Mobilizing Voices for Science to Promote Trust in Vaccination:

The *Health Ambassadors* Multi-site Randomized Controlled Trial

Statistical Analysis Plan

Health Ambassadors Research Consortium:

Arsene Brice Bado*, CERAP Abidjan

Yapo Felix Boa, U. Félix Houphouët Boigny and CERAP Abidjan

Horace Gninafon, U. Laval and U.C. Berkeley/EGAP

Paul Kawale, Kamuzu U.

Fortunate Machingura, CeSHHAR Zimbabwe

Cyrus Samii, New York U.

Samba Corr Sarr, Senegal Ministry of Health

Mame Mor Anta Syll, U. Gaston-Berger

Mavuto Thomas, Malawi Ministry of Health

Maarten Voors, Wageningen U.

VERSION: July 12, 2024

* *Study authors listed in alphabetical order. All authors are contributors to the formulation, design, and implementation of the research.* *This research was organized as part of the Social Science Research Council's Mercury Project with generous funding from the Rockefeller Foundation.*

Analysis Plan

The study protocol describes the primary and secondary outcome variables, treatment variables, covariates, and design variables in detail. Here we explain how those variables will be used in the statistical analysis.

Estimating Average Effects with Household Data

For the vaccine trust and other attitudinal measures in the individual survey, we will use a hierarchical index strategy.[1] All outcome measures will first be standardized in terms of control group standard deviations. That is, we will construct a standardized sum of scores index for each of the 4 groupings listed above. Then, we will construct a global index that is a standardized sum of the grouping indices. Indexes will be oriented so that more positive values indicate a movement in the beneficial direction. For a given index, we will estimate average treatment effects using the following regression specification:

$$y_i = \alpha + \beta\, Treated_i \; + Rand.Block\; FE_{b[i]} \; + X_i'\gamma \; + \epsilon_i \;\; , \qquad (1)$$

where *i* indexes the individual-survey respondent for household *i*, and $Rand.Block\; FE_{b[i]}$ and the covariate vector X_i are defined the same way as in equation (1). We calculate cluster-robust standard errors taking the error term ϵ_{vji} to be clustered at the level of the randomization cluster. Again, we hypothesize that $\beta > 0$, and we test using a two-sided test (allowing for the possibility that effects could go in the opposite direction as what is hypothesized) with 95% confidence. We will use a hierarchical statistical significance testing strategy:

- For the global index, we will use the standard asymptotic p value and a two-sided null hypothesis with 95% confidence.
- If the global p-value rejects the null, then for the grouping-specific indices, we will use a Benjamini-Hochberg FDR correction for the 4 p-values.
- For any grouping index that rejects the null, we will use the Benjamini-Hochberg FDR correction for the item level p-values.

For estimating peer effects, we will use the same specification (2), with *I* indexing the peer respondent and $Treated_i$ indicating the treatment status of the person that recommended them.

For our behavioral outcome measured in the household roster, we will estimate the average treatment effect on the pooled vaccination rate using the following regression specification:

$$y_{vji} = \alpha + \beta\, Treated_i \; + Vaccine\; FE_v + Rand.Block\; FE_{b[i]} \; + X_{ji}'\gamma \; + \epsilon_{vji} \;\; , \qquad (2)$$

where *v* indexes the vaccine, *j* indexes the individual adult in the household roster, *i* indexes the household, and *b*[*i*] indexes the randomization block for household *i*. The variable $Treated_i$ is an indicator for whether the household is in a treated cluster or not. Then,

[1] The approach follows that of Casey et al. (2012).

$Vaccine\ FE_v$ are fixed effects (dummy variables) for the different vaccines and $Rand.Block\ FE_{b[i]}$ are fixed effects (dummy variables) for the randomization blocks. The covariate vector X_{ji} includes the following elements:

- Indicator for female of individual *j*,
- Age for individual *j*,
- Level of education for individual *j*,
- Pre-intervention vaccination rate in the household's cluster.

We calculate cluster-robust standard errors taking the error term ϵ_{vji} to be clustered at the level of the randomization cluster. We hypothesize that $\beta > 0$, and we test using a two-sided test (allowing for the possibility that effects could go in the opposite direction as what is hypothesized) with 95% confidence. If we reject the null, we will then test effect on each individual vaccine, using the Benjamini-Hochberg FDR correction for the vaccine-level p-values.

Effect Heterogeneity with Household Data

We propose an exploratory analysis of background conditions that may moderate the size of the treatment effects on behavior and attitudes. These include:

Respondent attributes:

- Gender,
- Age,
- Level of education,
- Ethnic group (coded for collectivist norms),
- Proximity to health center.

Health ambassador attributes:

- Gender,
- Age,
- Type of degree program,
- Co-ethnicity with respondent,
- Personality measure of empathy (empathy assessment scale).

This analysis will take advantage of the random assignment of health ambassadors to households. We will perform a "dummy" random assignment for the control group so that we know, for each control group household, which of the health ambassadors *would* have treated them, counterfactually.

These analyses will be conducted as separate one-way interaction regressions with the following specifications:

$$y_{vji} = \alpha + \beta_1\, Treated_i\ + \beta_2 W_{ji} + \beta_3 Treated_i * W_{ji} + Vaccine\ FE_v + Rand.Block\ FE_{b[i]} + X_{ji}'\gamma\ + \epsilon_{vji}\ , \qquad (4)$$

$$y_i = \alpha + \beta_1\, Treated_i\ + \beta_2 W_i + \beta_3 Treated_i * W_i\ + Rand.Block\ FE_{b[i]} + X_i'\gamma\ + \epsilon_i\ , \qquad (5)$$

where W_{ji} or W_i refers to the background characteristic in question. We will use FDR control on the p-values for the respondent attribute (as one family of estimates) and then the health ambassador attributes (as a separate family) interaction effect estimates.

Mechanisms through which the intervention may increase vaccine uptake

To identify the mechanisms through which the intervention may increase the primary outcome of interest, we estimate the following two equations:

$$M_{ji} = \alpha^M + \beta^M Treated_i + Rand.Block\ FE_{b[i]} + X_{ji}'\gamma^M + \epsilon_{vji}{}^M \quad (6)$$

$$Y_{ji} = \alpha^Y + \beta^Y Treated_i + \kappa M_{vji} + Rand.Block\ FE_{b[i]} + X_{ji}'\gamma^Y + \epsilon_{vji}{}^Y \quad (7)$$

The "first stage" effect of the Health Ambassador treatment on the mediating variables is captured by β^M and then the extent to which a mediating variable is responsible for transmitting the effect of the HA treatment onto vaccine outcomes is given by $\beta^M\kappa$. We will use the Imai et al. "mediation" software in R to implement this analysis. We hypothesize that the intervention will improve vaccine trust through subjects' acceptance of the benefits of vaccination, subjects' access to vaccination and affordability. Additionally, the subjects are aware of the availability of vaccines and the benefits and risks associated with the vaccine and are motivated to seek vaccination. Moreover, the subjects accept the medical system as a trustworthy healthcare provider.

Descriptive Analyses of Process and Context

We will prepare simple descriptive summaries of the process data from households and Health Ambassadors and the context data from the health facility authorities. These summaries will be used to understand more precisely how the intervention functioned and also what contextual factors may help to explain both patterns in the control group and patterns in effects.

Cost Analysis

We will conduct a Cost Analysis using the detailed input cost accounting methodology proposed by Dhaliwal et al. (2013).

References

Casey, Katherine, Rachel Glennerster, and Edward Miguel. "Reshaping institutions: Evidence on aid impacts using a preanalysis plan." The Quarterly Journal of Economics 127.4 (2012): 1755-1812.

Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch. "Comparative cost-effectiveness analysis to inform policy in developing countries: a general framework with applications for education." Education policy in developing countries 17 (2013): 285-338.