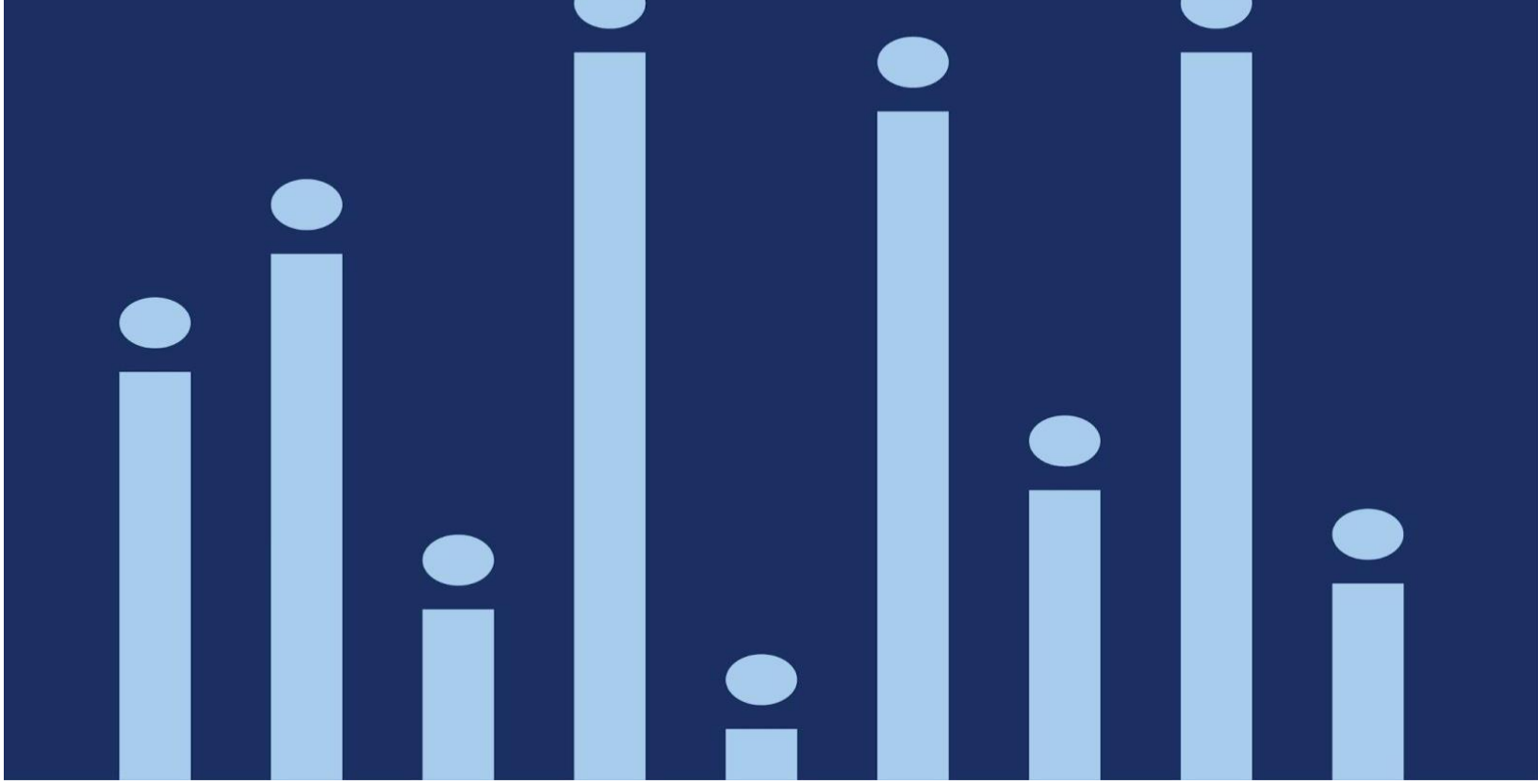


Pre-Analysis Plan - Testing the Effectiveness of Quality-Based WhatsApp Messaging on Vaccine Hesitancy and Childhood Immunization in a National Maternal Messaging Program (MomConnect)

IDinsight

September 20, 2023

NCT Number: NCT05505760

Table of Contents

Acronyms

ARM	Appointment reminder model
BCM	Browsable content model
RCM	Relevant content model
ICT	Information and Communication Technologies
MIS	Management Information Systems
PAP	Pre-Analysis Plan
RCT	Randomized Control Trial
TOC	Theory of Change
WA	WhatsApp

INTRODUCTION

OBJECTIVE

The enclosed Pre-Analysis Plan (PAP) outlines the objectives, methods, and analytical framework for the IDinsight evaluation of the effectiveness of Quality-Based WhatsApp Messaging on Vaccine Hesitancy and Childhood Immunization in a National Maternal Messaging Program (MomConnect) together with the organization implementing MomConnect, Reach (formerly Praekelt.org).

BACKGROUND & MOTIVATION

The case for mobile health (mHealth) has been gaining traction as a source to improve health outcomes of individuals. mHealth refers to the use of Information and Communication Technologies (ICT) to support health care (Lee et al. 2016). The ubiquity and penetration of mobile phones presents the opportunity to deliver health care services directly to citizens, with the greatest potential gains in under-resourced health ecosystems. Of what evidence does exist, literature reviews independently identify a wide range of instances where maternal health messaging is effective (Lee et al. 2015; Poorman et al. 2014). A recent small scale randomized control trial (RCT) of 177 mothers in one South African city found that enrolment in a maternal messaging program resulted in higher odds of administering all first-year child vaccinations and attending antenatal and postnatal clinic appointments (Coleman et al. 2020). Despite the widespread adoption of mHealth interventions, there has been relatively little research on their impact on maternal and infant health at a large scale in developing countries (Lee et al. 2015).

Operating for 8 years and supporting more than 1 million mothers through their pregnancies and early childhood care, MomConnect represents one of the largest maternal health messaging platforms in the world (Jahan et al. 2020). Implemented at the national level, MomConnect has been credited with being the first national-scale mHealth program of its kind and has won numerous international awards. While achieving some large starting successes, there is still scope for improving MomConnect's impact for mothers and their children (Mehl et al. 2018).

As a WhatsApp-delivered service, MomConnect has been limited by WhatsApp's historical terms of service. These limitations involve both the frequency and content of messages that the service can send mothers to start a conversation (i.e. "push messages"). In its current form, MomConnect can only start a conversation with mothers once a week by alerting them of their upcoming antenatal care (ANC) appointment. From these initial conversation starters, mothers can then engage with the platform by responding to the push message to learn more information relative to their stage of pregnancy or postpartum. MomConnect program aligns with the National Department of Health's National Development Plan: Vision 2030 third goal of reducing maternal and infant mortality. As MomConnect's content is approved by the National Department of Health (NDOH), it avoids women needing to navigate complicated, contradictory, or unverified information available on the internet. Recent changes to WhatsApp's terms of service have removed the restriction around push messages for certain programs, including MomConnect. Messages have historically been restricted to weekly conversation starters and carry only generic administrative reminder-style information. As a result, it is possible that many mothers are failing to engage with the content as much as they could. If mothers were more engaged, it could greatly increase exposure to and knowledge of accurate maternal and infant health content. With the updated terms of service, there is an opportunity to test different ways of engaging mothers.

There have been some studies of the early model of MomConnect content to measure its effects on health outcomes (Jahan et al. 2020; LeFevre et al. 2018; Mehl et al. 2018; Xiong et al. 2018; Skinner et al. 2017). However, these studies

were largely qualitative, at smaller scales, and focused more on user satisfaction of MomConnect. These studies also preceded the changes in WhatsApp's terms of services. As such, more evidence is needed to rigorously evaluate the impacts of MomConnect on maternal and child health outcomes. **This study proposes to evaluate, through a randomized control trial (RCT), the relative effect of three new behaviorally-informed program models on key knowledge of and adoption of healthy behaviors.**

EVALUATION

INTERVENTION

The randomized control trial will consist of one control arm, resembling the status quo MomConnect messaging model, and three new treatment arms. The study will evaluate the relative effect of three new behaviorally–informed program models on key health outcomes. The table below summarizes the different models (treatment arms) proposed in this study:

Table 1: Treatment Arms

Name	Form	Description
ARM	WhatsApp	**The appointment reminder model (ARM) is the current MomConnect WhatsApp model (control).** Mothers receive weekly conversation starter messages reminding them about their upcoming clinic appointments, providing more comprehensive and relevant maternal health information only after mothers respond to the appointment reminder.
RCM	WhatsApp	In the relevant content model (RCM), Mothers receive weekly conversation starter messages, which carry both clinic appointment reminders **along with some maternal and infant health information relevant to their pregnancy/postpartum stage**. In addition, a list of "frequently asked questions" (FAQs) relevant to the week of pregnancy the mother is in are provided so that mothers can engage further with maternal health information topics relevant to them.
	SMS	Mothers receive **twice weekly conversation starter messages** of 160 characters each, which carry both clinic appointment reminders as well as **maternal and infant health information**, relevant to their stage of pregnancy or the age of their baby. Mothers can access the list of frequently asked questions relevant to their week of pregnancy via USSD.
RCM + BCM	WhatsApp	In addition to the RCM, this arm also includes the browsable content model (BCM). The BCM adds **an option to browse a menu of relevant maternal and infant health information** that mothers can explore themselves over WhatsApp. All in all, the RCM+BCM arm includes appointment reminders, clinical information, a browsable menu, and prompts to relevant stage-based topics. Mothers also receive weekly conversation starter messages as in all other models.

RESEARCH QUESTIONS

The table below highlights the key research questions that the study aims to answer.

Table 2: Research Questions

Category	Research Question	Decision Relevance
Primary Research Questions		
Health Outcomes	Which version of MomConnect improves key maternal and infant health outcomes?	Expansion of the most impactful MomConnect treatment arm across South Africa
Secondary Research Questions		
Message Delivery	Are maternal and infant health outcomes better if MomConnect is administered over SMS as compared to WhatsApp for the RCM model?	Decision on standardizing the RCM model in SMS or WhatsApp format
Directed Attention	Which aspects of the content models do women engage with the most?	Improved MomConnect message format
Perceived reliability	Which health-related topics do pregnant women find most reliable/trustworthy?	Improved MomConnect message content
Socio-economic characteristics	Which socio-economic characteristics are correlated with higher/lower effectiveness of MomConnect?	Adjusting content for groups for which MomConnect is less effective

STUDY DESIGN

The impact of MomConnect's models will be evaluated using a randomized controlled trial (RCT). Randomization will be at the individual level, where an individual refers to the "unique user", which is defined by the unique phone number used at the time of registration to MomConnect at health facilities. Randomizing control and treatment allocations ensures that the study groups will be comparable in terms of observable and unobservable characteristics in expectation. Therefore, statistical inference can shed light on the likelihood that any differences in outcome variables at the end of the intervention were caused by the intervention as compared to chance. We will also study the treatment effects on different outcomes between the different arms of the experiment - comparing outcomes across different treatment arms or outcomes in the control to outcomes in treatment arms will help us answer the aforementioned research questions.

STUDY POPULATION & INCLUSION CRITERIA

This study will recruit pregnant women as per the requirements listed in the table below:

Table 3: Inclusion Criteria

Inclusion Criteria	Justification
Pregnant women who visit	The relevant content has been developed for the 16th week of pregnancy and

health facilities and are between their 16th-30th week of pregnancy	beyond. Due to timeline and sample size considerations, the last date of eligibility will be of women in their 30th week of pregnancy. Additionally, this will ensure that respondents have used the platform for a time period long enough to have experienced an adequate dosage of the treatment effects.
Pregnant women with WhatsApp capability	Given that we do not want users in the SMS arm to be different from those in the WhatsApp arm, differences between control and treatment arms are mitigated under this criterion. We recognize that this will exclude a proportion of the population that do not own smartphones/cannot access WhatsApp and our results will only be representative of smart-phone-owning women whose phones have WhatsApp capabilities.
Pregnant women above 18 years	Consent needs to be received from an adult. For logistical reasons, it is not feasible for us to obtain consent from guardians for potential study participants who are below 18 at the start of the study. Therefore, the study will only enroll women above the age of 18 who can give consent to participation. While underage pregnancy is an issue in South Africa and MomConnect messages are available to regardless of age, underage pregnant mothers make up only 2% of MomConnect users. We therefore think the resulting loss of external validity from excluding users below the age of 18 is limited.

RECRUITMENT & RANDOMIZATION PROCEDURE

Pregnant women presenting at public health facilities across South Africa are asked by nurses, in accordance with national health guidelines, if they would like to join the MomConnect platform. This process has successfully registered 60% of pregnant women across the country (Benjamin et al. 2018). The recruitment process for this study will follow the same procedure and is unchanged by the study.

Women will register for MomConnect at a health facility by filling out a USSD recruitment survey. The survey asks specific questions about each pregnant woman. From the recruitment survey, we will learn the gestational age of pregnant women, how old they are, and which province they are from. Mothers will then be asked if they would like to take part in a research study to improve the MomConnect service and will be sent a consent message via WhatsApp. Consenting participants will then be assigned to different strata (one of three categories for gestational age: 16-20 weeks, 21-25 weeks, 26-30 weeks; one of two categories for age: 30 and under, over 31; and one of nine categories for province). Next, they will be sorted into one of 54 strata according to their unique combination of weeks pregnant, age, and province. Because recruitment is occurring on a rolling basis, we use sequential randomization within strata to improve balance within strata, across study arms. Participants will then be assigned one of the four messaging arms according to an ex-ante random schedule within each stratum. In this way, there will be a similar way of women assigned to each of the four arms within each stratum.

After consent and randomization participants will be invited to a (non-mandatory) baseline survey via WhatsApp (or USSD for the RCM-SMS arm) and will start to receive content based on which of the four study arm message models they were randomly assigned to. Women who do not give consent to be part of the study will receive the status-quo WhatsApp MomConnect service (equivalent to the Control arm) and be excluded from the study.

SAMPLE SIZE

Since we are comparing different versions of MomConnect, it is important that the experiment is well-powered with sufficient sample size to detect reasonably small differences between treatment arms. A sample size of 5,200 completed endline interviews enables the experiment to have 80% power to detect a reasonably small effect size of a 5 percentage point difference in binary outcomes between individual treatment arms.

The table below outlines a summary of the power calculations and the sample sizes for the study:

Table 4:Sample Sizes

Sample Sizes	
Per control arm (1 arm) at Endline	1300
Per treatment arm (3 arms) at Endline	1300
Total Sample Size Overall (expected at endline)	5200
Number of MomConnect users to enroll in the study	8667
Minimum detectable effect size between different treatment arms	5 percentage points
Assumptions	
There are around 8,667 women who will register for MomConnect while 16-30 weeks pregnant and consent to the study during the study period. This is based off of MomConnect data from 2022.	
Based on previous surveys with MomConnect users, we expect an endline completion rate of around 60%. Thus, with 8,667 enrolled participants, we should reach the desired sample size of 5,200.	
Based on previous data on the share of babies of mothers enrolled on MomConnect, we are assuming a control group estimate of 68% for full infant vaccination 7 weeks after birth.	
Parameters	
Alpha: the level of statistical significance	0.05 This is the research standard

Power: the probability of avoiding a Type II error ("false negative")	0.80 This is the research standard
Unit of Treatment Assignment	Individual
Predicted Mean of Outcome Variables: the estimated indicator value	0.68 We assume that 68% of the control group will have their children fully immunized with the NDOH-recommended vaccinations from birth to 6 weeks post-birth, measured during an endline interview around 7-10 weeks after birth.

We will have the same power for detecting a 5 pp. increase for our other primary outcome variable, a binary indicator taking on value 1 if a mother completed at least 8 ANC visits as recommended by the WHO and 0 otherwise. Our two secondary outcomes, the knowledge index and the behavior index, we will only measure for 2,200 endline respondents instead of the full sample of 5,200 respondents. With an alpha of 0.05 and power of 0.8, we will be powered to detect a MDES of 0.17 standard deviations for both indices.

DATA SOURCES & DATA MANAGEMENT

We will obtain data from the following sources for the analysis of study outcomes:

1. **MIS backend data** - the back-end of the WhatsApp Chat bot will provide details on how users engaged with the messages they received. Reach already collects this information and will provide relevant data to IDinsight. IDinsight will use this data to track outcomes including the responses users send and their engagement with the platform.
2. **Primary data** - IDinsight will design and administer surveys with pregnant women and mothers via WhatsApp, USSD, and phone. Women will be given a small incentive in exchange for participating in the surveys. These surveys will help measure primary and secondary outcome measures. There will be 3 rounds of primary data collection: baseline, midline, and endline. All baseline and midline surveys will be conducted via WhatsApp (USSD for SMS arm) and endline data will be collected via mobile phone surveys.

ANALYSIS FRAMEWORK

UNIT OF ANALYSIS

The unit of analysis is the individual.

OUTCOME INDICATORS

The measurement tools differ for each survey round. This is because the baseline and midline surveys are conducted over WhatsApp/USSD during pregnancy whereas the endline survey is conducted via phone surveys starting 7 weeks post-delivery. The tables below summarize the primary and secondary outcomes that will be analyzed for each round.

Primary Outcomes

MomConnect's theory of change posits that if pregnant women and mothers have access to content related to health guidelines, they will be equipped with knowledge about healthy pregnancy and postpartum behaviors, and this would translate to practicing these behaviors. The primary outcomes of interest are whether newborn babies have received all child immunizations recommended 6 weeks after birth and whether women attended at least 8 ante-natal care visits during her pregnancy.

For the number of immunizations, we ask mothers about the immunizations their baby received at birth (recommended to be OPV0 and BCG) and six weeks after birth (recommended to be OPV1, DTaP-IPV-Hib-HepB, RV, and PCV) during a phone survey starting 7 weeks after birth. Similarly, we will ask mothers about the number of antenatal care visits they went to at their clinic over the course of their pregnancy. Immunizations and ANC visits have been identified as the most decision-relevant outcomes to base future programming decision on by MomConnect stakeholders.

Secondary Outcomes

Our secondary outcomes include indices based on a number of questions related to both knowledge of healthy behaviors as well as actual adoption of healthy behaviors. We calculate these two indices using primary data from our endline survey. Other secondary outcomes relate to the user experience of the MomConnect platform, based on both backend data and some survey questions.

The knowledge and behavior outcomes will be coded as two summative indices with equal weights given to the different topic areas of interest. The topic areas and questions within topic areas were determined together with Reach based on the content of MomConnect. This is done for two reasons: first, MomConnect content covers a variety of topic areas, and other than immunizations and ANC visits, no single topic was identified as more important or decision-relevant than others; second, compared to the alternative of analyzing each topic area separately, collapsing discrete topic areas into two indices reduces the risk of false positive results from testing a large number of hypotheses. The success of each treatment arm relative to control will therefore be based on demonstrating impact on administered child immunizations six weeks after birth, the number of ANC visits pregnant women went to during their pregnancy, mothers' knowledge/attitudes, and mothers' behavior.

In the two summative indices, each topic area is given an equal weight in the overall index, regardless of how many questions are included in each topic area. A score multiplier will ensure that each topic area is equally weighted. We

chose to equally weigh each topic because all content areas are an equally important contributor to overall health. To ensure that each topic area has an equal weight, we do the following *for each outcome*:

- Count the number of topic areas (N)
- Count the total number of points that mothers could earn *within* each topic area (S)
- Sum the total number of points *across* all topic areas (T)
- Calculate the overall score per topic area, which is the sum of all points a mom received from survey response (A)

Each topic area score (A) will be multiplied by a multiplier:

$$A * \frac{T}{NS}$$

To calculate the overall score for the outcome, the above formula will apply to each topic area, and then scores will be summed across all topic areas. Therefore, the overall score for each outcome will be at maximum *T* and regardless of how many questions are asked within each topic area, each topic will be given at maximum $\frac{T}{N}$ points. The overall score will represent knowledge/attitudes or behaviors of healthy behaviors overall (across all topics).

The table below summarizes the primary outcomes across midline and endline:

Table 5: Knowledge and Behavior Indices

Outcome	Definition
Endline Outcomes	
Immunisations	Number of immunisations baby received at birth and 6 weeks after birth (continuous variable with range 0-6)
ANC Visits	Number of ANC visits mother attended at clinic during her pregnancy (continuous variable)
Knowledge of/ Attitudes toward Healthy Behaviors (Max score = 12)	$((Baby\ danger\ signs\ score * \frac{4}{7}) + (Breastfeeding\ score * \frac{4}{3}) + (Anemia\ score * \frac{4}{5})$ *where:* Baby danger signs score (0-7) = the sum of the following points: • 1 point if mothers know that it is dangerous for babies to sleep on their tummies and with a blanket • 1 point if mothers identify that her baby having difficulty eating is a danger sign • 1 point if mothers identify that her baby having difficulty breathing is a danger sign • 1 point if mothers identify that her baby having a high fever is a danger sign • 1 point if mothers recognize that her baby not moving is a danger sign • 1 point if mothers recognize that her baby having a seizure is a danger sign • 1 point if mothers recognize that her baby having jaundice is a danger sign Breastfeeding score (0-3) = the sum of the following points:

	• 1 point if mothers intend on giving her baby only breastmilk for 6 months • 1 point if mothers only intend on introducing her baby to water and food after 6 months • 1 point if mothers know they can get pregnant even while breastfeeding Anemia score (0-5) = the sum of the following points: • 1 point if mothers correctly identify at least one correct (and no incorrect) way(s) to prevent anemia • 1 point if mothers correctly identify that cereals, nuts, or raisins are rich in iron (and they do not mention any food not rich in iron) eggs, meat, fish, and green vegetables are rich in iron • 1 point if mothers correctly identify that meat, fish, or liver are rich in iron (and they do not mention any food not rich in iron) • 1 point if mothers correctly identify that green leafy vegetables or beetroot are rich in iron (and they do not mention any food not rich in iron) • 1 point if mothers correctly identify that lentils or beans are rich in iron (and they do not mention any food not rich in iron)
Adoption of healthy behaviors (Max score = 16)	$((Diet\ score) + (Pregnancy\ preparedness\ score * \frac{4}{7}) + (Breastfeeding\ \ score * 2) + (Anemia\ treatment\ score))$ *where:* Diet score (0-4) = the sum of the following points: • 1 point if mothers report eating fruits on a daily basis during pregnancy • 1 point if mothers report eating vegetables on a daily basis during pregnancy • 1 point if mothers report eating liver not more than twice a month during pregnancy • 1 point if mothers report not drinking alcohol during pregnancy Pregnancy preparedness (0-7) = the sum of the following points: • 1 point if mothers reported packing a pregnancy bag to be prepared for her delivery • 1 point if mothers packed her maternity record in her pregnancy bag • 1 point if mothers packed a towel, facecloth, or toiletries in her pregnancy bag • 1 point if mothers packed baby wipes, nappies, or bum cream in her pregnancy bag • 1 point if mothers packed clean clothes for herself in her pregnancy bag • 1 point if mothers packed clean clothes for her baby in her pregnancy bag • 1 point if mothers packed a blanket, hat, or socks in her pregnancy bag Breastfeeding score (0-2) = the sum of the following points: • 1 point if mothers are currently only feeding her baby breastmilk • 1 point if mothers attempted to breastfeed their baby within the first hour of birth Anemia treatment score (0-4) = the sum of the following points: • 1 point if mothers were tested for anemia during her pregnancy • 1 point if mothers possessed iron folic acid pills during her pregnancy • 1 point if mothers consumed iron folic acid pills during her pregnancy • 1 point if mothers consumed iron folic acid pills on a daily basis during her pregnancy

To capture the user experience on MomConnect, we will also analyze the following outcomes (which are based on a mixture of MIS data and user-experience data captured in primary endline surveys).

Table 6: Secondary Outcomes

Outcome	Definition
Survey enrollment	Indicator for whether a mother initiates surveys or feedback loops
Survey completion	Indicator for whether a mother completes the survey
Engagement	Total number of messages sent by mother
Rate of successful message delivery	Total number of messages successfully delivered / total number of message attempts planned
Rate of read messages	Total number of messages read by users / total number of messages delivered
Opt-out	Indicator for whether a mother opts-out of MomConnect
Conversation starter	Total number of conversation starters a mother reads
Clinic reminders	Total number of clinic reminders a mother receives
Keyword messages sent	Total number of keyword messages that a mother sends across her journey
Non-keyword messages sent	Total number of non-keyword messages mothers send
Internet usage	Indicator for whether a mother used the internet to find information on pregnancy, childbirth, and newborn care
Trustworthy score	Indicator for whether a mother rates MomConnect as a very or completely trustworthy as a source for information on pregnancy, childbirth, and newborn care
Trustworthy reasons	Indicator for whether a mother thinks MomConnect is a trustworthy source of information because of NDOH's push
	Indicator for whether a mother thinks MomConnect is a trustworthy source of information because of a doctor's recommendation
	Indicator for whether a mother thinks MomConnect is a trustworthy source of information because of the appointment reminders
	Indicator for whether a mother thinks MomConnect is a trustworthy source of information because of the responsiveness of the platform
	Indicator for whether a mother thinks MomConnect is a trustworthy source of information because of the level of detail in the content
	Indicator for whether a mother thinks MomConnect is a trustworthy source of information because it helped her have a safe pregnancy
Best message	Indicator for whether a mothers selected various aspects of each message as the best part

content	of the message
Message helpfulness	Indicator for whether a mother answered "somewhat helpful" or "very helpful" to each aspect of each message
Perceived support from MomConnect	Indicator for whether a mother answered "Very supported" or "somewhat supported" regarding the helpfulness of MomConnect messages

EMPIRICAL SPECIFICATION

The main analysis of the primary and secondary outcomes described above will be based on the "intent to treat" impact of the different MomConnect messaging arms. Because of stratified randomized assignment, we expect the only systematic difference between mothers in different study arms to be the type of messaging they receive. The "intent to treat" effect refers to the impact of signing up for MomConnect and being assigned to one of the treatment messaging arms on primary and secondary outcomes, relative to the status-quo arm. Each arm involves a different method of engaging with content about healthy pregnancies. Given the arbitrary nature of defining "treatment compliance" in this context (i.e. mode of engagement varies message to message) - our analysis will focus on estimating average treatment impact among those who signed up for and *were sent* the messages. The ITT estimate is the decision-relevant estimator as this provides a picture of the impact of introducing the program at scale. The analysis of primary and secondary outcomes will follow from the model specified below.

Primary Specification

For our primary research question, we will estimate the following regression specification in order to understand the impact of each treatment arm on the four primary outcomes (immunizations, ANC visits, knowledge/attitudes index, behavior index):

$$Y_{is} = \alpha + \sum_{j} \beta_j T_{jis} + \delta_s + \epsilon_{is}$$

- Y_{is} is the midline or endline primary outcome of mother i in stratum s
- T_{ijs} is the treatment indicator for mom i and treatment arm j in stratum s. β_j is the impact of treatment arm j relative to the status quo model (control).
- δ_s is a vector of dummies corresponding with each stratum (province by age-range category by weeks pregnant category at time of recruitment)
- ϵ_{is} is the individual level error term. We will estimate heteroskedasticity-robust standard errors using the Huber-White sandwich estimator.

In primary analysis, we will use the separate specification to test the differences in outcomes between the WhatsApp ARM control arm and each of the two other WhatsApp treatment arms (RCM, and RCM + BCM) as well as the difference between RCM and RCM+BCM and the difference between RCM and RCM-SMS.

External Validity and Gestational Age

We will weight observations in our analysis sample by the inverse of the ratio of the number of users of their gestational age in our sample and the number of users of their gestational age in the population of MomConnect users who are above 18. Due to our enrolment schedule, our sample will overrepresent pregnant women who register for MomConnect at a later gestational age than is the case for the overall population of women on MomConnect. Since these women may be different on unobserved characteristics correlated with treatment effectiveness and women registering later will be exposed to MomConnect for a shorter time, post-stratification weights will potentially improve the external validity of the estimates when extrapolating results beyond the specific sample to the population of all MomConnect users.

Subgroup analysis

We will also do some exploratory subgroup analysis based on the following characteristics:

1. **Age:** We might expect mothers who are above 35 (advanced maternal pregnancy) to act differently than mothers who are below 35 given the higher risks associated with pregnancies for mothers above 35.
2. **The number of children the mom has**: We might expect first-time mothers to rely more on MomConnect than incumbent mothers who could pull from previous experience.
3. **Education level:** Since education levels are associated with socioeconomic status, we might expect mothers with less formal education to have less access to knowledge and resources on how to adopt healthy behaviors
4. **Income quartile:** We might expect mothers from poorer households to have less access to knowledge and resources on how to adopt healthy behaviors.
5. **Province:** The health service infrastructure differs strongly across South African provinces. We will explore in which provinces MomConnect is particularly effective.
6. **Weeks pregnant at sign-up:** The earlier in their pregnancy mothers sign-up with MomConnect, the longer they are exposed to messaging. We would therefore expect mothers signing up earlier to see larger benefits. We will use the stratification bins for weeks pregnant for this analysis.

LIMITATIONS AND CORRECTIONS TO DATA

Attrition

Attrition due to opt outs

There is a possibility that mothers opt out of receiving MomConnect messages. Under South Africa's Protection of Personal Information (POPI) Act, if users want to opt out of messaging, they can choose to delete their information from MomConnect, including their contact information. This is a threat to internal validity if the mothers who drop out are different from the mothers who continue to receive messages in ways correlated to knowledge and adoption of healthy behaviors.

Based on previous data from on opt out rates for the status quo of MomConnect, this is a very low proportion of mothers. However, in order to further reduce the risk of attrition bias, Reach has introduced a question in the user flow for mothers to specify what they are opting out of. If they are opting out of messages, they can therefore still be contacted for our surveys.

Attrition due to non-response

There is a high likelihood that some mothers will choose not to respond to some or all questions in the surveys. We expect a response rate of at least 60% for phone surveys. Bias can occur if the mothers who don't respond to surveys differ meaningfully from those who do respond across treatment arms. In order to maximize response rates, we will incentivize survey responses by offering airtime credits. We will also utilize different callback protocols during our endline phone survey to maximize response rates and make up to 7 attempts for each study participant.

Checks for attrition

We will check for differential attrition to see if mothers who respond are different across messaging arms. In order to do this, we will regress a binary variable representing response on treatment status and use the same controls as defined in the main specification. We will also check if mothers who respond to our endline survey differ across baseline variables across treatment arms. We will also check for selective attrition to see if baseline characteristics differ across responders and non-responders in general. This will be informative about the generalizability of our results with respect to the whole user base of MomConnect. We will also apply bounds assuming worst and best-case scenarios following Horowitz and Manski (2000) to bound treatment effect estimates.

Missing Values

Amongst mothers who do respond to our survey, some might respond "don't know" or "refuse." For knowledge based questions, we will treat "don't know" responses and refusals as incorrect answers as it is a fair assumption that mothers who skip responding to a quiz question are choosing to skip because they do not know the correct answer. For behavioral questions, we will treat "don't know" responses and refusals as missing given the likelihood that the mom might legitimately not recall doing a particular behavior.

Multiple Hypotheses Adjustments

The most relevant decision from this study to inform scale-up of the program will be informed by which WhatsApp treatment arm is best for key primary health outcomes. We will test the following 8 null hypotheses:

- The difference in $outcome_i$ between ARM and RCM is zero
- The difference in $outcome_i$ between ARM and RCM + BCM is zero
- The difference in $outcome_i$ between RCM and RCM + BCM is zero
- The difference in $outcome_i$ between RCM and RCM + SMS is zero

where the subscript *i* captures the two main outcomes, immunizations and ANC visits.

Given that there are multiple hypotheses to test, standard statistical significance levels ($\alpha = 0.05$) would likely result in finding significant outcomes by chance alone. Indeed, the probability of finding at least one false significant result (rejecting the null when it is true), would be $1 - (1 - 0.05)^8 = \sim 34\%$.

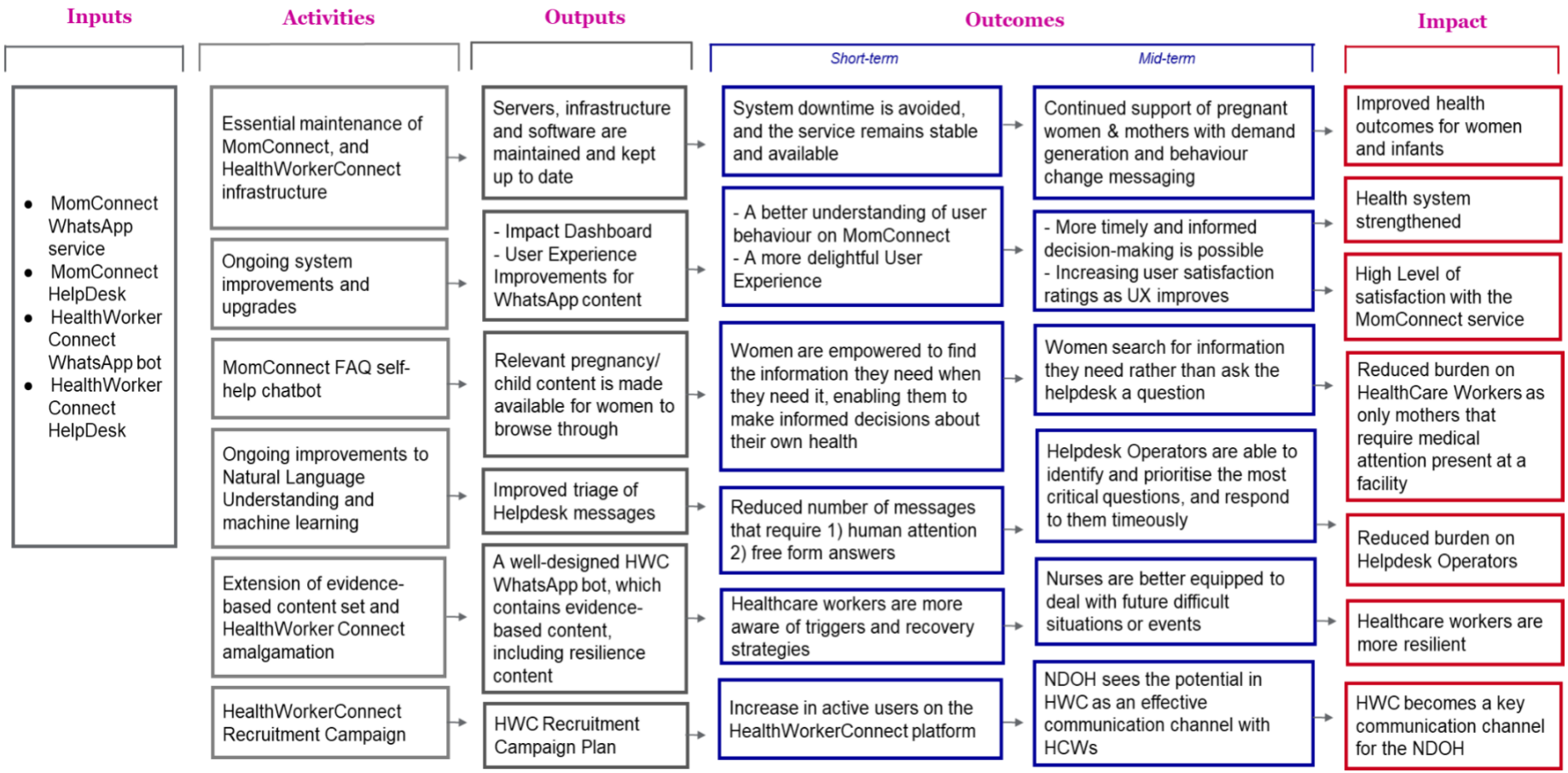
To correct for this, we will adjust for multiple hypotheses by applying a false discovery rate adjustment (FDR) following Benjamini et al. (2006). In particular, we will use the sharpened q-values discussed in Anderson (2008). This correction will control for the expected proportion of Type 1 errors. In our analysis, we report both corrected and uncorrected p-values.

These multiple hypothesis adjustments will focus on analysis of the four treatment arms and two primary outcomes specified given that these hypotheses affect potential scale-up decisions. Secondary outcomes will be exploratory.

APPENDIX

Appendix A: MomConnect Logic Model

Figure 1: MomConnect Logic Model



REFERENCES

Abiodun, M.O, Ljaiya, M.A., Aboyeji, P.A. "Awareness and knowledge of mother-to-child transmission of HIV among pregnant women." Journal of the National Medical Association 99(2). 2007. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2574348/pdf/jnma00206-0060.pdf

Anderson, M. L. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association, 103(4*84), 1481-1495. 2008. https://doi.org/10.1198/016214508000000841

Almutairi, W.M., Alsharif, F., Khamis, F., Salla, L.A. et al. "Assessment of mothers' knowledge, attitudes, and practices regarding childhood vaccination during the first five years of life in Saudi Arabia." Nurs. Rep. 11. 2021. https://doi.org/10.3390/nursrep11030047

Baleta, A. "Dramatic drop in SA's immunization rates." https://www.spotlightnsp.co.za/2020/06/24/dramatic-drop-in-sas-immunization-rates/

Belloni, A., Chernozhukov, V., & Hansen, C. "Inference on treatment effects after selection among high-dimensional controls." The Review of Economic Studies, 81(2), 608-650. 2014. https://doi.org/10.1093/restud/rdt044

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. "Adaptive linear step-up procedures that control the false discovery rate." Biometrika 93, no. 3: 491-507. 2006. https://EconPapers.repec.org/RePEc:oup:biomet:v:93:y:2006:i:3:p:491-507

Coleman, J., Black, V., Thorson, A.E. et al. "Evaluating the effect of maternal mHealth text messages on uptake of maternal and child health care services in South Africa: a multicentre cohort intervention study." Reprod Health 17(160). 2020. https://doi.org/10.1186/s12978-020-01017-3

Domek, G.J., O'Leary, S.T., Bull, S., Bronsert, M. et al. "Measuring vaccine hesitancy: field testing the WHO SAGE working group on vaccine hesitancy survey tool in Guatemala." Vaccine 36(35). 2018. https://doi.org/10.1016/j.vaccine.2018.07.046

Horowitz, J. L., & Manski, C. F. "Nonparametric analysis of randomized experiments with missing covariate and outcome data." Journal of the American statistical Association, 95(449), 77-84. 2000. https://doi.org/10.1080/01621459.2000.10473902

Lakshman, R.R., Landsbaugh, J.R., Schiff, A., Hardeman, W., et al. "Development of a questionnaire to assess maternal attitudes towards infant growth and milk feeding practices." International Journal of Behavioral Nutrition and Physical Activity 8(35). 2011. https://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-8-35

Lee, S.H., Nurmatov, U.B., Nwaru, B.I., Mukherjee, M., Grant, L., Pagilari, C. "Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: systematic review and meta-analysis." Journal of Global Health 6(1). 2016. https://doi.org/10.7189/jogh.06.010401

LeFevre, A.E., Dane, P., Copley, C.J., et al. "Unpacking the performance of a mobile health information messaging program for women (MomConnect) in South Africa: evidence on program reach and messaging exposure." BMJ Global Health. 2018. https://gh.bmj.com/content/bmjgh/3/Suppl_2/e000583.full.pdf

Leshi, O., Samuel, F.O, Ajakaye, M.O. "Breastfeeding knowledge, attitude and intention among female young adults in Ibadan, Nigeria." Open Journal of Nursing 6. 2016. https://www.scirp.org/pdf/OJN_2016011816100937.pdf

Jahan, R., Zou, P., Huang, Y., & Jibb, L. "Impact of MomConnect program in South Africa: a narrative review." Online Journal of Nursing Informatics. 2020. https://www.himss.org/resources/impact-momconnect-program-south-africa-narrative-review

Mehl, G.L., Tamrat, T., Bhardwaj, J., et al. "Digital health vision: could MomConnect provide a pragmatic starting point for achieving universal health coverage in South Africa and elsewhere?" BMJ Global Health. 2018. https://gh.bmj.com/content/bmjgh/3/Suppl_2/e000626.full.pdf

National Department of Health, South Africa. "Expanded Programme on Immunization - EPI (SA) Revised. Childhood Immunization Schedule from December 2015." 2015. https://www.health.gov.za/immunization/

Ndwandwe, D., Nnaji, C.A., Mashunye, T., Uthman, O.A., Wiysonge, C.S. "Incomplete vaccination and associated factors among children aged 12-23 months in South Africa: an analysis of the South African demographic and health survey 2016." Human Vaccines & Immunotherapeutics 17(1). 2021. https://doi.org/10.1080/21645515.2020.1791509

Opel, D.J., Mangione-Smith, R., Taylor, J.A., Korfiatis, C. et al. "Development of a survey to identify vaccine-hesitant parents: the parent attitudes about childhood vaccines survey." Human Vaccines 7(4). 2011. https://doi.org/10.4161/hv.7.4.14120

Peter, J., Benjamin, P., LeFevre, A.E. et al. "Taking digital health innovation to scale in South Africa; ten lessons from MomConnect." BMJ Global Health. 2018. https://doi.org/10.1136/bmjgh-2017-000592

Poorman, E., Gazmararian, J., Parker, R.M., Yang, B., Elon, L. "Use of Text Messaging for Maternal and Infant Health: A Systematic Review of the Literature." Maternal Child Health Journal, 19(5). 2015. https://doi.org/10.1007/s10995-014-1595-8

Pop-Eleches, C., Thirumurthy, H., Habyarimana, J.P. et al. "Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders." AIDS 25(6). 2011. https://doi.org/0.1097/QAD.0b013e32834380c1

Skinner, D., Delobelle, P., Pappin, M., et al. "User assessments and the use of information from MomConnect, a mobile phone text-based information service, by pregnant women and new women in South Africa." BMJ Global Health. 2017. https://gh.bmj.com/content/bmjgh/3/Suppl_2/e000561.full.pdf

StatsSA. "Maternal Health Indicators: Further Analysis of the 1998 and 2016 South Africa

Tamrat, T., and Kachnowski, S. "Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world." Maternal and Child Health Journal 16(5). 2012. https://link.springer.com/article/10.1007%2Fs10995-011-0836-3

United Nations. Sustainable Development Goal 3: Good Health and Wellbeing [Internet]. 2015. Available from: https://www.un.org/sustainabledevelopment/health/)

United States Agency of International Development. “Guidance for formative research on maternal nutrition.” Infant & Young Child Nutrition Project. 2012.
http://www.iycn.org/files/IYCN_Maternal_Nutrition_Research_Guidance_022112.pdf

Van Lier, A., Ferreira, J.A., Mollema, L. et al. “Intention to vaccinate universally against varicella, rotavirus gastroenteritis, meningococcal B disease and seasonal influenza among parents in the Netherlands: an internet survey. Bio Medical Central Res Notes 10. 2017. https://doi.org/10.1186/s13104-017-3004-z

Vander Wyst, K.B., Vercelli, M.E., O’Brien, K.O., Cooper, E.M., Pressman, E.K., Whisner, C.M. “A social media intervention to improve nutrition knowledge and behaviors of low income, pregnant adolescents and adult women.” Plos One. 2019. https://doi.org/10.1371/journal.pone.0223120

Xiong, K., Kamunyori, J., Sebidi, J. "The MomConnect helpdesk: how an interactive mobile messaging programme is used by women in South Africa." BMJ Global Health. 2018.
https://gh.bmj.com/content/bmjgh/3/Suppl_2/e000578.full.pdf