Title: Health and Wellbeing of Pregnant and Post-Partum Women During the COVID-19 Pandemic

Document: Statistical Analysis Plan

NCT number: 04385238

Date: 04 May 2020

An anonymous, online, cross-sectional survey targeting pregnant and postpartum women will be conducted in 64 countries between May 26 2020 and June 13 2020. Participation in this study will be voluntary. All potential participants will be informed about the research objectives and standards of confidentiality regarding the use of the data. The survey, hosted on the Pregistryplatform for COVID-19 studies (https://corona.pregistry.com), will be advertised predominantly in social media channels and online parenting forums. Advertisements and the survey will be available in twelve languages (Arabic, Chinese, English, French, German, Italian, Korean, Portuguese, Russian, Spanish, Turkish, and Urdu) by human translators. We will seek to obtain at least 100 responses from each of the countries with the highest number of COVID-19 cases at the time of recruitment. Interested participants will be invited to follow a link to take the survey. The survey will collect standard demographic data and included questions that addressed topics such as COVID-19 exposure and worries, lifestyle changes, media exposure, protective factors, and mental health.

Participants. Women who self-identified as being 18 years or older at the time of the survey and as currently pregnant or having given birth within the past 6 months will be eligible to participate. The study was classified exempt by the Harvard Longwood Campus Institutional Review Board (HLC IRB) per the regulations found at 45 CFR 46.104(d)(2) on the basis that it poses no greater than minimal risk and that the recorded information cannot readily identify the subject (directly or indirectly).

Measures

COVID-19 Assessments

COVID-19 Exposure. Questions assessing whether participants were tested, diagnosed, or in contact with an individual who had COVID-19 will be adapted for this study based on those formulated by the US Centers for Disease Control and Prevention for the Household Pulse Survey.

COVID-19 Information Seeking. Participants will be asked about the frequency (never, <1x/day, 2-4x/day, 5-8x/day, 9-16x/day, and >16x/day) of their interactions with various sources of information, including the news, social media, and interpersonal discussions about COVID-19 using a measure modified from other published studies. For analyses, interactions will be categorized as never, <1x/day, 2-4x/day, 5+x/day.

COVID-19 Worries. Participants will be asked to rate their overall level of worry about COVID-19 on a Likert Scale ranging from 1 for “not worried at all” to 4 for “very worried”. They will be then asked to endorse fifteen specific worries on a list developed for this study. Exploratory factor analyses will be conducted to identify domains within the questionnaire, with oblimin rotation on a tetrachoric correlation matrix due to the binary nature of the variables. Worries will be categorized into the following domains: social (parents/grandparents unable to visit, family unable to visit, not able to have a baby shower, not able to attend a funeral), COVID-19 infection-related (participant or partner will bring infection home, family or friends will get COVID-19), child-related (no adequate childcare, other children will get COVID-19), delivery-related (partner not present during delivery, changes to delivery plan, unborn baby will get COVID-19, not able to breastfeed), economic (significantly affect economic situation/finances), and missing doctor appointments.

COVID-19 prevention behaviors. Participants will be asked to endorse seventeen behaviors they had engaged in to protect themselves from COVID-19 from a list developed for this study based on WHO recommendations and media reports. Behaviors will be classified into the following categories: hygiene-related (mask-wearing, washing hands, disinfecting surfaces), physical distancing (avoiding public places, restaurants and other people, canceling personal engagements, work or school and

working at home), canceling travel (for work or pleasure), stockpiling essential resources (food or water, hand sanitizer, medication), postponing medical care, and prayer.

Mental Health Outcomes

Depression and Anxiety will be assessed via the Patient Health Questionnaire-4 (PHQ-4) a four-item inventory rated on a four-point Likert-type scale. Items are drawn from the first two items of the 'Generalized Anxiety Disorder–7 scale' (GAD–7) and the 'Patient Health Questionnaire-8' (PHQ-8). The overall PHQ-4 score is a sum of the four items (0 = not at all, 1 = several days, 2 = more than half the days, 3 = nearly every day). A PHQ-4 score of >=6 is considered clinically significant.

Posttraumatic stress symptoms will be assessed via a modified version of the Impact of Events Scale - 6 (IES-6). Participants will be asked to report how bothered they will be by each symptom from 0 (Not at all) to 4 (Extremely) over the past seven days. The symptom statements will be modified to assess impact related to the COVID-19 pandemic. The scale is scored by calculating the mean of the five items used in this study, with the original IES-6 cutoff score of 1.75 yielding 0.88 sensitivity and 0.85 specificity for posttraumatic stress disorder (PTSD) diagnosis. Loneliness will be evaluated using the UCLA Three-Item Loneliness Scale (UCLA-3). Participants are asked how often they experience the following: feeling that they lack companionship, feeling left out, and feeling isolated from others. Questions will be framed so that participants described their feelings since the start of the COVID-19 pandemic. Responses are rated as 1 (Hardly Ever), 2 (Some of the Time), or 3 (Often), with the overall score as the sum of the three responses. A score of 3 is considered low, 4-5 medium, and >= 6 high. High loneliness has been associated with poorer mental and physical health over the life course.

Sociodemographic factors and potential confounders

Standard socio-demographic measures will be collected including age, education (categorized as never attended school, elementary school, some high school, high school graduate or general equivalency diploma (GED), some college/university, college diploma or university degree, master's degree, professional degree, doctoral degree), self-identified race/ethnicity (categorized as White/Caucasian, Latina/Hispanic, Asian, South Asian, Black, Middle Eastern, Native Hawaiian or Other, Pacific Islander, American Indian or Alaska Native, Other/Multiracial), employment status (healthcare worker in a hospital or clinic, worked in a nursing home, essential/key worker (as defined by the government), none of these, don't know), medical coverage status, marital status (married, living with partner, divorced, separated, single, widowed), weeks pregnant/postpartum (first trimester 0 to <13 weeks, second trimester 13 to <28 weeks, third trimester 28+ weeks, postpartum). Participants indicated their country of residence which was classified by region for analytic purposes.

Statistical analyses

Descriptive statistics for socio-demographic characteristics, COVID-19 exposure variables, mental health, and loneliness will be calculated by perinatal stage, classified as first, second, or third trimester, or post-partum, for all participants. A series of multivariable logistic regressions will be then run to test the study hypotheses. For these analyses, PHQ-4, IES-6 and UCLA-3 outcomes will be specified as binary dependent variables using recommended cutoffs. The first set of models included pregnancy stage, socio-demographic variables and COVID-19 exposure as independent variables. The second set of models examined the type and amount of information seeking about COVID-19 as independent variables. The third and fourth set of models focused on COVID-19 worries and COVID-19 behaviors as independent variables. Models for information seeking, COVID-19 worries, and

COVID-19 behaviors will be adjusted for age, level of education, race, survey region, marital status, and pregnancy stage.

Models will be estimated using the GLM function in the R statistical program, version 3.6.2. Due to the number of models run, interpretations are focused on effects that met the threshold of $p < .001$. Due to low levels of missingness (667 [9%]) a complete case analysis was conducted. The raw data used for the study are publicly available.