Analysis Plan

Project Name: **Increasing IPTp uptake in Nigeria**

Project Code: 1715

Date Finalized: March 5, 2018

This document serves as a basis for distinguishing between planned (confirmatory) analysis and any unplanned (exploratory) analysis that might be conducted on project data. This is crucial to ensuring that results of statistical tests will be properly interpreted and reported. In order that the Analysis Plan fulfill this purpose, it is essential that it be finalized and date-stamped before we begin looking at the data — ideally, before we take possession of the data. Once this plan is finalized, a date is entered above, and the document is posted publicly on our team website.

Outcome Variables to Be Analyzed:

Primary:

- Whether a pregnant mother did or did not receive at least one IPTp dose (binary)
- Number of IPTp doses completed by pregnant women

Secondary:

- Number of ANC visits by women in the treatment and control wards
- Correlation between ANC attendance and IPTp uptake
- Heterogeneous treatment effects on subpopulations if data is available on household characteristics (four particular variables were of greatest interest: prior ANC attendance, prior IPTp receipt, age at baseline, and education level at baseline). Note that this would constitute exploratory analysis, as the study was not powered to explore these relationships as central aspects of the analysis.
- Frequency of experiencing stock-outs in treatment wards via completed information on the Record Cards
- Drop off in ANC attendance after each numbered ANC visit (e.g. ANC visit #2 attendance conditional on having attended ANC visit #1, etc.

Data on IPTp uptake was collected from referral forms and health facilities via the DHIS2. Visit data was obtained from referral forms.

Additional Data:

- Demographics (e.g., age, religion, whether the woman is the only wife of the husband, education, prior ANC/IPTp uptake)
- HC3 Meeting attendance and effectiveness
- Pregnancy process, outcome, and complications
- Women's opinion on ANC/IPTP effectiveness
- Presence/absence of the husband at baseline.
- Facility data from the Record Card (right hand side).

Statistical Models & Hypothesis Tests:

Because there was random assignment at the ward level, the data analysis needed was relatively straightforward, involving regression analysis to estimate average treatment effects. Specifically, we will use the following specification:

$$Y_{iw} = \beta_0 + \beta_1 X_w + \beta_2 Z_i + \beta_3 \Upsilon_w + \beta_4 M_i$$

Where:

- Y_{iw} : Outcome Y for individual i, in ward w
- X_w : A dummy variable for whether mother's ward, w, was a treatment ward
- Z_i : A vector of individual-level demographic controls measured at baseline
- Υ_w : Controls for ward size (a continuous variable) and for LGA of the ward (fixed effects)
- M_i : Month-of-delivery measured as months after experiment initiation

In the analysis, we will use clustered standard errors at the ward level.

In the specification above, β_1 represents the estimate of interest - the causal effect of being in a treatment ward on the key individual-level outcomes in the study, controlling for individual characteristics, ward characteristics, and month of delivery relative to experiment initiation.

Data was stored on password-protected devices in the field and imported into Stata, which was used for analysis. Data was collected in three distinct waves: a baseline survey (first visit with mothers), a midline survey (two weeks after the baseline visit), and an endline survey (at the end of the intervention).

There were some intermediate outcome variables included in the midline survey, but the majority of the outcome variables analyzed will be those from the endline survey.

Note that we will not be able to fully disentangle whether it is the Record Card or the conversation with the husband that drives any treatment effects that we observe. That said, we will use the endline variables we collect on reported (and actual) use of the Record Card, as well as data on reported conversations with the husband about ANC/the Record Card, to conduct some exploratory analysis to try to better understand what might have been driving the effects.

We also will conduct a robustness check to determine the extent to which individuals in control and treatment wards differed on observable variables measured at baseline. This consists of a regression, whereby a dummy variable for treatment is regressed on individual-level, observable baseline variables; that is:

$$\text{Treatment} = 0 + 1Z_i + 2M_i$$

Where "Z_i" represents a vector of individual-level baseline characteristics and M_i represents month-of-delivery fixed effects. We would then use an F-test to assess the joint significance of the baseline characteristics in predicting assignment to treatment.

For robustness, and to provide a very conservative estimate, we will also group data at the ward level (compute means of outcome variables at the ward level) and conduct the analysis above for ward-level data only (with 72

data points). We do not anticipate this analysis will find a significant result, due to how much of the data will be collapsed, but we feel it a useful guide regarding how to interpret the analysis based on individual-level data.

Limitations:

The biggest limitation in the design was the relatively small number of clusters (72) inhibiting the power of the intervention. In contrast, the subjects per cluster is not an issue, barring a massive decrease in the monthly number of pregnant women seen per ward. Another limitation was the data, which is more survey-based than would be ideal, because of on-the-ground and technological constraints.