Cash Benefits and Reproductive/Perinatal Health

Protocol and Analysis Plan

March 23, 2023

ClinicalTrials.gov Identifier: NCT05782660

Study Team

Sumit Agarwal, MD, MPH, Brigham and Women's Hospital, Harvard Medical School

Jeffrey Liebman, PhD, Harvard Kennedy School

Table of Contents

A. Introduction

During the first two years of the COVID-19 pandemic, the City of Chelsea, Massachusetts held a lottery to allocate cash benefits to its residents for ten months. In one of the original analyses of the Chelsea Eats program, the survey results showed that there was a higher rate of pregnancy in the treatment group compared to control. The underlying mechanism, however, remained unclear; the cash benefit could theoretically have induced changes in family planning or in maternal health, leading to the observed changes in pregnancy rates. Using data from the Chelsea Eats program, we propose to study the impact of the cash benefit on reproductive and perinatal health.

This proposal and analysis plan detail the intervention, randomization, and data source. Importantly, while this analysis plan is specified after the conduct of the trial, it does pre-specify the outcomes and statistical analysis prior to analysis of the data. No analyses of post-intervention outcomes were conducted prior to gaining access to the data.

B. Treatment

Chelsea, Massachusetts, a city of 40,000 people just north of Boston, was among the places in the country hardest hit by COVID-19, both from a health and an economic perspective. Its heavily Latino population is concentrated in sectors of the economy that were shut down when the pandemic hit, and Chelsea residents are also disproportionately likely to be front-line service workers exposed to infection risk. In April 2020, local community organizations and the City of Chelsea responded to the economic crisis facing jobless Chelsea residents by mounting an unprecedented food distribution effort.

In September 2020, after five months of running its food distribution sites, the City redirected its efforts toward distributing financial support so that residents could purchase their own food through a program called Chelsea Eats. By combining city general revenue funds, state aid, and philanthropic contributions, the City assembled enough resources to distribute Chelsea Eats debit cards to approximately 2,000 households and to replenish the cards on a monthly basis for a total of six months, that was later extended to ten months. The card amounts vary with household size. Most households received $400 per month, but one- and two-person households received $200 and $300, respectively. Spending from the cards was not restricted to food but could be spent on anything and anywhere Visa was accepted. In total, 3,615 households applied for the cards, and 2,074 were chosen to receive the cash assistance cards via a lottery. The debit cards were credited with the first payment on November 18th, 2020 and the second payment on December 18th, 2020. The program continued with monthly credits through August 2021.

C. Randomization and Participant Eligibility

Between July 27, 2020 and August 17, 2020, the City accepted applications for the Chelsea Eats cash assistance cards. Multilingual information about the cards was distributed to individuals using the city-run food distribution sites. Additionally, information and applications were disseminated to community-based organizations, food pantries, faith groups, health care organizations, and low- and moderate-income housing complexes. Direct outreach to residents by city staff occurred in multiple locations, including food pantry lines, COVID-19 testing lines, and social service agencies. Applications could be submitted online via the city's website or on paper by dropping them off or mailing them to City Hall. In practice, most applications were submitted by residents attending one of the various city or community partner food pantries, where city staff, equipped with tablets, assisted residents with the application.

Eligibility criteria included:

- Resident of Chelsea, Massachusetts
- Household income at or below 30% of the U.S. Department of Housing and Urban Development's Area Median Income

The lottery was a weighted lottery. Specifically, a household could receive additional lottery tickets by meeting any of the following criteria:

- No one in the household was currently working
- The household was not receiving unemployment insurance
- The household was not receiving food assistance (e.g., SNAP benefits)
- There was a disabled household member
- There was a household member over 65 years of age
- There was a household member who was a veteran
- There was a household member who was under 6 years of age
- There was a household member between the ages of 6 and 17

The total number of lottery tickets per application ranged from 1 to 8. Although it was the city's intention to restrict households to a single application, some households managed to enter the lottery with more than one application. In these cases, the household still received only a single Chelsea Eats card. Records for the duplicate applicants will be combined into a single record, summing the lottery tickets across the duplicate records to determine the household's overall probability of winning the lottery. We account for the differential probability of winning the lottery across households in the statistical analysis (see section E.2).

Of note, some households never picked up their Chelsea Eats cards. The city then gave these unused cards to households from a randomly generated waitlist that was created at the time of random assignment. All waitlisted households ultimately received Chelsea Eats cards and are therefore part of the treatment group. The waitlist households received their cards approximately one month later than the primary lottery winners, but the cards were credited with both the payments for the first and second months – so the total payments received by waitlist households are the same as initial lottery winners.

D. Data, Outcomes, and Other Variables

D.1. Data

Female study participants of childbearing age (typically defined as age 15-44) will be matched to their electronic health record at Cambridge Health Alliance (CHA), Mass General Brigham (MGB), and East Boston Neighborhood Health Center (EBNHC) using a probabilistic algorithm based on name, date of birth, gender, address, and phone number.

D.2. Planned Outcomes

Outcomes will be assessed on a per-participant basis and measured using the electronic health record data.

1. Primary Outcome
    a. Pregnancy
        i. We will first confirm the result from the survey data by assessing for pregnancies using the electronic health record data.
        ii. Pregnancies will be identified in the electronic health record as a positive urine or blood pregnancy test, ultrasound, and/or diagnosis codes over the 10 months of the trial.
        iii. We will then use the electronic health record identify live child births delivered vaginally or via Cesarean section. These can occur after the 10-month duration of trial.
        iv. Any pregnancy without an eventual delivery (not carried to term), the primary outcome for our main analysis, could be indicative of miscarriage, abortion, or fetal death, which will be assessed further with the secondary outcomes below.

2. Secondary Outcomes
    a. Miscarriage
        i. We will use diagnosis codes and documentation within clinical notes to subsequently identify miscarriages that present to the health care setting.
    b. Abortion
        i. We will also identify procedural abortions and prescriptions for medical abortions. Using documentation within clinical notes to distinguish between miscarriage-related abortions (i.e., those performed after early pregnancy loss, incomplete abortions, and ectopic pregnancies) and induced abortions, this outcome will focus specifically on induced abortions.
    c. Prenatal vitamin prescriptions prior to pregnancy
        i. To characterize family planning among participants, we will assess new use of prenatal vitamins, including multivitamins. Physicians can prescribe these vitamins or document over-the-counter use of them in

the electronic health record. We will exclude those that were prescribed or documented during pregnancy and focus specifically on those that occurred prior to any pregnancy over the 10 months of the trial.

d. Contraception
    i. To further understand family planning among participants, we will assess utilization of long-acting reversible contraceptive methods (intrauterine devices or implants) and prescriptions for hormonal birth control methods (pill, patch, ring).
e. Number and timing of prenatal visits before delivery
    i. In addition to quantity of visits, we will also assess for initiation during the first trimester and calculate the Adequacy of Prenatal Care Utilization (APNCU) Index, which characterizes the adequacy of a patient's prenatal care as inadequate, intermediate, adequate, or adequate plus based on the timing of their prenatal care initiation and the number of prenatal visits after initiation.
f. Composite of birth outcomes
    i. Following McConnell et al. (2022), we will construct a composite of at least one of: low birth weight, preterm birth, small for gestational age, or perinatal mortality. These are described in further detail below.
    ii. *Birth weight* is a continuous measure of the infant's birth weight in grams. *Low birth weight* is defined as a birth weight less than 2500 grams.
    iii. *Gestational age* at birth is a continuous measure of the infant's gestational age at birth. *Preterm birth* is defined as an infant born before a gestational age of 37 weeks.
    iv. *Birth weight for gestational age* is a continuous measure of the infant's birth weight relative to the average in the population for a given gestational age at birth. *Small for gestational age* is defined as a birth weight below the 10th percentile for infants of the same gestational age.
    v. *Perinatal mortality* is a binary indicator for whether there was a fetal death at or after 20 weeks of gestation or mortality within the first 7 days of life.

D.3. Patient Characteristics

The data includes information on participants' household size, disability status, veteran status, work status, income, receipt of other benefits or assistance, and utilization/spending prior to randomization. These characteristics come from the lottery application form or electronic health records and are assessed prior to randomization. In the event that there are high levels of missingness for any key covariate (i.e., >2%), multiple imputation methods will be used.

E. Statistical Analysis

E.1. Evaluation of randomization, balance, and attrition

We will test for balance between treatment and control based upon observable baseline characteristics for the overall study population.

Because differential attrition correlated with treatment, for example due to death, could introduce bias into our results, we will also evaluate the attrition rate and assess for balance between treatment and control based upon both baseline characteristics for the final analytic sample (and attritors) and potential causes of attrition.

E.2. Statistical specification

Our primary analytic approach is an analysis based on the intent-to-treat principle that compares outcomes for those randomized into the treatment group to those who were randomized into the control group. Specifically, we will estimate the following linear regression model:

$$y_i = \beta_0 + \beta_1 \mathrm{CASH}_i + \beta_3 \mathbf{X}_i + \epsilon_i,$$

where y_i is the utilization or clinical outcome for individual or household i. See section D.2. for a list of our primary and secondary outcomes. "$CASH_i$" is an indicator for whether individual i won the lottery and was thus randomized into the treatment group. $\boldsymbol{X_i}$ is a vector of covariates, specifically patient characteristics, including baseline (pre-randomization) values of the outcome variable, which are not explicitly necessary since they should be unrelated to treatment status, but they may increase the precision of our estimates to the extent that they explain some of the variance in the outcome.

The coefficient on CASH_i, β_1, is our main coefficient of interest; it provides the difference in means between the treatment groups and the control group. In addition to our primary specification above that corresponds to testing for a level shift, we will also include an interaction term between an indicator for time and treatment status to test for changes in slope.

As described in section C, a household could receive additional lottery tickets by meeting specific criteria. Observations will thus be weighted by the inverse probability of winning the lottery so that β_1 is an unbiased estimate of the relationship between winning the lottery and the outcome. An alternative specification that includes indicators for number of lottery tickets is discussed below. Standard errors will be clustered at the household level and adjusted for heteroskedasticity.

E.3. Subgroup analyses

To examine heterogenous effects of the cash benefits, we will repeat our analyses for four pre-specified subgroups, each defined based on data from the pre-randomization period. The subgroups are: (1) chronic disease at baseline, (2) financial distress at baseline, and (3) poor self-reported health at baseline, or (4) any of the prior.

Specifically, we use electronic health record data to define participants by whether they have a known chronic disease (e.g., diabetes, hypertension, depression, anxiety). For a subset of the participants, we have baseline survey data that enables subgroup analyses by level of financial distress and self-reported health. Financial distress is defined as an affirmative answer on a survey question asking the participant whether she/he/they had any bills, expenses, or needs that they were unable to pay. Poor self-reported health is defined as a positive screening (score≥3) on the 2-question version of the Patient Health Questionnaire, a positive screening (score≥3) on the 2-question version of the Generalized Anxiety Disorder questionnaire, or a fair/poor response on the 5-point scale of self-rated health.

E.4. Alternative specifications and sensitivity analyses

Our primary specification includes patient characteristics and baseline values of the outcome in the model to improve power as well as any chance imbalance between the study arms after randomization. As described in section E.1., we will compare covariates between the treatment and control groups, and we will explore whether our results are sensitive to inclusion of these covariates in the model. Importantly, our primary specification accounts for differential probability of winning the lottery using inverse probability weights. As an alternative approach, we will include the set of covariates that are correlated with treatment probability into the model, specifically indicators for number of lottery tickets. Finally, as a sensitivity check, we will also conduct our analyses on the subpopulation of participants with established primary care at one of the clinical sites of care.

To ensure our estimates are robust to method of estimation, we will also estimate generalized linear models assuming a negative binomial distribution for the count utilization outcomes and a generalized linear model with a Bournoulli distribution and logit link function for the binary outcomes. In a final robustness check, we will Winsorize the utilization measures to ensure that our estimates are not sensitive to outliers.

E.5. Statistical significance and adjustments for multiple comparisons

Statistical significance was defined as two-sided $P<0.05$ for the primary outcome. Because we have one prespecified primary outcome, we will not make any adjustments for multiple inference. For our secondary outcomes, we will the Benjamini-Hochberg procedure or similar to calculate adjusted p-values that account for testing of multiple outcomes.