Specifically, we will create an indicator for any mental health related emergency or inpatient visits during pregnancy or the 12 months postpartum based on all-listed diagnoses (i.e. primary or secondary) for depression/anxiety/stress reaction, excluding inpatient claims on the day of delivery. We will also create a count variable for the number of mental health related emergency or inpatient visits during pregnancy or the 12 months postpartum based on all-listed diagnoses for depression/anxiety/stress reaction, excluding inpatient claims on the day of delivery. We examine these hospital-related outcomes because public insurance pays for six of every ten inpatient stays related to mental health,[25] and they could indicate inadequate outpatient treatment for perinatal depression and anxiety. We may also look at diagnoses for depression/anxiety/stress on the day of delivery, as prior research has found that a large proportion of women have diagnoses on this day.[23] These diagnoses may represent cases where women had subclinical or previously unidentified symptoms during pregnancy that were identified during the hospital stay.[23] Finally, we may construct an outcome that captures the proportion of women who have an emergency or inpatient stay related to mental health without any medication or outpatient psychotherapy treatment during pregnancy or the 12 months postpartum.

### 1.3. Timing of Study analyses

We will wait to complete this analysis until all study births have taken place and have completed administrative records which have been sent to the study team and matched to

existing analytic files. We will first publish our analysis of the primary outcome (described above) together with secondary outcomes related to maternal and infant health that can shed light on understanding impact estimates for the primary outcome. Following the publication of those results, we plan to produce separate analyses exploring the impact of NFP on maternal and infant outcomes related to the period of childbirth and the postpartum period, antenatal care and mental health care utilization. At 24 months, we plan to examine women's substance abuse, experience of violence, and use of social services, as described above. We may also examine mechanisms of these outcomes, for example, via nurse referrals. If Medicaid coverage is not differential between treatment and control group mothers at 24 months postpartum and the groups remain comparable, we may also explore outpatient mental health treatment patterns for up to two years after birth since the trajectory of depression symptoms in the postpartum period is increasingly understood to vary widely.[26,27]

### 1.4. Attrition and Missing Outcome Data

For the purposes of tracking our sample, we will separately describe the share of mother-baby dyads who fall in three distinct categories where we may observe an index pregnancy, but no index birth or fetal death as follows:

1. *Probable loss of pregnancy,* which is defined as the absence of a matched birth certificate for the index birth and either (a) a matched birth certificate for a subsequent pregnancy where the birth record indicates the mother has had 0 previous live births or (b) the presence of a matched Medicaid claim or hospital discharge record containing a diagnosis or procedure code indicating a probable pregnancy loss during the index pregnancy. Below are the validated ICD-10 diagnosis codes that we use to identify probable pregnancy losses (Table 1.7).

Table 1.7. Probable pregnancy loss diagnosis codes

| ICD-10-CM Code | Code Description |
|---|---|
| O02.0-O03.9 | Spontaneous abortion |
| O00 | Ectopic pregnancy |
| O01 | Molar pregnancy |

2. *Mother-baby linking error* is defined as the absence of a matched birth certificate for the index birth despite a matched record for the index birth in either Medicaid claims or hospital discharge records. The matched record in Medicaid claims or hospital discharge data must have an admission date within 120 days of the estimated due date reported on the baseline survey and an ICD-10 diagnosis code indicating a live birth according to ICD-9 claims identified by Kuklina et al. (2008) and widely used in claims-based research on birth outcomes.[28]

3. *Unmatched pregnancy outcome* which is defined as the absence of an administrative record for a birth or pregnancy loss episode for the index pregnancy in vital records, Medicaid claims, and hospital discharge data. Some of these "unmatched" cases may eventually become probable pregnancy loss if mothers later match to a birth certificate reporting a first-time birth.

We will report rates of missingness of outcome data for each category listed across the full sample and by treatment arm in order to characterize sample attrition and identify differential attrition across treatment arms in each of these categories.

### 1.5. Endogeneity of Data Sources and Sensitivity Analysis

Our primary analytic sample for the adverse birth outcome will focus on mothers with fetal deaths or index births observed in vital records, because these records contain more complete data on birth outcomes. In cases where a vital record exists, we will calculate the adverse birth outcome in the manner described in section 1.1. However, a small percentage of mothers who experience a fetal death or live index birth may have an error linking the mother's records to their child's vital records. To accommodate this possibility, we will perform a robustness check that utilizes the sample of mothers with a fetal death or index birth observed in either vital records, Medicaid claims or hospital discharge records.

#### 1.5.1 Calculating Adverse Birth Outcomes Without Vital Records Data

In this section, we define the calculation of the adverse birth outcome in instances where there is a matched record for the index birth or fetal death in hospital discharge record or Medicaid claims, but not in vital records.

We identify births in hospital discharge records and Medicaid claims using the ICD-10 diagnosis code definition of live birth in Kuklina et al. (2008) and listed in Table 1.8.[7] We will use the ICD-10 diagnosis code Z3A to calculate gestational age. In cases where an ICD-10 code for gestational age is absent from the hospital or Medicaid claim record, we will calculate gestational age at birth using the admission date on the record and the expected due date as given in the baseline survey. The preterm birth component of the adverse birth outcome will be coded as 1 if the gestational age on the health record is less than 37 weeks, or, if the record contains a diagnosis code P07.30-P07.39. The low-birth weight component will be coded as 1 if the record contains a diagnosis code P07.0-P07.18. The SGA component will be coded as 1 if the record contains a diagnosis code of P051. Fetal deaths will be identified with the codes in Table 1.9 when accompanied by an ICD-10 diagnosis code Z3A for gestational age of at least 20 weeks.

Table 1.8. Codes to Identify Live Births

| ICD-10 Code(s) | Code Description | Exclusion Code |
|---|---|---|
| O80, O82 | Encounter for delivery | No |
| Z37 | Outcome of delivery | No |
| 10D00Z0, 10D00Z1, 10D00Z2 | Extraction of products of conception | No |
| 10D07Z3, 10D07Z4, 10D07Z5, 10D07Z6, 10D07Z7, 10D07Z8 | Forceps/Vacuum/Breach | No |
| 0W8NXZZ | Episiotomy | No |
| 10A07ZZ, 10A00ZZ, 10A03ZZ, 10A04ZZ, 10A07ZX, 10A07ZW, 10A07Z6, 10A08ZZ | Abortion of products of conception | Yes |
| O019 | Hydatidiform mole, unspecified | Yes |
| O048 | Shock following (induced) termination of pregnancy | Yes |
| O02 | Missed abortion | Yes |
| O00 | Tubal pregnancy without intrauterine pregnancy | Yes |
| O03 | complete or unspecified spontaneous abortion without complication | Yes |
| O08 | Complications following ectopic and molar pregnancy | Yes |

Table 1.9. Fetal Death Codes

| ICD-10 Code(s) | Code Description | Exclusion Code |
|---|---|---|
| Z37.1 | Single stillbirth | No |
| Z37.4 | Twins, both stillborn | No |
| Z37.7 | Other multiple births, all stillborn | No |
| P95 | Stillbirth on child claims (occasionally appears on Mother claims) | No |
| O36.4 | Maternal care for intrauterine death | No |

### 1.6. Subgroups

Our primary sub-group is described in the study protocol. Briefly, this subgroup comprises women who have an indicator of poor mental health, are under 19 years of age, or have not

completed high school/received a General Education Development (GED) certificate by the time of enrollment.

In addition to the planned subgroup described above, we may also consider other potential subgroups of particular importance for outcomes related to pregnancy, childbirth, and maternal health. In particular, we may examine the program's impact on the outcomes of Black mothers who are disproportionately at risk for adverse pregnancy and birth outcomes compared to white mothers[29,30] and who receive postpartum mental health care at disproportionately lower rates.[31] We are also interested in focusing on births where the period of pregnancy and birth occurred prior to the start of the COVID19 pandemic (defined here as March 23, 2020). We expect that the pandemic altered NFP program implementation, clinical practices and utilization of care surrounding pregnancy, as well as families' mental health.[32]

### 1.7. Accounting for Multiple Hypothesis Testing

For our measures of guideline-recommended antenatal care, we will report a summary index of the outcomes alongside the individual outcomes to be more parsimonious in the number of hypotheses we are testing. The summary index of the four guideline-recommended care outcomes will equal the proportion of the services that the enrollee received. For example, the index will be equal to 0.5 if they received two out of the four services, and equal to 1 if they received all four services.

**Nurse-Family Partnership Evaluation Analysis Plan Addendum 2: Child Health Outcomes**

*Note: This version of Addendum 2 reflects plans for study design as of January 20, 2021. The first unblinding is planned for January 21, 2021 when interim analyses will be performed for Pay-for-Success payments.*

The goal of this analysis plan addendum is to enumerate analysis specific to the child health outcomes.

## 2. Methods

A full treatment of our empirical approach, primary study outcomes, planned subgroup analyses, and statistical power is given in sections 2.4-2.10 of McConnell et al. (2020).[1] The following sections provide additional context and details of planned analyses for child health-related outcomes.

### 2.1. Defining the Sample for Analysis

Our sample for analysis of our primary child-health related outcome will be restricted to mother-baby dyads with an identified index birth as defined in section A.1 of the pre-analysis plan. All analysis will take place at the mother level, where randomization occurs. We discuss how outcomes will be operationalized in the case of multiple births below. We will not restrict the sample based on matching to Medicaid data as the outcomes include utilization that will be observable in all-payer hospital discharge data. For outcomes that rely primarily on Medicaid data (i.e. outcomes related to preventive child health utilization), we may restrict the sample to mother-baby dyads where children retain either partial or continuous Medicaid coverage through the child's first or second year of life. We anticipate that the children of most mothers who enroll in the trial will be eligible for Medicaid for their entire first year of life but we anticipate that more children may drop off of Medicaid during their second year of life, whether because of the need to document eligibility or because of changes in eligibility in the second year of life during some periods of the study (see Table 2). We discuss the potential differences in Medicaid enrollment between control and treatment arms and planned sensitivity analyses and alternative specifications in section 2.4.

### 2.2. Construction of Study Outcomes

For our primary study outcome, we will assess the likelihood of experiencing injury, abuse, or neglect during early childhood. This outcome will be defined as a composite measure indicating a health care encounter or mortality associated with International Classification of Diseases (ICD) codes indicating either a major child injury or suspicion of abuse or neglect. We will identify major injury as any medical claim or mortality case that includes an ICD code associated with injury excluding superficial injuries, injuries related to medical care, and injuries stemming from allergic reactions. ICD codes indicating suspected abuse or neglect are derived from validated methods described in Schitzer et al. 2011 and Hooft et al. 2013.[33,34] Data on early

childhood injury outcomes and suspected abuse or neglect will come from South Carolina all-payer hospital discharge records, Medicaid inpatient and outpatient claims and mortality records. Secondary outcomes related to injuries, suspected abuse or neglect and mortality will decompose the primary outcomes into their composite parts. In the case of multiples, we will consider a mother to experience the outcome if any of her children do and we will average the number of events in cases where the outcome is continuous (i.e. the number of injuries).

### 2.2.1 Secondary Outcomes

We specify the diagnosis and procedure codes used to measure the utilization of preventative care for children under 2 in Table 2.1. We will construct these outcomes using Medicaid claims data. We define receiving the share of recommended well-child visits as a binary indicator equal to one if a child has received at least the number of well-child visits in the first 2 years of life as recommended by the American Academy of Pediatrics.[35] In the case of multiples we will consider binary outcomes to have occurred if all children from the index birth meet the criteria.

Table 2.1. Preventative Care Utilization Measures

| Outcome | Outcome Definition | Code(s) |
|---|---|---|
| Well-child visits | Binary indicator equal to 1 if the child has received 9 well-child visits by 24 months | ICD-10: Z00.1; CPT: 99391, 99392 |
| Lead screening | Received at least one lead screening by 24 months | ICD-10: Z13.88; CPT: 83655 |
| Developmental screening | Received at least one developmental screening by 9 months | ICD-10: Z13.41, Z13.42; CPT: 96110 |
| Dental visit | Received at least one dental visit by 12 months | NA |
| Fluoride treatments | Binary indicator equal to 1 if the child has received at least one fluoride treatment by 24 months[iii] | ICD-10: Z293; CPT: 99188 |

[iii] While our protocol paper specified this outcome as receiving at least four fluoride treatments, there are inconsistent guidelines regarding the recommended number of visits. Therefore, we have modified the definition of this outcome to be a binary variable indicating whether any fluoride treatments were received. See https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/dental-caries-in-children-from-birth-through-age-5-years-screening for details on fluoride varnish recommendations.

### 2.3. Timing of study analyses

We will complete analysis of the primary outcome and related secondary outcomes once all study births have occurred, and at least 24 months have elapsed following the study birth. We will add additional lags to account for the time it takes for administrative records to be complete, sent to the study team, and matched to existing analytic files. If we make additional changes to the analysis plan before starting the analysis, we will indicate those changes with a new version number and date. We plan to publish analyses of the primary outcome and related secondary outcomes first. A separate manuscript is planned to explore outcomes related to preventative health care utilization for children. We also plan to conduct follow-up analyses of similar outcomes as children age, including beyond the time when families would participate in NFP. Analytical plans for follow-up longitudinal analyses will be added as they are developed.

### 2.4. Attrition and Missing Outcome Data

We will report on the following categories of mother-baby dyads for whom we may not be able to see outcomes related to child health, reporting on the sample of children tracked to 12 months and 24 months separately:

a. *Mortality prior to 12 or 24 months of age:* We will use vital records to identify children born into the study who experience mortality prior to the age of 12 months.
b. *Unmatched to Medicaid Eligibility through 12 or 24 months of age:* We will consider children who never match to Medicaid eligibility in any of their first 12 months of life or 13-24 months of life respectively.

We will report rates of missingness for each category listed here for the whole sample and by treatment arm in order to characterize missing outcome data and identify differential rates of missing outcome data across treatment arms in each of these categories.

### 2.5. Endogeneity of Data Sources and Sensitivity Analysis

As discussed above, because child preventative health outcomes will be observed exclusively in Medicaid claims data, we will not observe these outcomes among children who are not enrolled in Medicaid. Because NFP may affect maintenance of Medicaid coverage through the child's first two years of life, child enrollment rates in Medicaid may differ between control and treatment groups.

We will conduct several analyses to determine whether rates of enrollment in Medicaid differ between treated and control groups. First, we will compare enrollment through the first year of life and the first two years of life across treatment and control groups. Even if we do not observe differences in Medicaid enrollment rates between the two groups, the type of children retaining coverage may be influenced by participation in NFP. Therefore, we will also test

whether there are differences in the characteristics of participants between control and treatment groups over the two time periods (12 months and 24 months). We will use a joint F-test to compare characteristics measured at baseline across these two groups using the list of baseline covariates specified in section B.1.

When conducting study analyses, we will report outcomes observed over the first year of life, where more children are likely to be enrolled in Medicaid as well as outcomes over the first 24 months of life. As specified in our study protocol in Section 2.12, [1] we will report our primary outcome excluding data that comes from Medicaid to assess the robustness of our results to this potential source of endogeneity. If we observe significant differences in Medicaid enrollment across treatment and control groups in the second year of life, we may focus more on outcomes that can be measured within the first year of life.

### 2.6. Subgroups

Our primary sub-group is described in the study protocol. Briefly, this subgroup is composed of women who report mental health challenges, are under 19 years of age, or have not completed high school/received a GED certificate by the time of enrollment.

In addition to the planned subgroup described above, we may also consider other potential sub-groups of particular importance for outcomes related to child health. In particular, we are interested in focusing on child outcomes where the first year of the child's life was completed prior to the start of the COVID-19 pandemic (defined here as March 23, 2020). We expect that the pandemic altered the way that the NFP program was implemented. In addition, it is likely that the pandemic substantially affected clinical practices and utilization of pediatric care.

### 2.7. Accounting for Multiple Hypothesis Testing

We have planned for the analysis of several outcomes related to utilization of preventative care for children. To account for multiple related secondary outcomes in the domain of preventative care we will construct an index of preventative care-seeking, similar to the approach proposed by Kling, Liebman and Katz 2007.[36] We will report the index in addition to other outcomes.

## Nurse-Family Partnership Evaluation Analysis Plan Addendum 3: Alter Maternal Life Course

*Note: This version of Addendum 3 reflects plans for study design as of January 15, 2021. The first unblinding is planned for January 18th, 2021 when outcomes will be analyzed for Pay-for-Success payments.*

The goal of this analysis plan addendum is to enumerate analysis specific to the outcome domain focused on birth spacing and family planning utilization.

### 3. Methods

A full treatment of our empirical approach, primary study outcomes, planned subgroup analyses, and statistical power is given in sections 2.4-2.10 of McConnell et al. (2020).[1] The following sections provide additional context and details on auxiliary analysis for the outcomes related to altering the maternal life course including birth spacing and family planning utilization.

#### 3.1. Defining the Sample for Analysis

The sample will be restricted to women with a fetal death or index birth as defined in section A1 of the pre-analysis plan.

#### 3.2. Construction of Study Outcomes

We outline the construction of study outcomes across two domains: outcomes related to birth spacing and outcomes related to the utilization of contraceptive methods.

##### 3.2.1. Construction of birth spacing outcomes

Data will be obtained from vital statistics birth records. In the case of multiples, birth spacing will be measured as the time between the date of birth of the last child from the index pregnancy to the date of birth of the first child from the subsequent pregnancy. There may be cases where the subsequent birth is implausibly close to the index birth. Births that occur less than or equal to 90 days from the index birth will be assumed to be multiple gestation based on guidelines provided by the National Center for Health Statistics. Subsequent births that occur between 90 days and 21 weeks (147 days) after the index birth will be considered outliers and will be dropped from the analysis. We define birth outcomes of interest in Table 3.1.

Table 3.1. Birth spacing outcomes

| Birth spacing outcomes | Follow-up |
|---|---|
| Inter-birth interval of < 21 months (primary outcome) | 21 months |
| Inter-birth interval of < 24 months | 24 months |

| | |
|---|---|
| Inter-birth interval of < 15 months | 15 months |
| Inter-birth interval (continuous) | 60 months |

### 3.2.2. Construction of contraception outcomes

Table 3.2 summarizes outcomes related to the utilization of contraceptive methods. The primary data source used for the contraception outcomes will be Medicaid claims data. Medicaid claims capture contraceptive provision among Medicaid enrollees that takes place in the hospital before discharge (inpatient), during outpatient hospital visits, and during ambulatory clinical visits. We will rely primarily on Medicaid claims data because most postpartum contraceptive provision occurs after hospital discharge. However, we will also supplement Medicaid discharge data with hospital discharge data to capture any inpatient contraceptive provision that may be missing from Medicaid claims. Any family planning related counseling or service will include any diagnosis code, procedure code, CPT code, or HCPCS codes for contraceptive counseling, the intrauterine device, implant, injectable, contraceptive pill, patch, ring or diaphragm (See Table 3.3 below) and National Drug Codes for the contraceptive pill, patch, and ring identified by the Office of Population Affairs Contraceptive Care Measures. Receipt of a moderately effective method of contraception will include all of the above methods but will not include contraceptive counseling without provision of a contraceptive method. Immediate postpartum contraception will include diagnosis, procedure, CPT or HCPCS code for an intrauterine device or contraceptive implant.

Contraceptive counseling and/or contraceptive receipt within six weeks of hospital discharge will meet the criteria for the 6-week contraceptive outcomes. Contraceptive counseling and/or receipt within 12 months of the date of hospital discharge will meet the criteria for the 12-month contraceptive outcomes.

Table 3.2. Contraceptive outcomes

| Variables | Follow-up |
|---|---|
| Any family planning related counseling or service | 6 weeks |
| Received a highly or moderately effective method of contraception[iv] | 6 weeks |
| Immediate postpartum long-acting reversible contraception | 6 weeks |
| Any family planning related counseling or service | 12 months |
| Received a highly or moderately effective method of contraception[16] | 12 months |
| Postpartum intrauterine device insertion | 12 months |

[iv] CDC defines highly effective contraception to include implant, Immediate Post-Partum-Long-Acting Reversible Contraception, Long-Acting Reversible Contraception, or sterilization and moderately effective contraception to include path, ring, diaphragm, injectables and contraceptive pills.

| | |
|---|---|
| Time to first family planning counseling or service (months from pregnancy) | 24 months |
| Time to first utilization of highly effective contraceptive methods (months from discharge) | 24 months |

In Table 3.3 we define codes used to identify specific contraceptive methods that will be used in analyses. Codes will be identified in both Medicaid claims and hospital discharge records.

Table 3.3. Codes used to identify contraception

| Contraceptive method | Diagnosis codes | Procedure | CPT | HCPCS |
|---|---|---|---|---|
| Female sterilization | Z30.2 | 0U574ZZ, 0U578ZZ, 0UL74CZ, 0UL74DZ, 0UL74ZZ, 0UL78DZ, 0UL78ZZ | 58565, 58600, 58605, 58611, 58615, 58670, 58671 | A4264 |
| Intrauterine device | Z30.430, Z30.014 | 0UH97HZ, 0UH98HZ, 0UHC7HZ, 0UHC8HZ, 0UH90HZ | 58300 | J7300, J7301, J7302, S4989, Q0090, S4981, J7297, J7298 |
| Contraceptive implant | Z30.017, Z30.46 | 0JHD0HZ, 0JHD3HZ, 0JHF0HZ, 0JHF3HZ | 11981 | J7306, J7307 |
| Contraceptive injectable | Z30.013, Z30.42 | -- | -- | J1050 |
| Contraceptive pills, patch, ring | Z30.011, Z30.41, Z30.016, Z30.45, Z30.015, Z30.44 | -- | -- | S4993, J7304, J7303 |
| Diaphragm | -- | -- | 57170 | A4266, A4261 |

### 3.3. Timing of Study analyses

We will complete this analysis once all study births have occurred, and at least 24 months have elapsed following the last study birth. We will add additional lags to account for the time it takes for administrative records to be complete, sent to the study team, and matched to existing analytic files. Based on the estimated delivery date collected in the baseline survey, the latest expected study delivery date is November 7, 2020. As we define births within a 120-day window of the estimated gestational age as a study index birth, the latest possible date when a

study index birth could occur is March 7, 2020. We will allow for twenty-four months of follow-up from the time of the final possible study index birth (March 7, 2022), six additional months for the study outcomes to be fully incorporated into South Carolina administrative data (September 7, 2023), and an additional three-month buffer for the administrative data to be matched with the study dataset (December 7, 2023). Therefore, we anticipate that analyses of these outcomes will begin in December 2023.

This study includes four birth spacing outcomes. The first three are birth interval indicators defined as having no subsequent live birth observed within 21 (primary outcome), 15 and 24 months (secondary outcomes) of the index birth. The fourth is a continuous outcome defined as the number of months before any subsequent birth within 60 months of the index birth. The analysis for this outcome will include only women who went on to have another birth within 60 months of their index birth. The first three of the birth spacing outcomes described in Table 3.1 will be reported in the main manuscript for the maternal life course objective. The continuous birth interval outcome will be included in a later manuscript which will include longer-term maternal outcomes.

### 3.4. Attrition and Missing Outcome Data

As described in section 1.4 some mother-babies will be matched to a childbirth-related Medicaid claim or discharge record but unmatched to a birth certificate record or fetal death vital record. We will conduct a sensitivity analysis that will include these study participants, in addition to the main sample following procedures defined in section 1.6.1. For mother-babies without a birth certificate or fetal death record, we will use hospital discharge data as the date of birth or fetal death.

### 3.5. Endogeneity of Outcome Data and Sensitivity Analyses

Because outcomes related to contraceptive uptake will rely primarily on Medicaid claims data, we cannot observe contraception use in outpatient settings among women who are not enrolled in Medicaid. As a non-expansion state, all women with Medicaid pregnancy coverage in South Carolina, except undocumented and recent immigrants, retain Medicaid coverage for at least 60 days following childbirth. Women who qualify for other payment categories (e.g. foster-care, disability), may retain coverage past 60 days. Because NFP may affect maintenance of Medicaid coverage past 60 days postpartum, including maintenance of enrollment in Medicaid Family Planning coverage, we may not be able to observe treatment and control group mothers in the Medicaid claims data at equal rates.

To determine whether continuous Medicaid coverage differs between treatment groups, we will first compare continuous Medicaid coverage of any type (coverage through the low-income

families, qualified disabled workers, pregnancy and family planning eligibility pathways), and family planning Medicaid coverage specifically, between treatment and control groups at the three contraception outcome time points (6 weeks, 12 and 24 months). Continuous coverage will be measured using Medicaid enrollment data and will be defined as enrollment during all months, starting with the month of childbirth through 6 weeks, 12 and 24 months postpartum. It is also possible that while average Medicaid enrollment will not differ between the two groups, the type of women retaining coverage will be affected by NFP. Therefore, using a joint F-test, we will also test whether there are differences in the characteristics of participants between control and treatment groups enrolled in Medicaid during the three-outcome time points (60 days, 12 months, 24 months).

When conducting the study analysis, we will divide outcomes between those that we observe within 60 days after childbirth (when nearly all women retain Medicaid), and those measured after 60 days postpartum. Our main analysis will include the full sample of women regardless of postpartum Medicaid coverage after 60 days, but we will also examine the outcomes in the subgroup of women who retained coverage for 24 months postpartum. If we find evidence that NFP affected average Medicaid enrollment or the type of woman enrolling in Medicaid, we will place more emphasis on the 6-week contraception outcomes when Medicaid coverage loss is minimal and more likely to be balanced between groups. Alternatively, if we find little evidence of differential coverage loss at 12 and 24 months, we will consider contraception outcomes at all three time points.

### 3.6. Subgroups

Our primary sub-group is described in McConnell et al. (2020).[1] Briefly, this subgroup comprises women who have an indicator of poor mental health, are under 19 years of age, or have not completed high school/received a General Education Development certificate by the time of enrollment.

In addition to the planned subgroup described above, we may also consider several other potential subgroups. We are interested in focusing more narrowly on groups who are most likely to benefit from increased access to family planning including teens, because pregnancies in this subgroup are more likely to be unintended than among adult women, and women who reported at baseline that they did not want to have another birth for at least two years. We are also interested in focusing on births that took place at least 24 months before the start of the COVID-19 pandemic (defined here as March 23, 2020). We expect that the pandemic altered the way that the NFP program was implemented. In addition, it is likely that the pandemic affected the availability of contraceptive services, fertility preferences, and employment; all factors on the causal pathway between NFP and birth outcomes.

[1] McConnell MA, Zhou RA, Martin MW, et al. Protocol for a randomized controlled trial evaluating the impact of the Nurse-Family Partnership's home visiting program in South Carolina on maternal and child health outcomes. *Trials*. 2020;21(1):997. doi:10.1186/s13063-020-04916-9

[2] Medicaid and CHIP Income Eligibility Limits for Pregnant Women, 2003-2020. KFF. Published April 1, 2020. Accessed January 7, 2021. https://www.kff.org/medicaid/state-indicator/medicaid-and-chip-income-eligibility-limits-for-pregnant-women/

[3] Medicaid, Children's Health Insurance Program, & Basic Health Program Eligibility Levels | Medicaid. Accessed January 7, 2021. https://www.medicaid.gov/medicaid/national-medicaid-chip-program-information/medicaid-childrens-health-insurance-program-basic-health-program-eligibility-levels/index.html#footnote1

[4] Medicaid Income Eligibility Limits for Parents, 2002-2020. KFF. Published April 1, 2020. Accessed January 15, 2021. https://www.kff.org/medicaid/state-indicator/medicaid-income-eligibility-limits-for-parents/

[5] Medicaid, Children's Health Insurance Program, & Basic Health Program Eligibility Levels | Medicaid. Accessed January 7, 2021. https://www.medicaid.gov/medicaid/national-medicaid-chip-program-information/medicaid-childrens-health-insurance-program-basic-health-program-eligibility-levels/index.html#footnote1

[6] Medicaid Income Eligibility Limits for Infants Ages 0 – 1, 2000-2020. KFF. Published April 1, 2020. Accessed January 7, 2021. https://www.kff.org/medicaid/state-indicator/medicaid-income-eligibility-limits-for-infants-ages-0-1/

[7] Medicaid Income Eligibility Limits for Children Ages 1 – 5, 2000-2020. KFF. Published April 1, 2020. Accessed January 7, 2021. https://www.kff.org/medicaid/state-indicator/medicaid-income-eligibility-limits-for-children-ages-1-5/

[8] Medicaid Income Eligibility Limits for Children Ages 6 – 18, 2000-2020. KFF. Published April 1, 2020. Accessed January 7, 2021. https://www.kff.org/medicaid/state-indicator/medicaid-income-eligibility-limits-for-children-ages-6-18/

[9] Puma MJ, Olsen RB, Bell SH, Price C. *What to Do When Data Are Missing in Group Randomized Controlled Trials. NCEE 2009-0049*. National Center for Education Evaluation and Regional Assistance; 2009. Accessed January 8, 2021. https://eric.ed.gov/?id=ED511781

[10] Han D, Khadka A, McConnell M, Cohen J. Association of Unexpected Newborn Deaths With Changes in Obstetric and Neonatal Process of Care. JAMA Netw Open. 2020;3(12):e2024589. doi:10.1001/jamanetworkopen.2020.24589

[11] Ford JB, Roberts CL, Algert CS, et al. Using hospital discharge data for determining neonatal morbidity and mortality: a validation study. BMC Health Serv Res. 2007;7(1):188. doi:10.1186/1472-6963-7-188

[12] How Does CDC Identify Severe Maternal Morbidity? | CDC. Published December 26, 2019. Accessed January 15, 2021. https://www.cdc.gov/reproductivehealth/maternalinfanthealth/smm/severe-morbidity-ICD.htm

[13] Bailey, B. A., McCook, J. G., Hodge, A., & McGrady, L. (2012). Infant birth outcomes among substance using women: why quitting smoking during pregnancy is just as important as quitting illicit drug use. Maternal and child health journal, 16(2), 414-422.

[14] Shah, P. S., & Shah, J. (2010). Maternal exposure to domestic violence and pregnancy and birth outcomes: a systematic review and meta-analyses. Journal of women's health, 19(11), 2017-2031.

[15] Mills, G., & Urban Institute. (2014). Understanding the rates, causes, and costs of churning in the supplemental nutrition assistance program (SNAP). United States Department of Agriculture, Food and Nutrition Service, Office of Policy Support.

[16] Jarlenski, M., Krans, E. E., Chen, Q., Rothenberger, S. D., Cartus, A., Zivin, K., & Bodnar, L. M. (2020). Substance use disorders and risk of severe maternal morbidity in the United States. Drug and alcohol dependence, 216, 108236.

[17] Davidov, D. M., Davis, S. M., Zhu, M., Afifi, T. O., Kimber, M., Goldstein, A. L., ... & Stocks, C. (2017). Intimate partner violence-related hospitalizations in Appalachia and the non-Appalachian United States. PLoS one, 12(9), e0184222.

[18] Schafer, S. D., Drach, L. L., Hedberg, K., & Kohn, M. A. (2008). Using diagnostic codes to screen for intimate partner violence in Oregon emergency departments and hospitals. Public Health Reports, 123(5), 628-635.

[19] Siega-Riz AM, Bodnar LM, Stotland NE, Stang J. The Current Understanding of Gestational Weight Gain Among Women with Obesity and the Need for Future Research. NAM Perspectives. Published online January 6, 2020. doi:10.31478/202001a

[20] Horovitz, J.L. & C. F. Manski (2000). "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data." Journal of the American Statistical Association, 95, 77-84.

[21] Fairbrother N, Young AH, Janssen P, Antony MM, Tucker E. Depression and anxiety during the perinatal period. BMC Psychiatry 2015; 15: 206.

[22] Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: Results from the national comorbidity survey replication (NCS-R). JAMA. 2003;289(23):3095.

[23] Sherman LJ, Ali MM. Diagnosis of Postpartum Depression and Timing and Types of Treatment Received Differ for Women with Private and Medicaid Coverage. Women's Health Issues 2018; 28: 524–9.

[24] Stewart, D. E., & Vigod, S. (2016). Postpartum depression. New England Journal of Medicine, 375(22), 2177–2186

[25] Saba DK, Levit KR, Elixhauser A. Hospital Stays Related to Mental Health, 2006. 2008; : 13.

[26] McCall-Hosenfeld JS, Phiri K, Schaefer E, Zhu J, Kjerulff K. Trajectories of Depressive Symptoms Throughout the Peri- and Postpartum Period: Results from the First Baby Study. J Womens Health 2002 2016; 25: 1112–21.

[27] Putnick DL, Sundaram R, Bell EM, et al. Trajectories of Maternal Postpartum Depressive Symptoms. Pediatrics 2020; published online Oct 1. DOI:10.1542/peds.2020-0857.

[28] Kuklina EV, Whiteman MK, Hillis SD, et al. An enhanced method for identifying obstetric deliveries: implications for estimating maternal morbidity. *Matern Child Health J*. 2008;12(4):469-477. doi:10.1007/s10995-007-0256-6

[29] Leonard SA, Main EK, Scott KA, Profit J, Carmichael SL. Racial and Ethnic Disparities in Severe Maternal Morbidity Prevalence and Trends. *Ann Epidemiol*. Published online February 28, 2019. doi:10.1016/j.annepidem.2019.02.007

[30] Howell EA, Brown H, Brumley J, et al. Reduction of Peripartum Racial and Ethnic Disparities: A Conceptual Framework and Maternal Safety Consensus Bundle. *Obstet Gynecol*. 2018;131(5):770-782. doi:10.1097/AOG.0000000000002475

[31] Kozhimannil KB, Trinacty CM, Busch AB, Huskamp HA, Adams AS. Racial and ethnic disparities in postpartum depression care among low-income women. Psychiatr Serv Wash DC 2011; 62: 619–25.

[32] Davenport MH, Meyer S, Meah VL, Strynadka MC, Khurana R. Moms Are Not OK: COVID-19 and Maternal Mental Health. Front Glob Womens Health 2020; 1. DOI:10.3389/fgwh.2020.00001.

[33] Schnitzer PG, Slusher PL, Kruse RL, Tarleton MM. Identification of ICD codes suggestive of child maltreatment. *Child Abuse Negl*. 2011;35(1):3-17. doi:10.1016/j.chiabu.2010.06.008

[34] Hooft A, Ronda J, Schaeffer P, Asnes AG, Leventhal JM. Identification of Physical Abuse Cases in Hospitalized Children: Accuracy of International Classification of Diseases Codes. *J Pediatr*. 2013;162(1):80-85. doi:10.1016/j.jpeds.2012.06.037

[35] Turner K. Well-Child Visits for Infants and Young Children. Am Fam Physician. 2018;98(6):347-353.

[36] Kling J, Liebman J, Katz L. Experimental Analysis of Neighborhood Effects. *Econometrica*. 2007;75(1):83-119.