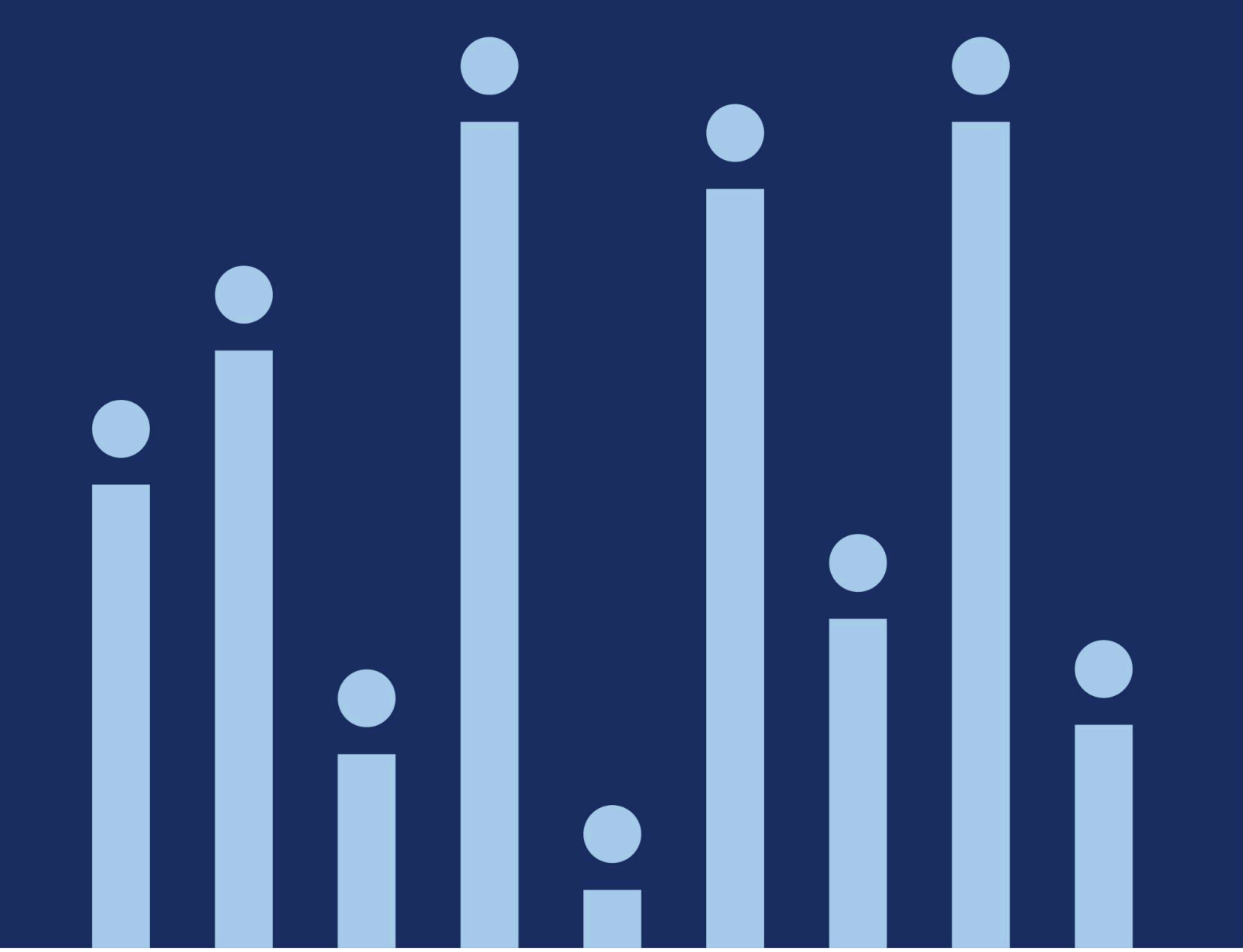


MomConnect - Quality-Based Messaging RCT

Final Study Report

April 2024

health
Department:
Health
REPUBLIC OF SOUTH AFRICA

Meta

REACH Health made possible

IDinsight

IDinsight

MomConnect - Quality-Based Messaging RCT

April 2024

Authors
Philipp Kastrau: **Philipp.Kastrau@IDinsight.org**
Lorry Symington**: Lorry.Symington@IDinsight.org**
Anjani Balu: **Anjani.Balu@IDinsight.org**
Rachel Lusava**: Rachel.Lusava@IDinsight.org**
Alison Connor**: Alison.Conner@IDinsight.org**

Key Contributors to the study
Lauren Kotze: **Lauren@reachdigitalhealth.org**
Kris Lemon: **Kris@reachdigitalhealth.org**
Desi Nair: **Desi@reachdigitalhealth.org**
Keith Mangam: **Keith@reachdigitalhealth.org**
Sr. Jane Sebidi: **jane.sebidi@health.gov.za**

Acknowledgements
We would like to especially thank the National Department Of Health and Sister Jane Sebidi for their facilitation of the MomConnect platform for the last 10 years and for her oversight of the study presented in this report. We would also like to acknowledge the contribution of Meta, who funded this study.

About IDinsight
IDinsight uses data and evidence to help leaders combat poverty worldwide. Our collaborations deploy a large analytical toolkit to help clients design better policies, rigorously test what works, and use evidence to implement effectively at scale. We place special emphasis on using the right tool for the right question, and tailor our rigorous methods to the real-world constraints of decision-makers. IDinsight works with governments, foundations, NGOs, multilaterals and businesses across Africa and Asia. We work in all major sectors including health, education, agriculture, governance, digital ID, financial access, and sanitation.

We have offices in Dakar, Lusaka, Manila, Nairobi, New Delhi, Rabat, and Remote. Visit **www.IDinsight.org** and follow on Twitter **@IDinsight** to learn more.

About Reach Digital Health
Reach Digital Health uses digital tools to help users access tailored healthcare information and services. Using mobile technologies like SMS (Short Message Service) and WhatsApp we are able to reach large populations in cost-effective ways, including citizens and health workers.

Based in South Africa, our solutions have reached millions of people in over 200 low and middle-income countries. We work closely with implementation and research partners, as well as national governments and global public health organisations, such as the World Health Organization.

Learn more at **www.reachdigitalhealth.org** or follow us on social media **@ReachHealth** on X, **@ReachDigitalHealth** on LinkedIn and Instagram.

Table of Contents

Acronyms

ANC	-	Antenatal Care
ARM	-	Appointment Remind Model
BCM	-	Browsable Content Model
CONSORT	-	Consolidated Standards of Reporting Trials
DHS	-	Demgraphic Health Survey
DTP	-	Diphtheria, Tetanus, and Pertussis
FAQs	-	Frequently Asked Questions
FDR	-	False Discovery Rate
ICT	-	Information and Communication Technologies
ITT	-	Intent To Treat
NDOH	-	National Department of Health
pp.	-	Percentage Points
RCM	-	Relative Content Model
RCT	-	Randomised Controlled Trial
SD	-	Standard Deviation
SMS	-	Short Messaging Service
TOT	-	Treatment-On-the-Treated
USD	-	United States Dollar
USSD	-	Unstructured Supplementary Service Data
WA	-	WhatsApp
WHO	-	World Health Organisation
ZAR	-	Zuid-Afrikaanse Rand (South African Rand)

Executive Summary

This report presents details and results of a randomised controlled trial (RCT) of different messaging models implemented on the MomConnect platform in South Africa. The primary objective of the study is to determine the impact of different messaging models on key health knowledge, behaviour, and healthcare utilisation outcomes in order to inform a decision about which model to utilise across MomConnect as the standard messaging model. The key research question that we address is what impact the different messaging models have on attendance of antenatal care (ANC) visits, infant immunisation (through six weeks of age), knowledge of maternal and infant health behaviours, and adoption of maternal and infant health behaviours compared with the standard messaging model.[1]

Mobile health (mHealth) interventions are increasingly recognised for their potential to improve health outcomes, particularly in settings with limited healthcare infrastructure. MomConnect is an mHealth initiative supported by the South African National Department of Health (NDOH). MomConnect's Theory of Change posits that if pregnant women and mothers have access to high-quality, relevant content related to maternal and infant health, they will be equipped with knowledge about healthy pregnancy and postpartum behaviours, which would translate to practising these behaviours, and thus, contribute to improvements in maternal and infant health outcomes.

Through an RCT, we assess the impact of one status quo delivery model compared to three alternative delivery models on health knowledge, behaviours, and healthcare utilisation outcomes. Our control group is the Appointment Reminder Model (ARM; the standard MomConnect messaging model), and our three treatment groups are:

1) **Relevant Content Model on WhatsApp (RCM-WA)**
2) **Relevant Content Model on SMS (RCM-SMS)**
3) **Relevant Content Model + Browsable Content Model (RCM+BCM)**

The RCM model provides maternal health information and frequently asked questions (FAQs) relevant to the mother's stage of pregnancy or postpartum; the RCM-SMS model covers the same health topics as RCM-WA but is slightly condensed due to lower character limitations of Short Message Service (SMS) messages and the associated costs with that; and the RCM+BCM model has the same maternal information as the RCM model, but provides an option to browse a menu of additional maternal health information in addition to the stage-relevant FAQs.

We split the health-related outcomes into primary and secondary outcomes. **Our primary outcomes are close to final health outcomes: 1) number of ANC visits, measured by a binary indicator of whether participants attended at least eight ANC visits; and 2) infant immunisation status, measured by a binary indicator of whether the baby received all six recommended vaccinations at six weeks old (two at birth and four at six weeks old).** Our secondary outcomes

[1] We did not have any participants who were not exposed to MomConnect at all, so we are not able to make any conclusions about the impact of the MomConnect platform, itself.

are more intermediate: 1) an index of maternal and infant health behaviours knowledge (knowledge index), 2) an index of adoption of healthy maternal and infant health behaviours (behaviour index), and 3) a number of outcomes measuring participants' user experience with MomConnect.

Our analysis estimates the intent-to-treat effects of the different messaging models on the primary and secondary outcomes listed above. Specifically, we conduct hypothesis tests of the equivalence of outcomes between the status quo ARM model and the RCM-WA and RCM-BCM models, between the RCM-WA model and the RCM-BCM model, as well as between the RCM-WA model and the RCM-SMS model (to isolate the impact of delivering messages via SMS instead of WhatsApp which is highly relevant for lower-resource settings with low smartphone ownership rates).

Eligibility for the study was determined based on the age of potential study participants, their gestational age at enrolment, whether participants had access to a mobile device that can download and use WhatsApp, their willingness to enrol in the platform, and whether participants were first-time users of MomConnect. Enrolment had to be staggered based on gestational age to permit a data collection period that was manageable given resource constraints and decision-making timelines about scale-up. This resulted in the oversampling of participants with a later gestational age at enrolment, participants who are slightly older than the population of all MomConnect users, and participants of specific South African provinces (especially Gauteng and Mpumalanga). We conduct robustness tests of our results applying post-stratification weights to strengthen the external validity with respect to the overall population of MomConnect users.

We conducted three rounds of primary data collection - baseline and midline surveys conducted via WhatsApp/SMS (depending on treatment group), and a longer endline survey conducted over the phone. Additionally, Reach collected user engagement data through MomConnect.

We conduct balance checks on a number of demographic characteristics and find very small and largely insignificant differences across treatment arms, which is consistent with the randomization procedure having been implemented as intended[2] and our estimates are internally valid.

Our results show very few meaningful differences in primary outcomes. For ANC visits, we find no significant effects between ARM and RCM-WA or RCM-BCM, nor between RCM-WA and RCM-SMS and can rule out effect sizes of more than +/- 5 percentage points with 95% confidence. For immunisations, we find that the RCM-BCM model leads to a 3 percentage point decrease in full immunisation coverage (p-value of 0.018) compared with the ARM model's value of 91%, significant at 5% significance level. However, when using a multiple hypothesis correction, this finding is no longer significant at 5% or 10% levels (sharpened q-value of 0.16). Since we are unsure whether RCM-BCM really caused a reduction in immunisation coverage, but there is evidence that it may have, we believe the cautionary approach is to advise against making the RCM-BCM model the new default model of MomConnect.

[2] Note that randomization was done within the MomConnect platform and was not implemented by IDinsight, therefore ensuring randomization appeared to be successful was an important task

We find some moderate differences across messaging models for our secondary outcomes. For the knowledge index, we find no significant effects between ARM and RCM-WA or RCM-BCM at conventional significance levels, but we do find an increase in knowledge for RCM-WA compared with RCM-SMS of ~0.13 standard deviations (p-value of 0.039), significant at the 5% level. For the behaviour index, we find that both RCM-WA and RCM-BCM have a positive and significant (at the 10% level) impact of ~0.1 standard deviations (p-values of 0.100 and 0.073) compared to the ARM model. For user experience, we find that RCM-WA and RCM-BCM lead to an increase in the likelihood of finding the baby information component of MomConnect most useful (mainly instead of finding appointment reminders most useful) of 4.2 and 6.3 percentage points (p-values of 0.106 and 0.021), respectively. For user engagement, we see that RCM-BCM send about 6 messages, on average, to MomConnect compared with ARM (p-value of 0.017) and that the number of messages sent by RCM-SMS is substantially less than all other groups, including 67.2 fewer messages compared with RCM-WA (p value of 0.000).

We identified three main limitations with our study linked to the response rate, recruitment strategy, and self-reporting. First, while our response rate of 63.7% is higher than we expected for a phone survey, it poses a potential threat to extrapolating the results from the specific study sample to the population of all MomConnect users. Second, the recruitment strategy of the study led to a study sample that had a substantially higher gestational age at the time of registration, slightly different average age, and some differences in distribution of province compared to the broader population of MomConnect users, although this was mitigated by the use of post-stratification weights. Third, all of our outcomes rely on self-reported survey data, which can lead to random measurement error, potentially leading to attenuation bias, or an underestimation of treatment effects.

Overall, our recommendation is to scale up the RCM-WA model as the new default model for MomConnect due to its significant positive impact on certain secondary outcomes without any negative impact on primary outcomes. However, there are two main reasons that we consider this a “soft” recommendation: 1) the lack of detectable effects on primary outcomes and the relatively moderate effect sizes of secondary outcomes; and 2) that the primary result that contributes to the choice of RCM-WA over RCM-BCM is not significant when conducting multiple hypothesis corrections. Our secondary recommendations are: 1) Adapt the MomConnect SMS Core messaging model to include some elements of the RCM model; 2) Consider conducting further qualitative investigations to understand what may be driving some of the results, or lack of results, presented; 3) Consider conducting an impact evaluation of the MomConnect platform as a whole, as it is rolled out to a new country where non-users of MomConnect can provide a valid counterfactual; and 4) Consider strengthening health facility adherence to registering women on MomConnect during ANC visits in health facilities outside of Gauteng.

1. Introduction

This study examines different potential messaging models of the MomConnect program, an mHealth initiative in South Africa owned by the South African National Department of Health (NDOH). The primary objective is to determine which of the messaging models improves health knowledge, behaviours, and healthcare utilisation outcomes the most compared to the status quo model, and therefore should be considered for expansion across MomConnect as the standard messaging model. This is relevant for Reach Digital Health and NDOH in their decision-making about MomConnect, as well as providing some learnings for implementation of mobile health (mHealth) interventions.

This report first presents a literature review and background to the project in section 1. Section 2 is a detailed overview of the study design which includes a description of the different models tested, eligibility criteria, randomisation strategy, data collection strategy, and ethical considerations. Section 3 presents the empirical approach, including specification of regression, primary and secondary outcomes, external validity and multiple hypothesis adjustments. Section 4 is the descriptive statistics and balance checks, including a large table with all relevant information. Section 5 presents the results for primary and secondary outcomes, as well as attrition. Finally, Section 6 is a discussion of the results, limitations of the study, and our core recommendations.

1.1. Literature Review & Background to the Project

The case for mHealth has been gaining traction as a source to improve health outcomes of individuals. mHealth refers to using 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 healthcare services directly to citizens, with the greatest potential gains in under-resourced health ecosystems. Existing literature reviews identify a wide range of settings where maternal health messaging is effective (Lee et al. 2015; Poorman et al. 2014). A recent small-scale RCT of 177 mothers in one South African city found that enrollment 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 nine years and supporting 4.8 million mothers through their pregnancies and early childhood care, MomConnect represents one of the world's largest maternal health messaging platforms (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 successes, there is still scope for improving MomConnect's impact on mothers and their children (Mehl et al. 2018). The 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 NDOH, it avoids women needing to navigate complicated, contradictory, or unverified information available on the internet.

As a primarily 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 the status quo ARM messaging model, MomConnect can only start a conversation with mothers once a week by alerting them of their upcoming 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. 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 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.

The most recent data that we could find to provide comparison for our primary outcomes is from the 2016 Demographic and Health Survey (DHS).[3] The DHS estimates that 76% of pregnant mothers made at least 4 ANC visits (NDOH & ICF, 2019). The DHS does not report for full immunisation coverage at six weeks (as we will in this paper) but we can use the DTP-containing vaccine, administered at six weeks and for which there is data both in the DHS and in our endline survey, as a comparison point. The DHS reports 88% of the infant population received the DTP-containing vaccine in 2016 (NDOH & ICF, 2019). Note that there could be more recent data available from DHIS2 data, although we were unable to access this data at the time of writing this report.

This study was initially commissioned to take place between May 2022 - March 2023. However, on November 8th 2022, the original study was stopped for reasons related to implementation fidelity concerns. After discussions between IDinsight, Reach, and Meta (who funded the study) in early 2023, a decision was made to relaunch the study in May 2023, with endline data collection taking place from October 2023 - February 2024.

[3] We are working with the NDOH to get access to administrative health data from DHIS2

1.2. Stakeholders

Four key organisations are involved in this research study: IDinsight, Reach Digital Health, Meta, and the South African NDOH. IDinsight is the research partner who leads the research design, endline data collection, data analysis, report writing and drafting of recommendations based on study results. Reach Digital Health are the implementer of MomConnect; other than MomConnect implementation, they collect and share program data, facilitate baseline and midline WhatsApp surveys, support endline data collection, and lead the key decision-making process. MomConnect is an NDOH initiative integrated into the public health systems of South Africa; therefore, NDOH is a key decision maker in the future implementation of MomConnect. Meta is the funder for the project, who shares research findings, and uses research results for future WhatsApp projects.

2. Study Design

2.1 MomConnect Background

MomConnect is an mHealth initiative supported by the South African NDOH that was launched in 2014. With 4.8 million mothers reached as of early 2024, this platform has demonstrated significant initial successes in maternal and infant health communication, as mentioned in section 1.1. The most recent standard MomConnect model involves sending appointment reminders to pregnant women and mothers via WhatsApp with a prompt provided to see some additional maternal health information by a simple message response, or bi-weekly Short Message Service (SMS) messages based on her stage but with no engagement from the user.

According to guidelines by NDOH, (specifically the "Road map for the provision of a maternal and child health package of care for the first 1000 days"), pregnant women in South Africa are supposed to be informed about MomConnect and encouraged to join the platform during their first ANC visit at a public health facility, and each subsequent visit.[4] This is typically done by clinic staff as pregnant women attend their first ANC visit at a public health facility. While compliance with the guideline is imperfect, MomConnect had 243,952 new registrations in 2022, roughly one quarter of the 998,362 registered births in South Africa in 2022.[5] In 2023, MomConnect reported a similar number of registrations (224,745).

2.2 Intervention Background

The four experimental groups being assessed in this study are:

1. **Control: Appointment Reminder Model (ARM)** - delivered through WhatsApp, this model sends mothers conversation starter messages reminding them about their upcoming clinic appointments (1 week before and another 2 days before scheduled ANC and PNC meetings) and then **provides more comprehensive and relevant maternal health**

[4] https://www.health.gov.za/wp-content/uploads/2023/05/Roadmap-SA-Maternal-Child-26-August-2019.pdf
[5] Government data on the number of births in 2023 is not available, hence the use of 2022 data, which can be found at https://www.statssa.gov.za/?p=16902.

information after mothers respond "YES" to the appointment reminder, which ends asking them if they'd like to learn more.

2. **Treatment 1: Relevant Content Model (RCM-WA)** - delivered through WhatsApp, this model sends mothers weekly conversation starter messages reminding them about their upcoming appointments **along with some maternal health information relevant to their pregnancy/postpartum stage.** Additional frequently asked questions (FAQs) are also provided relevant to their stage, accessed by responding with a number, so they can engage further with other topics.

3. **Treatment 2: Relevant Content Model SMS (RCM-SMS)** - delivered through SMS, this model sends mothers weekly conversation starter messages, which carry **both clinic appointment reminders as well as maternal health information relevant to their pregnancy/postpartum stage** but are shorter due to limitations on SMS length. These SMS messages are approximately 480-640 characters, whereas WhatsApp messages are limited to 1,024 characters. Additional FAQs are also provided and can be accessed via dialling an Unstructured Supplementary Service Data (USSD) code.

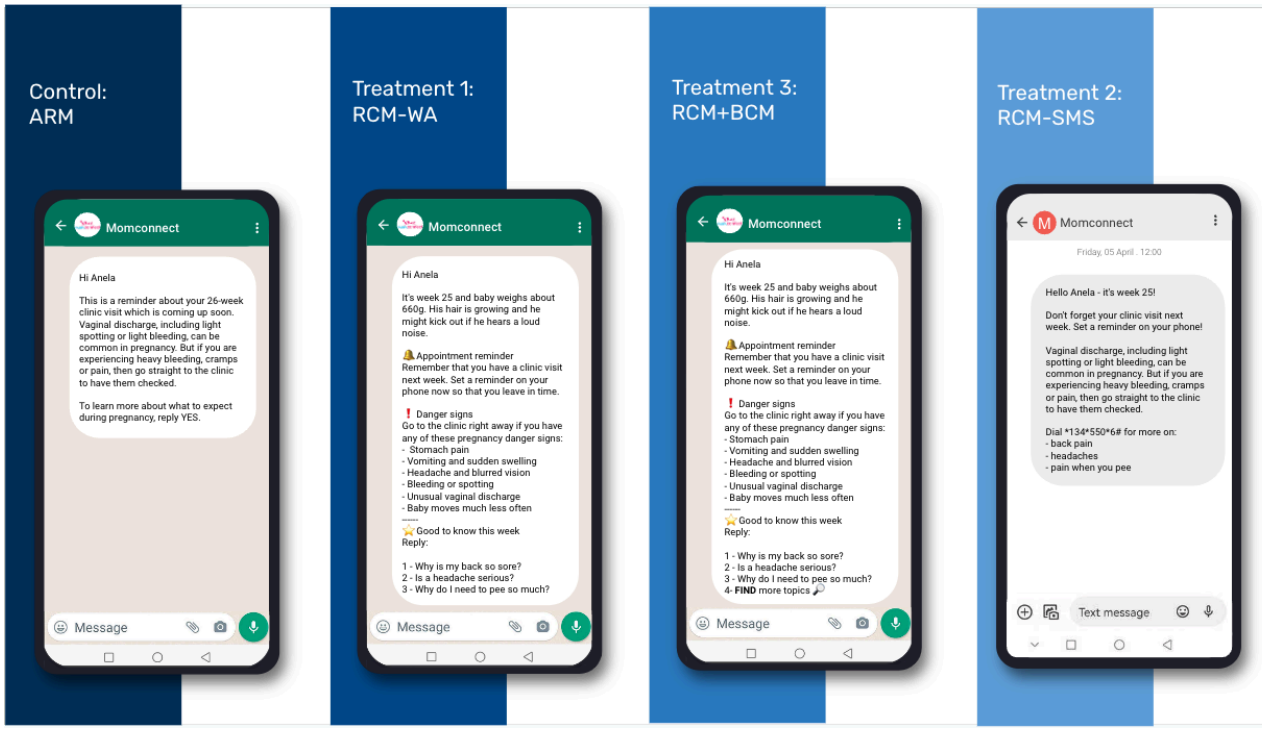
4. **Treatment 3: Relevant Content Model + Browsable Content Model (RCM+BCM)** - delivered through WhatsApp, **in addition to the RCM details described above, the RCM+BCM model includes an option to browse a larger menu of maternal health information topics** with the push message containing relevant content. This means that participants receive the FAQs relevant to their stage of pregnancy, but can also browse all other FAQs in the system.

The study compares the three new messaging models to the status quo MomConnect messaging model as well as to each other. All treatment models differ from the ARM control model as they 1) proactively provide maternal and infant health information with the initial appointment reminder and not only once a response is received; and 2) send messages bi-weekly rather than just when appointments are scheduled. The RCM model provides maternal health information and FAQs relevant to their stage; the RCM-SMS model covers the same health topics as RCM-WA but is slightly condensed due to lower character limitations of SMS messages compared to WA messages; and the RCM+BCM model has the same maternal information as the RCM model, but provides an option to browse a menu of additional maternal health information in addition to the FAQs. See **Figure 1** for an example of the messages sent for different treatment arms.

The study evaluates the relative effectiveness of the three new health messaging models on ANC visits and infant immunisation of vaccinations recommended at birth and six weeks after birth (data collected when baby is 7+ weeks old).[6]

[6] These are the primary outcomes. We also study a number of secondary outcomes. See section 3.2 for details on the complete set of outcomes included in this study.

Figure 1: Examples of the Four Different Messaging Models



2.3 Eligibility Criteria

Once a new user registered for MomConnect during our study enrollment period from 31 May to 8 October 2023, we checked whether she satisfied the eligibility criteria of the study. To be eligible, a woman had to be 18 years or older, be a first time user of MomConnect, have a phone that allows her to install the WhatsApp application on which the standard MomConnect model sends messages, and be within a gestational age range of 16-30 weeks, depending on the specific week of the enrollment period. We introduced the eligibility restriction on gestational age at registration to 1) ensure that we had at least 9-10 weeks of exposure to the intervention pre-birth; and 2) to ensure that all study participants will be at least 7 weeks postpartum within our endline data collection period, which is why the enrolment period is staggered as shown in **Figure 2**. The figure also presents when someone enrolled at a specific gestational age and in a specific enrollment week (shown on the left side of the red line) was expected to be contacted for the endline survey (shown on the right side of the red line).

Figure 2: Gestational Age Eligibility by Week of Study Enrollment

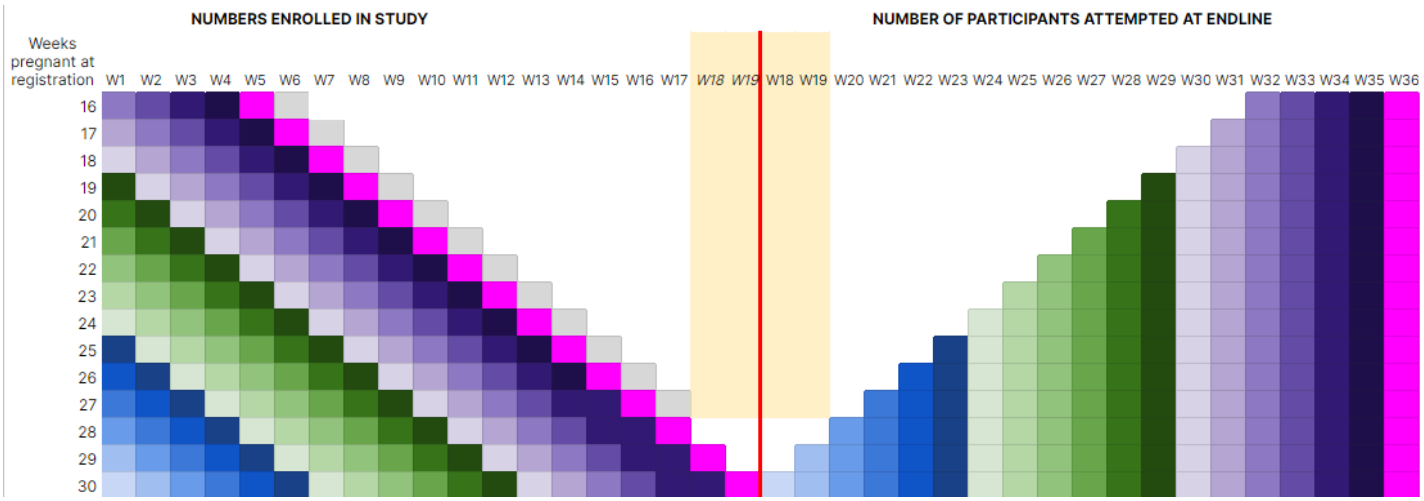


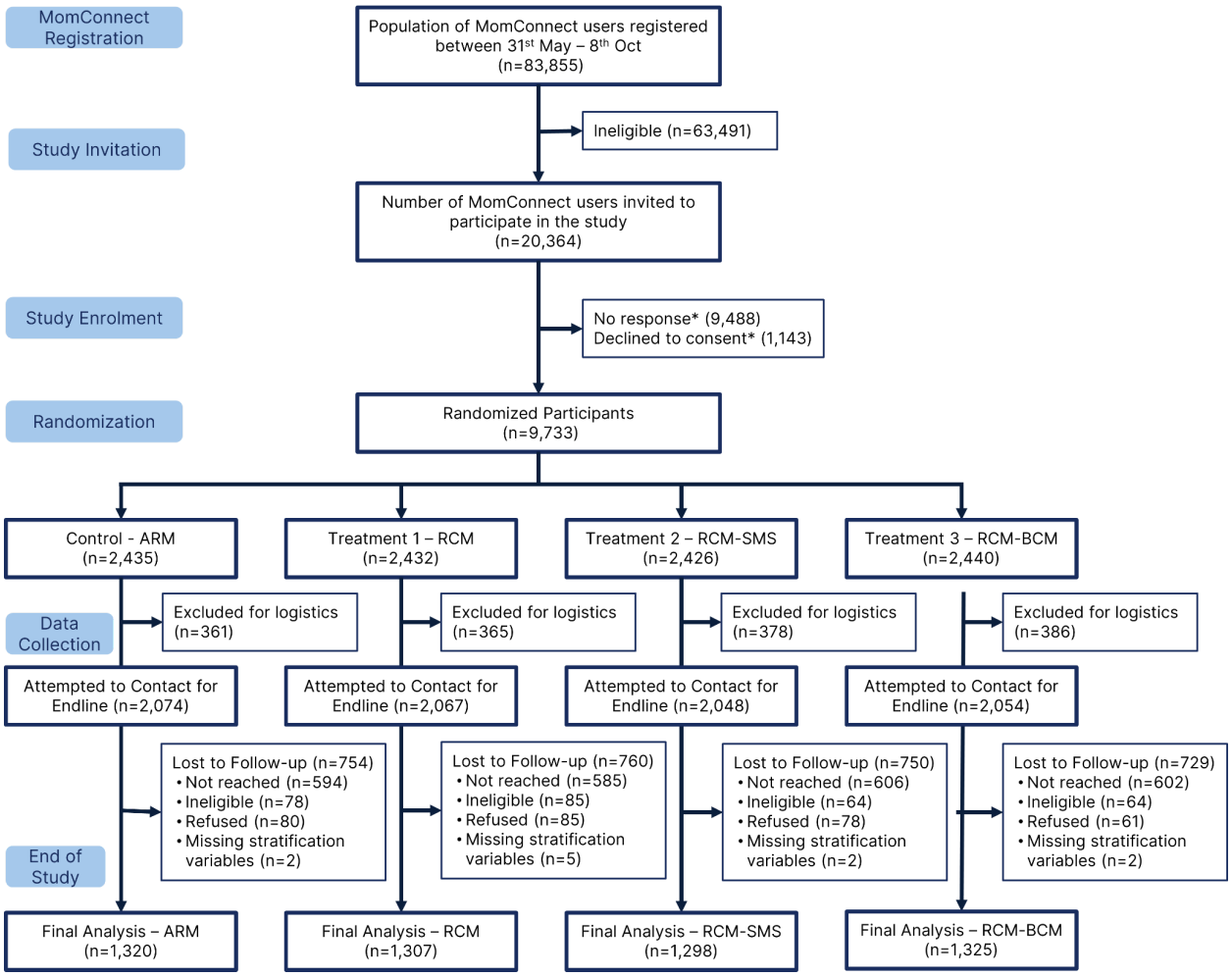
Figure 2 shows weeks of the data collection period on the horizontal axis and the gestational age of mothers at registration on the vertical axis. Weeks 1-18 on the horizontal axis are the enrollment period and weeks 18-36 are the data collection period. Note that weeks 18 and 19 are repeated on the horizontal axis in both the enrollment period (left of the red line of Figure 2) and the data collection period (right of the red line in Figure 2) due to the overlap between the end of enrollment and the start of data collection. The rows represent the gestational age of the participants at enrollment. The colours indicate how the week a mother registers combines with their gestational age to result in the week that they are surveyed during the endline data collection period.

We employed a staggered enrollment strategy to ensure the data collection timeline did not extend beyond the target date of February 9, 2024. Eligibility for enrollment was therefore determined by the participant's gestational week, creating a controlled and tiered enrollment cutoff system, as seen in **Figure 2**. For example, participants who were 16 weeks pregnant (and therefore were only expected to be eligible for the endline interview 31 weeks later) were only enrolled during the first six weeks of the enrollment period, whereas mothers who were 30 weeks pregnant (and therefore were expected to be eligible for the endline interview 17 weeks later) were enrolled for the entire enrollment period. Since this enrollment protocol created an overrepresentation of mothers enrolling at a later gestational age in our sample, we conduct robustness tests of our results applying post-stratification weights in our analysis in section 5. Note that this diagram is just illustrative and information on actual study enrollment can be found in Section 2.4.

Due to a higher than expected response rate, we were able to exclude the pink and darkest purple line in **Figure 2** from the endline data collection in order to complete our data collection on time whilst still meeting our target sample. This is also represented as “Excluded for logistics” in **Figure 3** below.

2.4 Study Enrollment and Randomisation

Figure 3: CONSORT diagram for the study



*Note that the numbers in this box are based on percentages of no response and declined to consent and the difference between the box above and below as the totals. This is because the tracking dashboards from which this data came updated on a continuous basis, and MomConnect users can request to leave the platform and have their data deleted. This means that at the time of writing the report, the numbers in the MomConenct database do not match the numbers at the point of enrolment and in the endline dataset.

The CONSORT diagram in Figure 3 shows the progress of study participants through the different study stages. Out of 83,855 women registering for MomConnect during our study enrollment period from 31 May to 8 October 2023, 20,364 (24%) were eligible according to the criteria of the study - 18 years or older, a first time user of MomConnect, have a phone that allows her to install the WhatsApp application, and be within the eligible gestational age range given the specific enrollment week. Out of these 20,364 eligible women, 9,733 (48%) consented to be part of the

study.[7] We randomly assigned the 9,733 consenting study participants either to MomConnect's status quo Appointment Reminder Model or one of three new messaging models, stratified by age, province, and gestational age at the time of registration.[8] Of those that were allocated to a specific messaging model, between 361-385 were excluded due to data collection logistics. Specifically, our response rate was higher than expected so we excluded the final two weeks of data collection participants without biasing our sample, once we hit the target of at least 1,300 per treatment arm. [9] The majority of participants lost to follow up were due to contacting them seven times without response, following the protocol outlined in Appendix 3, as well as some being ineligible[10] or refusing to participate in the survey.

2.5 Data Collection

We conducted three rounds of primary data collection for this study and also analysed back-end engagement data.

Baseline survey: Immediately after consenting to the study and being randomly assigned to one of the four study arms, we invited study participants to a baseline survey, incentivised with a USD $0.26[11] airtime token. The baseline survey was sent via WhatsApp from the MomConnect account for all treatment arms and contained 13 questions, primarily covering healthy behaviours during pregnancy and intentions after birth, as well as some demographic questions. The full survey instrument can be found in Appendix 1A.

Midline survey: Once study participants reached the 35th week of their pregnancy, we invited them to participate in a midline survey, also incentivised with a USD $0.26 airtime token. We sent the survey through WhatsApp from the MomConnect account for participants in the ARM, RCM-WA, and RCM-BCM arms and through USSD for participants in the RCM-SMS arm. The midline survey contained 16 questions on healthy behaviours and practices during pregnancy and intentions after birth, including nutrition behaviour, knowledge of potential danger signs, breastfeeding intentions, and opinions on infant vaccinations. The full survey instrument can be found in Appendix 1B. We do not report treatment effects estimates for any midline outcomes since we find that there is differential attrition in midline survey completion across treatment arms as shown in Table 13 in section 5.3.

Endline survey: Once study participants reached their seventh week postpartum (based on the estimated date of delivery which was provided in the MomConnect platform upon enrollment), we called them on their phone number registered with MomConnect and asked them to participate in

[7] The whole study registration and consent flow took place on WhatsApp for all study participants, including those who were later assigned to the RCM-SMS arm. The actual content messages were subsequently delivered through SMS for the RCM-SMS arm and through WhatsApp for the other three arms.
[8] For age, we grouped participants by whether they are below or above 30 years old. For province, we grouped participants by the nine provinces in South Africa. For stage of pregnancy, we grouped participants by whether their gestational age at study registration was between weeks 16-20, weeks 21-25, or weeks 26-30. Participants were then stratified based on these groupings, which yielded a total of 54 unique strata.
[9] Note that the RCM-SMS only has 1,298 in the analysis due to missing stratification variables for 2 respondents
[10] Reasons for ineligibility included the baby passing away, the owner of the number dialed reported having no child, or a baby that is too old that they were born before the study period.
[11] The amount was ZAR 5 which equals USD 0.26 according to April 15th 2024 exchange rate

our endline survey, offering a participation incentive of USD $1.59[12] worth of airtime credit.[13] Our enumerators made up to seven attempts when trying to reach each study participant. The endline survey asks participants about ANC visits, their baby's immunisations at birth and six weeks after birth, health knowledge, health behaviour, and their user experience with MomConnect. All respondents received questions on immunisations, ANC visits and demographics, whereas only a subset of participants received the questions on health knowledge, behaviour and user experience (n=2,200). If the baby was under 7 weeks old, then an appointment was scheduled with the mother for when the baby was old enough. The full survey instrument can be found in Appendix 1C. Enumerators were randomly assigned to respondents stratified by treatment status of respondents, in order to avoid enumerator-effects influencing results and to enable us to check the robustness of our results to the inclusion of enumerator fixed effects in our analysis. Endline data collection was facilitated by IDinsight whereas the baseline and midline, which was still designed by IDinsight, were directly from the MomConnect account.

Engagement data: The MomConnect platform collects data throughout the study period for each participant on their engagement with the platform. This includes: the number of messages received from MomConnect, the number of messages sent to MomConnect, the number of days on which the participant messaged MomConnect, and the number of days on which MomConnect messaged the participant.

2.6 Ethics & IRB

The study was registered, including a pre-analysis plan, with ClinicalTrials.gov under the title "Testing Content Deliver Models for MomConnect (MQR)" with ID number: NCT05505760.[14] All study details were included in the pre-analysis plan and uploaded before the endline data collection started and before we conducted any analysis of outcome variables.

Meta funded this study, and is also the owner of the WhatsApp platform. They contributed to the research questions being assessed, and we kept them updated on the design of and progress of the study, however, they did not have any active role in the study design, data collection, or data analysis.

The study was given ethical approval by PharmaEthics in South Africa (ref no: 220224545). This included a few minor amendments and an annual re-approval following the relaunch of the study.

Informed consent was provided at two levels:

1. **Participation in the study** - Since there are four different messaging models randomly assigned to study participants, participants were informed that participation in this study could mean receiving messages different from the standard MomConnect messages (i.e., the status quo ARM model). Participants were also informed that participation in the study

[12] The amount was ZAR 30 which equals USD 1.59 according to April 15th 2024 exchange rate
[13] Our endline data collection period started on 16 October, 2023, and lasted until 9 February, 2024.
[14] https://clinicaltrials.gov/study/NCT05505760

means they will be asked to complete surveys about their health knowledge and behaviours.

2. **Participation in outcome surveys** - For the baseline and midline surveys, study participants were surveyed either through WhatsApp or USSD and a consent message to participate in the survey was sent before each round through the MomConnect platform (WhatsApp or SMS). For the endline survey, study participants were surveyed over the phone and verbal consent was asked at the start of the survey, after enumerators read a comprehensive consent statement to study participants (this can be found in Appendix 1C as part of the endline survey).

3. Empirical Approach

3.1 Regression Specification

Our analysis estimates the intent-to-treat (ITT) effects of the different MomConnect messaging models on the primary and secondary outcomes described in Section 3.2. Since we randomly assign study participants to messaging models, we can interpret any differences in outcomes across messaging models as causal. Our analysis focuses on the ITT effect because i) treatment compliance is unobservable (i.e. we do not know whether participants read messages) and ii) the ITT is more policy-relevant than the treatment-on-the-treated (TOT) effect.[15] ITT analysis is more policy-relevant as it includes every subject who is randomised and who receives the intervention, ignoring noncompliance, protocol deviations, withdrawal, and anything that happens after randomisation, similar to what would happen if the intervention was scaled up (Gupta, 2011).

We estimate the following regression specification in order to understand the impact of the different messaging models on our outcomes:

Equation 1

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

Where Y_{is} is the outcome of individual *i* in stratum *s*, T_{ijs} is the treatment indicator for individual *i* and treatment arm *j* in stratum *s*, δ_s is a vector of strata dummies, and ϵ_{is} is the individual-level error term. We will estimate heteroskedasticity-robust standard errors using the Huber-White sandwich estimator. The estimated treatment coefficients β_j capture the causal effect of the respective treatment on the outcome Y_{is}.

We will also estimate a version of equation 1 controlling for demographic characteristics of individuals and enumerator fixed effects.

[15] The key policy question is which of the different messaging models to adopt based on how they affect key health knowledge, behaviours, and healthcare utilisation outcomes.

3.2. Outcomes

Primary Outcomes

The two primary outcomes of the study estimate ANC visit attendance of pregnant women and immunisation status of infants six weeks after birth. Our primary outcome for ANC visits is a binary variable measuring whether a woman attended at least eight ANC visits during her pregnancy as currently recommended by the World Health Organisation (WHO) and adopted by NDOH in 2017 (Hlongwane, Tsakane Mag et al, 2017).[16] Our primary outcome for infant immunisation status is a binary variable measuring whether the infant received the six vaccinations, two given at birth and four at six weeks post-birth prescribed by the NDOH in South Africa, with data collected on the immunisation status at the time of interview (mean was at a baby age of nine weeks).[17,18] We do not report results for vaccination intentions, which we measured at midline, since we find that there is differential attrition in midline survey completion across arms (see Table 13 in section 5.3). For ANC visit attendance, we also look at a binary variable measuring whether a woman attended at least four ANC visits during her pregnancy as previously recommended by the WHO and a continuous variable measuring the number of ANC visits a woman attended during her pregnancy. For infant immunisation status, we also look at a continuous variable measuring the number of vaccinations that are recommended at birth and at six weeks after birth, that the infant received.

Generally speaking, the primary outcome variables may suffer from measurement error (which would lead to a downward bias of our estimates coefficients due to attenuation bias) since they are self-reported and respondents may either not recall the correct information or intentionally mis-report information (for example, due to experimenter demand bias). Therefore, we ask mothers whether they have their pregnancy card (contains details of ANC visits) and their immunisation card (contains details of vaccinations their baby received) with them during the survey, and if so, to fetch it so that they can read off the information rather than relying on recall. Very few mothers had the pregnancy card (1.7%), however the majority of mothers did have their immunisation card (70.6%), allowing for some mitigation of this measurement error.

Secondary Outcomes

Our secondary outcomes include a knowledge and a behaviour index based on a number of survey questions we asked at endline. The other secondary outcomes we look at relate to respondents' user experience on the MomConnect platform, measured through endline survey questions, as well as to users' engagement with the platform, measured through backend data collected through the platform.[19]

[16] https://www.who.int/news/item/07-11-2016-new-guidelines-on-antenatal-care-for-a-positive-pregnancy-experience
[17] We pre-specified the construction of these two primary outcomes in our pre-analysis plan.
[18] The NDOH prescribes that babies in South Africa should receive the OPV0 and BCG vaccinations at birth and the OPV1, DTaP-IPV-Hib-HepB, RV, and PCV vaccinations six weeks after birth.
https://www.health.gov.za/wp-content/uploads/2020/11/epi-schedule.pdf
[19] We pre-specified the construction of the knowledge and behaviour indices as well as the user experience questions and the engagement data measures. Details from the pre-analysis plan can be found at the clinical trial registration page - https://clinicaltrials.gov/study/NCT05505760

For the knowledge and behaviour outcomes, we construct two summative indices with equal weights given to the different topic areas of interest. The two indices measure respondents' knowledge of healthy behaviours (knowledge index) and adoption of healthy behaviours (behaviour index), and include the following topic areas: knowledge about baby danger signs, breastfeeding, and anaemia for the knowledge index; and self-reported behaviour related to nutrition, pregnancy preparedness, breastfeeding practice, and anaemia prevention for the behaviour index. The topic areas, questions within topic areas, and index weights were determined together with Reach based on the content of MomConnect with a focus on areas where content varied between messaging models. The construction of the indices was pre-specified in the pre-analysis plan.

We measure knowledge and behaviour through composite indices for two reasons: first, MomConnect content covers a variety of topic areas, and other than immunisations and ANC visits, no single topic was identified as more important or decision-relevant than others; second, compared to the alternative of analysing each topic area separately, collapsing discrete topic areas into two indices reduces the risk of false positives resulting from testing a large number of hypotheses. Full details on the construction of the indices can be found in Appendix 4.

To capture the user experience of MomConnect, we also analyse the following outcomes, which are based on a mixture of MomConnect user engagement data and user-experience data captured in primary endline surveys:

Outcome	Definition
Most useful component	Indicator for whether mother thinks that **appointment reminders** are the most useful part of MomConnect Indicator for whether mother thinks that **pregnancy information** is the most useful part of MomConnect Indicator for whether mother thinks that **baby information** is the most useful part of MomConnect Indicator for whether mother thinks that **helpdesk for questions** is the most useful part of MomConnect Indicator for whether mother thinks that **something else** is the most useful part of MomConnect
Internet use	Indicator for whether a mother uses internet during her pregnancy (self-reported)
Trustworthy score	Indicator for whether a mother rates MomConnect as a very or completely trustworthy source for information on pregnancy, childbirth, and newborn care
Most helpful component	Indicator for whether a mothers selected **appointment reminder** as the most helpful part of the message Indicator for whether a mothers selected **baby growth information** as the most helpful part of the message

	Indicator for whether a mothers selected **baby health information** as the most helpful part of the message
Perceived support from MomConnect	Indicator for whether a mother answered "Very supported" or "somewhat supported" regarding the helpfulness of MomConnect messages
Platform preference	Indicator for whether a mother prefers to receive message on WhatsApp or SMS
Ease of finding information	Binary indicator (formed from Likert Scale question) for whether a mother finds it very easy to find information on MomConnect
Recommend MomConnect	Binary indicator (formed from Likert Scale question) for whether a mother is very likely to recommend MomConnect to other mothers
Read all MomConnect messages	Binary Indicator for whether a mother reports that she read all MomConnect messages in full
Frequency of messages	Binary Indicator for whether a mother felt they did not receive messages frequently enough

3.3. External Validity and Gestational Age

As section 4 shows, our sample differs somewhat in terms of their gestational age at registration, age, and province from the overall population of MomConnect users in 2023. As a robustness check of our main analysis, we apply post-stratification weights for all observations in our analysis sample by weighting observations by the inverse of the ratio of the number of users in each stratum in our sample and the number of users in the same stratum in the population of MomConnect users who are above 18.[20] Most notably, our enrollment schedule discussed in Section 2.2 has led to an overrepresentation of pregnant women who register for MomConnect at a later gestational age in our sample compared to the overall population of women on MomConnect. Women who enrol at a later gestational age on the MomConnect platform may be different from those who enrol earlier in pregnancy on unobserved characteristics correlated with treatment effectiveness.[21] Further, women registering later are exposed to MomConnect for a shorter time and therefore have less time to be affected by the intervention. While lower in magnitude, there are also statistically significant differences between the average age and the province distribution of users in our sample and the 2023 population of MomConnect users. Therefore, post-stratification weights likely improve the external validity of the estimates when extrapolating results beyond the specific sample to the population of all MomConnect users.

[20] We use data on the gestational age at registration, age, and province of MomConnect users who registered pre-birth provided by Reach for 2023.

[21] The majority of mothers who enrol on the MomConnect platform do so at their first ANC visit. Therefore, mothers who enrolled in the study at a later gestational age may have started ANC attendance later in their pregnancy, which could mean they are less likely to practise healthy behaviours during pregnancy, may have lower access to care, may have lower demand for information, etc. compared to those who start ANC earlier in pregnancy.

3.4. Multiple Hypothesis Adjustments

We designed this study to inform the decision around which messaging model should be scaled up for all MomConnect users. This decision will be informed by which treatment arm(s) improve the primary healthcare utilisation outcomes. We test the following eight 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: ANC visits and immunisations.

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

To correct for this, we adjust for testing multiple hypotheses by applying a false discovery rate adjustment (FDR) following Benjamini et al. (2006). In particular, we use the sharpened q-values discussed in Anderson (2008). This correction controls for the expected proportion of Type 1 errors. In our analysis, we report both adjusted and unadjusted p-values. These multiple hypothesis adjustments focus on the analysis of the four treatment arms and two primary outcomes specified, given that these hypotheses affect potential scale-up decisions. The analysis of secondary outcomes is more exploratory.

4. Descriptive Statistics & Balance Checks

4.1 Descriptive Statistics and Balance Checks

Table 1 shows descriptive statistics and balance tests for all participants registered for the study, all participants who responded to the endline survey, and all MomConnect users who registered before birth in 2023. For each of these three samples, the table includes age, gestational age at registration, and province - the three stratification variables. For the endline sample, we also include a number of demographic characteristics which we asked respondents in our endline survey. In order to ensure none of these demographic variables are affected by the treatments, we asked all respondents to respond to the demographic questions thinking back about the time they learnt about their most recent pregnancy, which was before they enrolled in this study.

We find that the sample of study participants who enrolled in the study (panel A) differs somewhat from the population of MomConnect users who registered before giving birth in 2023 (panel C). While the average age of all 2023 MomConnect users is 26.4, the average age of our study participants is 27.8. Due to the study enrollment structure described in section 2.3, the average gestational age at registration is substantially higher in our study sample (23.7 weeks) compared to all 2023 users (20.1 weeks). There are also some notable differences in the share of users from each province: users from Gauteng (45% in study sample, 40% in 2023 population) and Northern Cape (8.5% versus 1.4%) are overrepresented whereas users from Limpopo (7.4% versus 15.6%) and Mpumalanga (1.5% versus 8.5%) are underrepresented. The differences for all three variables are statistically significant between our sample and the 2023 user population at the 1% level.[22] Our post-stratification approach described in section 3.3 attempts to correct for these differences and to strengthen the external validity of the study with respect to the overall population of MomConnect users.

We do not find any substantial differences between the characteristics of study participants who enrolled in the study (panel A) and study participants who responded to our endline survey (panel B). The average age at study registration in the whole study sample is 27.8 years while for endline respondents it is 28.1 (the maximum difference within a treatment arm is 0.5 years). The average gestational age at registration is 23.7 weeks both in the whole study sample and for endline respondents (the maximum difference within a treatment arm is 0.1 weeks). Lastly, the maximum difference in the share of participants from a specific province between all participants and endline respondents is less than 1 percentage point. At least in terms of these three variables, **it does not seem to be the case that there is non-random attrition in our endline sample, although respondents and non-respondents may differ in terms of unobserved characteristics.** We have a more detailed discussion on differential attrition in section 5.3 and on the external validity of the study in section 6.

In addition to age, gestational age at registration, and province, we measure several other demographic characteristics of our endline sample: the average age of the baby at the time of the

[22] We conduct one-sample t-tests comparing the sample mean to the population mean for gestational age and age and a chi2 goodness of fit test between sample province proportions and population province proportions.

endline interview was 9.3 weeks which is within our target range of conducting interviews between 7-10 weeks after birth; 18% of our sample only have a primary education or less (grade 7 or less), 55% have a secondary education (finished grade 12), and 27% have a tertiary certificate or degree. At the time they learnt about their pregnancy, 43% of endline respondents were unemployed, and their average monthly household income was R5,000. 94% of our endline sample identify as Black African, with only small shares identifying as other ethnic groups. 18% of our sample report not having had a partner at the time they learnt about their pregnancy. Finally, 38% of our sample report that their most recent pregnancy has been their first pregnancy.

To assess the success of the randomisation procedure, we estimate the regression described in section 3.1, using the demographic characteristics shown in **Table 1** as outcomes. We report p-values of an F-test of joint significance of our treatment coefficients to assess the null hypothesis that the demographic characteristics are not equal across treatment arms. Out of 17 hypothesis tests, we only find one variable for which the p-value is below 0.05, namely the share of endline respondents who only obtained a primary education degree (p-value 0.044). The maximum difference between two treatment arms is 4 percentage points (15.7% in RCM-BCM versus 19.7% in ARM). Overall, we conclude that there are no substantial differences in observed characteristics across treatment arms, neither for the whole study sample nor for endline respondents. For most results that we report in section 5, we also conduct robustness checks controlling for the full set of demographic characteristics shown in panel B of **Table 1**.

Table 2 shows the proportion of births by province for our endline dataset, all MomConnect users in 2023, and all registered births in South Africa in 2022 (the latest available data on births). This comparison shows that MomConnect has a much higher proportion of registrations from Gauteng province, 45% in our sample and 40% on MomConnect in 2023, compared with the national proportion of births in Gauteng, 23% in 2022. This represents a 22 and 16 percentage point difference (our sample and all MomConnect registrations). This difference is mostly explained by a much lower representation of KwaZulu-Natal (10 and 12 pp.), Eastern Cape (6 and 6 pp.), and Mpumalanga (7 and 0 pp.). Prior to the study being launched, there was some expectation that this could be the case, driven by the fact that Gauteng is closer to the NDOH, and therefore health facilities in Gauteng may adhere more strongly to the official guidance of registering mothers on MomConnect. The differences in provinces may explain the composition of our endline sample, for example that women in our sample appear to be considerably more educated than the average South African. A potential policy recommendation arising from these findings would be for the NDOH to investigate the reasons for the discrepancies between birth shares and registration shares and potentially strengthen adherence to registering women for MomConnect during ANC visits in health facilities outside of Gauteng, particularly in KwaZulu-Natal and Eastern Cape provinces.

4.2 Descriptive Evidence on Changes from Baseline to Endline

Finally, we present some descriptive evidence on how MomConnect users compare between baseline and endline in terms of their self-reported diet as well as how their baseline intentions compare with behaviour or intentions at endline in terms of breastfeeding and immunisations.[23] We pool users from all four arms for this analysis and restrict it to the sample of users who responded to both the baseline and endline surveys. It is important to note that these results should not be interpreted as suggestive of the overall impact of MomConnect since we would expect behaviour and intentions to change over the course of pregnancy and beyond in absence of MomConnect, through information provided by nurses and doctors, for example. First, we find that users report slightly improved dietary behaviour at endline compared to baseline. Endline respondents are 4.7 pp. more likely to report that they ate vegetables at least once a week during their pregnancy than at baseline (an increase from 91.9% to 96.7%), 2.6 pp. more likely to report that they ate fruits at least once a week during their pregnancy than at baseline (an increase from 95.8% to 98.4%), and 6.2 pp. more likely to report that they ate liver at most twice a month during their pregnancy than at baseline (an increase from 57.6% to 63.8%). Second, we find that breastfeeding and immunisation intentions at baseline are closely aligned with intentions and behaviour at endline. At baseline 67.4% of respondents say that they intend to exclusively breastfeed their baby for six months compared to 65.5% at endline. Similarly, 88.5% of respondents say that the benefits of vaccinations outweigh the risks (implying that they intend to fully immunise their baby) and we find that 89.9% of respondents say that their baby received all six vaccinations scheduled at six weeks post birth. Again, it is important to note that an absence of a larger increase in these outcomes does not imply that MomConnect had no causal impact on these outcomes since there might have been a substantial gap between intentions and eventual behaviour in absence of MomConnect.

[23] We focus on these outcomes as they are the only ones measured in a comparable way both at baseline and endline.

Table 1: Descriptive Statistics and Balance

Sample:	(1) All		(3) ARM		(5) RCM-WA		(7) RCM-SMS		(9) RCM-BCM		(11) F-test of joint significance (p-value)
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	
A. Characteristics of study participants at registration (N=8243)											
Age (in years)	27.8	6	27.7	6.2	27.9	6	27.8	6	27.9	6	.374
Gestational age(in weeks)	23.7	4	23.7	3.9	23.7	4	23.7	3.9	23.7	4	.68
Province: Eastern Cape	.055	.227	.055	.229	.055	.228	.054	.226	.054	.226	
Province: Free State	.051	.219	.051	.219	.05	.218	.052	.222	.05	.218	
Province: Gauteng	.449	.497	.451	.498	.448	.497	.448	.497	.446	.497	
Province: KwaZulu-Natal	.117	.322	.114	.318	.117	.322	.118	.323	.12	.325	
Province: Limpopo	.074	.262	.074	.262	.076	.266	.072	.259	.074	.262	
Province: Mpumalanga	.015	.121	.015	.121	.015	.12	.015	.122	.015	.122	
Province: North West	.093	.291	.092	.289	.092	.289	.094	.293	.095	.294	
Province: Northern Cape	.085	.279	.088	.283	.085	.279	.085	.279	.083	.276	
Province: Western Cape	.061	.239	.06	.238	.061	.24	.06	.238	.062	.241	
B. Characteristics of endline sample (N = 5250)											
Age (in years)	28.1	5.9	28	5.80	28.4	5.9	28.1	6.10	28	5.9	.143
Gestational age (in weeks)	23.7	4	23.8	3.9	23.7	4	23.8	3.9	23.6	4	.636
Province: Eastern Cape	.053	.224	.053	.224	.053	.224	.05	.218	.056	.23	
Province: Free State	.05	.218	.048	.213	.05	.217	.053	.224	.049	.217	
Province: Gauteng	.449	.497	.461	.499	.448	.497	.447	.497	.439	.496	
Province: KwaZulu-Natal	.123	.328	.12	.326	.116	.321	.128	.334	.126	.332	
Province: Limpopo	.081	.273	.079	.27	.086	.28	.08	.271	.079	.27	
Province: Mpumalanga	.015	.122	.016	.125	.016	.126	.014	.116	.015	.12	
Province: North West	.091	.288	.089	.284	.088	.283	.094	.292	.093	.291	
Province: Northern Cape	.084	.278	.083	.276	.087	.282	.082	.274	.085	.279	
Province: Western Cape	.054	.227	.051	.22	.057	.231	.054	.225	.057	.232	
Child's age (in weeks)	9.30	2.4	9.4	2.4	9.30	2.4	9.4	2.5	9.30	2.4	.632

Sample:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	All		ARM		RCM-WA		RCM-SMS		RCM-BCM		F-test of joint significance (p-value)
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	
B. Characteristics of endline sample (N = 5250) - continued											
Education: Primary	.183	.387	.197	.398	.186	.39	.191	.393	.157	.364	.044
Education: Secondary	.547	.498	.534	.499	.552	.497	.532	.499	.571	.495	.144
Education: Tertiary	.265	.441	.264	.441	.259	.438	.269	.444	.267	.442	.964
Unemployed	.429	.495	.428	.495	.443	.497	.429	.495	.415	.493	.535
Monthly Income (in Rand)	5007	5472	5120	5731	4998	5364	4890	5411	5020	5375	.779
Ethnicity: Black African	.944	.23	.938	.241	.94	.237	.951	.216	.946	.226	.395
Ethnicity: Colored	.042	.201	.042	.202	.043	.203	.04	.196	.043	.203	.993
Ethnicity: Indian/Asian	.004	.066	.006	.078	.005	.068	.003	.055	.004	.062	.691
Ethnicity: White	.008	.087	.011	.102	.009	.095	.005	.073	.005	.073	.283
Relationship: No Partner	.178	.383	.185	.388	.17	.376	.169	.375	.19	.392	.39
First-time mothers	.38	.485	.387	.487	.366	.482	.38	.486	.387	.487	.774
C. Characteristics of 2023 MomConnect Users (N = 212025)											
Age (in years)	26.4										
Gestational age (in weeks)	20.1										
Province: Eastern Cape	.049										
Province: Free State	.042										
Province: Gauteng	.397										
Province: KwaZulu-Natal	.1										
Province: Limpopo	.156										
Province: Mpumalanga	.085										
Province: North West	.087										
Province: Northern Cape	.014										
Province: Western Cape	.07										

Note: Each row states the sample mean and standard deviation of a variable, and by treatment group. The last column reports the p-value from the F-test of joint significance of the treatment dummies. This is derived from a regression of each variable on the three treatment dummies, controlling for the stratification variables and using heteroskedasticity-robust standard errors.

Table 2: Difference between all registered births and both our endline sample and MomConnect users by province

Sample:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Endline Sample (N = 5,250)		2023 MomConnect Users (N = 212,025)		All Registered 2022 Births in South Africa (N = 998,362)		Difference between all registered births in SA and:	
	%	S.D.	%	S.D.	%	S.D.	Endline Sample	All MC Users
Province								
Eastern Cape	6%	.227	5%	*NA*	11%	*NA*	-6 pp.	-6 pp.
Free State	5%	.219	4%	*NA*	5%	*NA*	0 pp.	0 pp.
Gauteng	45%	.497	40%	*NA*	23%	*NA*	22 pp.	16 pp.
KwaZulu-Natal	12%	.322	10%	*NA*	22%	*NA*	-10 pp.	-12 pp.
Limpopo	7%	.262	16%	*NA*	12%	*NA*	-5 pp.	3 pp.
Mpumalanga	2%	.121	9%	*NA*	9%	*NA*	-7 pp.	0 pp.
North West	9%	.291	9%	*NA*	6%	*NA*	3 pp.	3 pp.
Northern Cape	9%	.279	1%	*NA*	2%	*NA*	6 pp.	-1 pp.
Western Cape	6%	.239	7%	*NA*	9%	*NA*	-3 pp.	-2 pp.

Notes: Each row states the percentage (and standard deviation in case our endline sample) of the total sample that is from that province. Data is presented for the endline sample (columns 1 and 2), all 2023 MomConnect users (columns 3 and 4), and all registered births in South Africa, as per statssa.gov.za (columns 5 and 6). Column (7) shows the difference between columns (5) and (1), and column (8) the difference between (5) and (3).

5. Results

This section presents the main results of this study. In section 5.1 we present the results on primary outcomes: ANC visits and immunisation rates. In section 5.2, we present the results on secondary outcomes: knowledge of healthy maternal and infant health behaviours and adoption of healthy maternal and infant behaviours.

We report p-values for four hypotheses in **Tables 2-12** : 1) RCM-WA = ARM (control); 2) RCM-BCM = ARM (control); 3) RCM-WA = RCM-BCM; and 4) RCM-SMS = RCM-WA. The fourth p-value compares RCM-SMS with RCM-WA rather than the control arm as the RCM-WA model is the most similar to the RCM-SMS model, allowing for the isolation of the effect of WhatsApp vs. SMS.

Within the context of decision-making around scale-up, we believe that a 10% significance level is sufficient to base decision-making on the results of this study although we understand that this is a departure from standard academic / publication definitions of statistical significance. We report p-values for every finding and clearly indicate whether each result is significant at the 5% or 10% level. However, our recommendations are based on considering 10% as the decision-relevant significance level rather than 5%.

5.1. Primary Outcomes

ANC visits

First, we assess the relative impact of the different messaging models on the number of ANC visits. **Table 3** shows the regressions results for the three ANC visit outcomes: the specifications in the odd-numbered columns control for all unique combinations of our three stratification variables and estimate heteroskedasticity-robust standard errors; in addition to that, the specifications in the even-numbered columns control for demographic characteristics and enumerator fixed effects.[24] **Figure 4** also presents the results of column 1 visually.

Among respondents in the ARM arm, 43.6% report having attended at least eight ANC visits.[25] The estimated coefficients for all three treatment arms are close to zero, statistically insignificant, and fairly precisely estimated; we can rule out effects exceeding +/- 5 percentage points with 95% confidence. The estimated coefficients are similarly small, also insignificant at 10% significance level, and similarly precisely estimated using at least four ANC visits, where the mean is 96.3%, or the continuous measure of ANC visits, where the mean is 7.34, as the outcome variable. The null results are also robust to controlling for demographic controls and enumerator fixed effects for all

[24] We collected demographic information from respondents at endline. However, all demographic characteristics we control for can either not possibly be affected by the treatments (whether the respondent's most recent pregnancy was her first pregnancy; ethnicity) or we asked respondents to report the information for the time before they learnt about their most recent pregnancy and therefore before they signed up for MomConnect (educational attainment; employment status; marital status).

[25] 96.1% of respondents in the ARM arm reported having attended at least four ANC visits. The mean (median) number of ANC reported by respondents is 7.3 (7).

three outcomes. **Overall, our results suggest the effectiveness in improving ANC visit attendance does not differ meaningfully across the four messaging models**.

The similarity across messaging models for ANC appointment reminders could be a potential reason for the statistically insignificant ANC findings. All four study arms receive the same appointment reminders at the beginning of the messages from MomConnect (see Appendix 4). It seems plausible that the four messaging models therefore have a similar impact on ANC visit attendance. Another consideration is that in order to sign up for MomConnect, a woman usually had attended at least one ANC visit, implying that the participants in this study were already connected to the health system and probably more likely to continue attending ANC visits even in the absence of MomConnect. If it was possible for MomConnect to register mothers without attending their first ANC visit, then further research could be conducted to see whether pregnant women who had not yet attended their first ANC visit are more impacted by additional maternal health information, especially nudges towards attending ANC visits.

We conducted sub-group analyses of the ANC visits for four different groups, which we believe are "low information" users. These are users who likely had less information about maternal and infant health at baseline, and therefore have a higher potential for informational treatments to work. The low information measures we construct are: 1) first time mothers; 2) mothers under 30; 3) gestational age lower than 20 weeks when they registered on MomConnect; and 4) level of education below high school.[26] We ran heterogeneous effects models for our primary ANC outcome, interacting our treatment arms with our measures of "low information". **Table 4** shows there are no significant effects for any of the interaction terms across the four measures. Thus, our finding that there are no meaningful differences in ANC visit attendance between the different messaging models also holds looking at the group of users who were most likely to respond to the additional information conveyed through MomConnect.

[26] While women registering at an earlier gestational age do not necessarily have lower information, they are exposed to MomConnect for longer, again increasing the potential for the information on the platform to be effective.

Table 3: Primary Outcomes Table - ANC Visits

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	At least 8 ANC visits		At least 4 ANC visits		Number of ANC visits	
RCM-SMS	.007	.006	.006	.006	.19	.2
	(.019)	(.019)	(.007)	(.007)	(.103)	(.101)
RCM-WA	.007	.01	0	0	.061	.081
	(.019)	(.019)	(.008)	(.008)	(.097)	(.096)
RCM-BCM	-.015	-.013	.004	.004	-.003	.021
	(.019)	(.019)	(.008)	(.008)	(.098)	(.098)
Observations	5142	5125	5142	5125	5142	5125
Mean Dep. Var.	0.436	0.436	0.963	0.963	7.341	7.341
Mean Dep. Var. in ARM	0.436	0.436	0.961	0.961	7.279	7.279
Stratification Variables	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	No	Yes	No	Yes
Enumerator Fixed Effects	No	Yes	No	Yes	No	Yes
p-value ARM = RCM-WA	.719	.592	.982	.992	.532	.401
p-value ARM = RCM-BCM	.429	.509	.634	.596	.974	.83
p-value RCM-WA = RCM-BCM	.249	.229	.618	.584	.521	.543
p-value RCM-SMS = RCM-WA	.986	.836	.421	.44	.212	.244

Notes: All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors. In addition, specifications in columns (2), (4), and (6) control for enumerator fixed effects as well as educational attainment, relationship status, employment status, ethnicity, and whether the respondent is a first-time mother.

Figure 4: Treatment Effects on "At Least 8 ANC Visits" Indicator

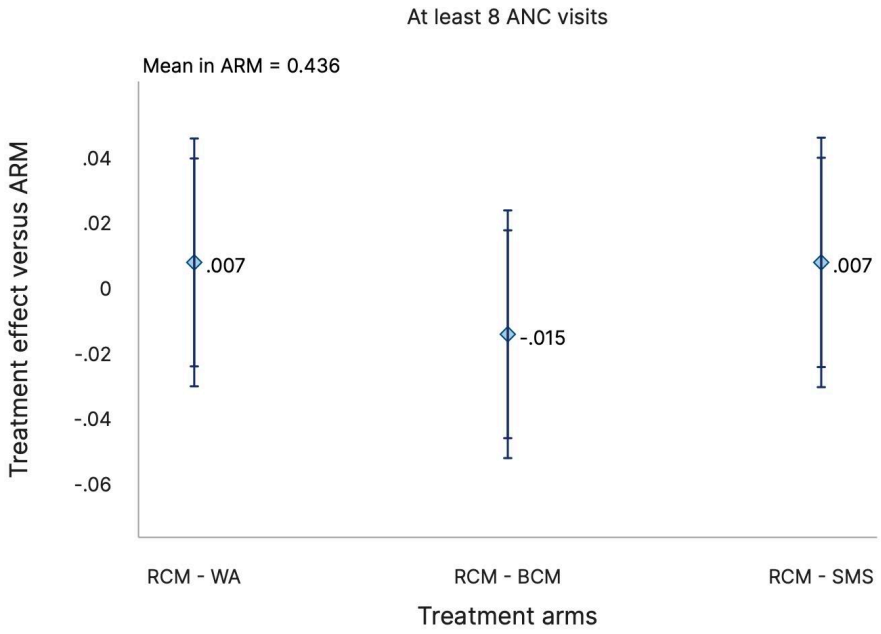


Figure 5 plots the difference in the proportion of mothers who attended at least 8 ANC visits between each treatment model and the ARM model (column 1 in Table 3). The coefficients are estimated from a regression of the binary outcome variable (respondent attended 8 or more ANC visits) on the treatment dummies, controlling for stratification variables and estimating robust standard errors. Bars show 90% and 95% Confidence Intervals, the narrower bars show 95% Confidence Intervals.

Table 4: Subgroup Analysis - ANC Visits

Dep. Var.	(1)	(2)	(3)	(4)
	At least 8 ANC visits			
Definition of covariate X	First-time mothers = {0,1}	Mother's age <= 30 = {0,1}	Gestational age > 20 = {0,1}	Has not finished grade school = {0,1}
RCM-SMS	-.003	-.02	.014	.003
	(.025)	(.034)	(.039)	(.022)
RCM-WA	.008	-.033	.037	.008
	(.024)	(.033)	(.039)	(.022)
RCM-BCM	-.028	-.037	-.013	-.007
	(.024)	(.033)	(.039)	(.022)
X	.024	-.063	-.111	-.006
	(.029)	(.029)	(.032)	(.035)
RCM-SMS*X	.026	.039	-.01	.025
	(.04)	(.042)	(.045)	(.051)
RCM-WA*X	.002	.062	-.04	-.005
	(.04)	(.041)	(.045)	(.049)
RCM-BCM*X	.036	.03	-.005	-.042
	(.04)	(.041)	(.045)	(.049)
Observations	5137	5142	5142	5132
Mean Dep. Var.	.436	.436	.436	.436
Mean Dep. Var. in ARM	.436	.436	.436	.436
Mean X	.38	.681	.743	.188
Stratification Variable	Age - Gestational Age - Province	Gestational Age - Province	Age - Province	Age - Gestational Age - Province
p-value ARM*X = RCM-WA*X	.955	.13	.365	.924
p-value ARM*X = RCM-BCM*X	.37	.466	.914	.391
p-value RCM-WA*X = RCM-BCM*X	.403	.435	.429	.449
p-value RCM-SMS*X = RCM-WA*X	.551	.59	.497	.568

Notes: The outcome variable for all columns is a binary indicator for whether the respondent reported that she attended at least 8 ANC visits during her most recent pregnancy. All specifications regress this outcome on treatment indicators, a variable indicated in the column heading meant to capture "low information" respondents, and their interaction terms. All regressions include stratification variables, excluding the variable indicated in the column heading where applicable, and use heteroskedasticity-robust standard errors.

Immunisations

Next, we assess the relative impact of the different messaging models on infant immunisation coverage. **Table 5** shows the results for immunisation outcomes. Like **Table 3** for ANC outcomes, the specifications in the odd-numbered columns control for the randomisation strata and estimate heteroskedasticity-robust standard errors while the specifications in the even-numbered columns also control for demographic controls and enumerator fixed effects. **Figure 5** displays the results of the primary specification (column 1) visually.

Among respondents in the ARM arm, 90.8% report that their baby received all six vaccinations that are recommended (two at birth and four at six weeks old), by the time of the interview.[27] **Our main finding is that the RCM-BCM arm shows a 3 percentage point reduction in full immunisation coverage compared to the ARM arm, with a p-value of 0.018**, and therefore significant at the 5% level. This result is robust to controlling for demographic characteristics and enumerator fixed effects, and, qualitatively, to using a continuous measure of the number of vaccinations the baby received as an outcome. While the coefficients for the other two messaging arms are also negative, the estimated effects are smaller (-1.7 pp. for RCM-SMS vs. ARM and -1.5 pp. for RCM-WA vs. ARM) and not significant at conventional significance levels (the p-value for RCM-SMS = RCM WA is 0.882 and the p-value for ARM = RCM-WA is 0.194). The difference between the RCM-BCM arm and the RCM-WA arm is also not significant at conventional significance levels (p-value 0.289).

For our immunisation outcomes, we can directly assess the degree of measurement error of self-reported outcomes: in our endline survey, we asked all respondents whether they have their baby's immunisation card available and - if yes - to read the information on past vaccinations the baby received directly from the card. Comparing coefficient estimates for the sub-sample who provided the information based on the card (73% of the sample) with the full sample (which includes 27% of respondents who provided information based on their recall) can shed light on the question whether answering based on recall introduces significant measurement error.[28] Thus, for both immunisation outcomes we report coefficients for the full sample and the sub-sample of respondents who answered based on the vaccination card. The results presented above are robust after controlling for only those with the vaccination card presently with them.

The lack of statistically significant results for RCM-WA and RCM-SMS could be driven by the already very high levels of immunisation across the study sample. The finding that across all arms, 89% of infants were fully immunised at six weeks is a useful contribution to understanding of infant health in South Africa due to a lack of recent data on immunisations at six weeks, despite only

[27] The mean age of babies at interview is 9.3 weeks and the mean (median) number of vaccinations respondents reported is 5.8 (6).

[28] It is still possible that respondents who answered based on the card did not tell the truth since we had no way of actually observing the card during the phone interviews. Still, we believe it is likely that measurement error is lower in the sub-sample that answered based on the card compared to the sub-sample that answered based on recall. Please also note that the fact that we do not find any evidence of substantial measurement error for recall-respondents for the immunisation outcome does not necessarily imply that this is true for other outcomes. For example, respondents might find it easier to recall vaccinations (that happened within six weeks of the interview) than the number of ANC visits they attended (some of which happened ten months prior to the interview).

representing a specific sample of women who are on MomConnect. As a comparison, 88% of babies in South Africa received DTP-containing vaccine, which is administered at 6 weeks, in 2016, while 99% of babies receive the same vaccination in our study sample. Since no more recent nationally representative data on coverage rates for vaccinations given at 6 weeks are available for South Africa, we are unsure how coverage for babies of MomConnect users compares to current national coverage rates. Additionally, the final 11% that are not fully immunised may be a particularly hard group to impact, and their barriers to immunisation may not be information, but for example access to vaccinations. Thus, it might be difficult for informational treatments to have an effect, even if the new messaging models would have had a positive effect on immunisation rates in populations with a lower baseline coverage rate.

The negative impact of 3 percentage points of RCM-BCM on immunisation outcomes is difficult to explain. One hypothesis is that the RCM-BCM provides access to more information, in general, and this could be overwhelming to users, or result in a lack of focus on the information provided that is deemed the most relevant information (for instance, on immunisations). The fact that participants are able to browse a menu of different maternal health information messages could feasibly result in less key information, such as the importance of vaccinations, being absorbed by the users. However, this explanation is speculative, and we are highly uncertain what could explain this result. Another possibility is that access to more information may reveal more risks or downsides of vaccines, like redness, swelling, or pain at the injection site. Whilst this is real information and is important for mothers to know, it may feasibly have had a negative impact on vaccination uptake.

Similarly to ANC visits, we conducted sub-group analyses of our primary immunisation outcome on four different “low information” groups (see the ANC sub-section above). We do find evidence that the RCM-WA and RCM-BCM arms are substantially more effective than the ARM arm for women below 30 as shown in column 2: in the group of women below 30, the RCM-WA model leads to an increase in full immunisation of 4.1 pp. (p-value 0.098) and the RCM-BCM model leads to an increase of 7.8 pp. (p-value 0.003) compared to women above 30. While this is again speculative, a possible explanation for this might be that younger women are more familiar with WhatsApp or spend more time on their phones, which might make them more likely to engage more with MomConnect or to discover additional content on the platform. This could in turn update their beliefs on the importance of vaccinations and affect their subsequent behaviour. We only find significant heterogeneous effects for users’ age, but not for any of our other three measures of low information.

Table 5: Primary Outcomes Table - Immunisations

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All 6 Vaccinations				Number of Vaccinations			
Sample	Full Sample		Has Vaccination Card		Full Sample		Has Vaccination Card	
RCM-SMS	-.017	-.018	-.019	-.021	-.012	-.015	-.01	-.014
	(.012)	(.012)	(.014)	(.014)	(.019)	(.019)	(.023)	(.023)
RCM-WA	-.015	-.019	-.011	-.015	-.021	-.027	-.02	-.028
	(.012)	(.012)	(.014)	(.014)	(.02)	(.02)	(.025)	(.025)
RCM-BCM	-.029	-.028	-.03	-.028	-.036	-.036	-.039	-.04
	(.012)	(.012)	(.015)	(.015)	(.02)	(.019)	(.024)	(.024)
Observations	4985	4969	3623	3612	4985	4969	3623	3612
Mean Dep. Var.	0.894	0.894	0.888	0.888	5.849	5.849	5.834	5.834
Mean Dep. Var. in ARM	0.908	0.908	0.903	0.903	5.865	5.865	5.851	5.851
Stratification Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Enumerator Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
p-value ARM = RCM-WA	.194	.113	.439	.294	.281	.171	.411	.26
p-value ARM = RCM-BCM	.018	.019	.045	.051	.068	.063	.107	.1
p-value RCM-WA = RCM-BCM	.289	.449	.227	.371	.474	.647	.459	.639
p-value RCM-SMS = RCM-WA	.882	.968	.583	.693	.649	.537	.695	.577

Notes: All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors. In addition, specifications in columns (2), (4), (6), and (8) control for enumerator fixed effects as well as educational attainment, relationship status, employment status, ethnicity, and whether the respondent is a first-time mother. Columns (3), (4), (7) and (8) restrict the sample to respondents who reported having access to a vaccination card during the endline interview.

Figure 5: Treatment Effects on "Baby Received all 6 Vaccines" Indicator

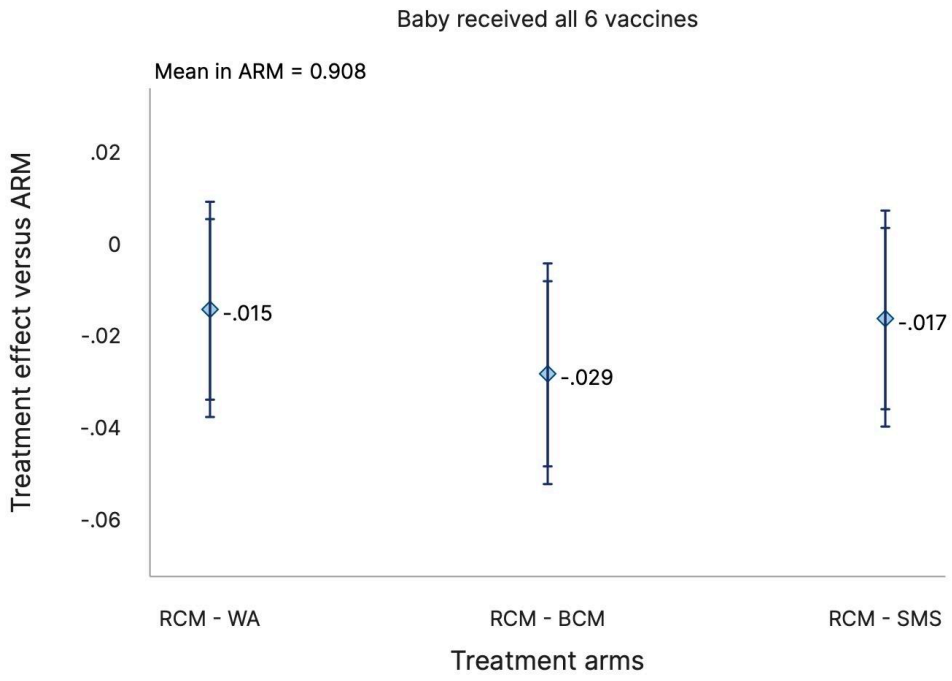


Figure 4 plots the difference in the proportion of children who have received all 6 vaccinations between each treatment model and the ARM model (column 1 in Table 5). The coefficients are estimated from a regression of the outcome variable (All 6 vaccinations) on the treatment dummies, controlling for stratification variables and estimating robust standard errors. Bars show 90% and 95% Confidence Intervals, the narrower bars show 95% Confidence Intervals.

Table 6: Subgroup Analysis - Immunisations

Dep. Var.	(1)	(2)	(3)	(4)
	All 6 vaccines			
Definition of covariate X	First-time mothers = {0,1}	Mother's age <= 30 = {0,1}	Gestational age > 20 = {0,1}	Has not finished grade school = {0,1}
RCM-SMS	-.001	-.028	-.009	-.015
	(.015)	(.019)	(.022)	(.013)
RCM-WA	-.014	-.042	-.021	-.017
	(.015)	(.019)	(.023)	(.013)
RCM-BCM	-.038	-.08	-.022	-.025
	(.016)	(.021)	(.023)	(.013)
X	.004	-.033	-.012	-.027
	(.017)	(.016)	(.018)	(.022)
RCM-SMS*X	-.044	.016	-.01	-.013
	(.025)	(.025)	(.026)	(.034)
RCM-WA*X	-.004	.041	.009	.006
	(.025)	(.025)	(.027)	(.032)
RCM-BCM*X	.027	.078	-.007	-.012
	(.025)	(.026)	(.027)	(.033)
Observations	4981	4985	4985	4976
Mean Dep. Var.	.894	.894	.894	.894
Mean Dep. Var. in ARM	.908	.908	.908	.908
Mean X	.38	.681	.743	.188
Stratification Variable	Age - Gestational Age - Province	Gestational Age - Province	Age - Province	Age - Gestational Age - Province
p-value ARM*X = RCM-WA*X	.869	.094	.746	.855
p-value ARM*X = RCM-BCM*X	.278	.003	.798	.726
p-value RCM-WA*X = RCM-BCM*X	.235	.177	.581	.61
p-value RCM-SMS*X = RCM-WA*X	.132	.337	.506	.588

Notes: The outcome variable for all columns is a binary indicator for whether the respondent reported that her newborn received all 6 vaccinations recommended by the South African immunisation schedule at six weeks after births. All specifications regress this outcome on treatment indicators, a variable indicated in the column heading meant to capture "low information" respondents, and their interaction terms. All regressions include stratification variables, excluding the variable indicated in the column heading where applicable, and use heteroskedasticity-robust standard errors.

5.2. Secondary Outcomes

Knowledge of Healthy Behaviours

First, we assess the relative impact of the different messaging models on respondents' knowledge of healthy behaviours. We measure this outcome through our pre-specified knowledge index aggregating knowledge questions on baby danger signs, breastfeeding, and anaemia. Columns 1 and 2 of **Table 7** show the results for the knowledge index. The specification in column 1 controls for the randomisation strata and estimates heteroskedasticity-robust standard errors while the specification in column 2 also controls for demographic controls and enumerator fixed effects. **Figure 6** displays the results of the primary specification of the knowledge index (column 1) visually.

The mean of the knowledge index in the ARM arm is 5.2 (with 0 being the minimum and 12 the maximum attainable value) with a standard deviation of 1.6. The coefficients for both the RCM-WA (0.119) and RCM-BCM (0.115) arms are positive and have a magnitude of approximately 0.07 standard deviations, but are not statistically different from the ARM arm at conventional statistical significance levels (the p-value for ARM = RCM-WA is 0.223 and the p-value for ARM = RCM-BCM is 0.235). The coefficient for the RCM-SMS arm is negative (-0.077) and, when compared with the RCM-WA value of 0.119, is statistically significantly different from the RCM-WA arm (p-value 0.039). **The RCM-WA arm improves respondents' knowledge by 0.196 index points (0.13 standard deviations) compared to the RCM-SMS arm, significant at the 5% level**.[29] All results are similar in terms of magnitude and statistical significance when controlling for demographic characteristics and enumerator fixed effects.

The difference between the RCM-WA and RCM-SMS arms suggests that receiving information via WhatsApp is slightly preferable for improving maternal and infant health knowledge. This might be explained by WhatsApp making it easier to access additional information. However, the overall magnitude of the difference is modest and in contexts where WhatsApp is not available, providing similar content via SMS does not seem substantially worse for improving maternal and infant health knowledge.

It could be questioned whether the small and mostly insignificant differences in the knowledge index across messaging models mask underlying heterogeneity of treatment effects on the different topic areas feeding into the aggregate index. **Table 8** shows the treatment effect estimates for the three index components separately. We find that five out of six coefficients for the RCM-WA and RCM-BCM arms are positive (the RCM-WA coefficient is negative for the anaemia score, but close to 0 and insignificant), but none are significantly different from the ARM arm at conventional significance levels. The overall significant difference between the RCM-SMS and the RCM-WA arms seems to be driven by the breastfeeding score which, when comparing the -0.07 of the SMS arm with the with the 0.069 of the RCM-WA arm, is substantially and statistically significantly (p-value 0.005) higher in the RCM-WA arm.

[29] The standard deviation of the knowledge index is 1.5 in the RCM-SMS arm.

Figure 6: Treatment Effects on Knowledge Index

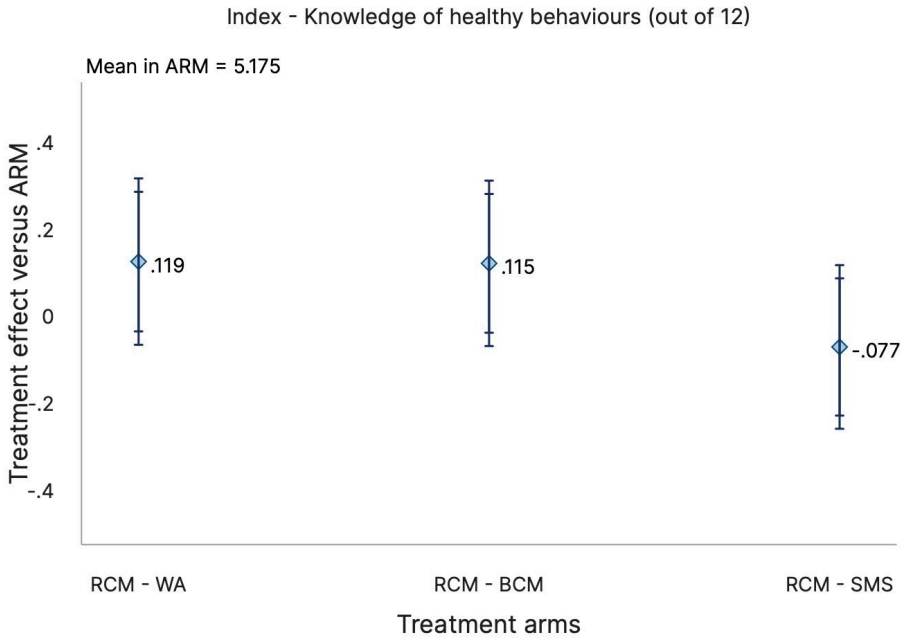


Figure 6 plots the difference in the index score for knowledge of healthy behaviors between each treatment model and the ARM model (column 1 in Table 7). The coefficients are estimated from a regression of the outcome variable (the knowledge index) on the treatment dummies, controlling for stratification variables and estimating robust standard errors.
Bars show 90% and 95% Confidence Intervals, the narrower bars show 95% Confidence Intervals.

Table 7: Secondary Outcomes Table - Knowledge and Adoption of Healthy Maternal Health Behaviours

Dep. Var.	(1) Knowledge of Healthy Behaviors	(2) Knowledge of Healthy Behaviors	(3) Adoption of Healthy Behaviors	(4) Adoption of Healthy Behaviors
RCM-SMS	-.077	-.137	.054	0
	(.096)	(.095)	(.135)	(.133)
RCM-WA	.119	.103	.216	.199
	(.097)	(.097)	(.131)	(.13)
RCM-BCM	.115	.107	.239	.244
	(.097)	(.095)	(.133)	(.13)
Observations	2178	2171	2178	2171
Mean Dep. Var.	5.222	5.222	11.485	11.485
Mean Dep. Var. in ARM	5.175	5.175	11.357	11.357
Stratification Variable	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	No	Yes
Enumerator Fixed Effects	No	Yes	No	Yes
p-value ARM = RCM-WA	.223	.289	.1	.126
p-value ARM = RCM-BCM	.235	.258	.073	.061
p-value RCM-WA = RCM-BCM	.97	.962	.867	.727
p-value RCM-SMS = RCM-WA	.039	.011	.233	.141

Note: The outcomes in this table are summative indices of different topics on healthy maternal behaviors. The outcome in columns (1) and (2) is a summative knowledge index capturing knowledge about baby danger signs, breastfeeding, and anemia. The maximum attainable value for the index is 12. The outcome in columns (3) and (4) is a summative behaviour index capturing behaviour related to diet, pregnancy preparedness, breastfeeding, and anemia. The maximum attainable value for the index is 16. All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors. In addition, specifications in columns (2) and (4) control for enumerator fixed effects as well as educational attainment, relationship status, employment status, ethnicity, and whether the respondent is a first-time mother.

Table 8: Secondary Outcomes Table - Detailed breakdown of Knowledge Index

Dep. Var.	(1)	(2)	(3)	(4)
	Knowledge of Healthy Behaviours			
	Baby danger signs (0-7)	Breastfeeding (0-3)	Anemia (0-5)	All Knowledge (0-15)
RCM-SMS	.042	-.07	-.011	-.038
	(.068)	(.05)	(.069)	(.11)
RCM-WA	.086	.069	-.028	.127
	(.069)	(.048)	(.07)	(.113)
RCM-BCM	.083	.025	.043	.151
	(.068)	(.048)	(.071)	(.112)
Observations	2178	2178	2178	2178
Mean Dep. Var.	2.534	2.018	1.354	5.906
Mean Dep. Var. in ARM	2.488	2.007	1.347	5.842
Stratification Variable	Yes	Yes	Yes	Yes
Demographic Controls	No	No	No	No
Enumerator Fixed Effects	No	No	No	No
p-value ARM = RCM-WA	.208	.152	.689	.26
p-value ARM = RCM-BCM	.225	.609	.543	.178
p-value RCM-WA = RCM-BCM	.962	.356	.308	.83
p-value RCM-SMS = RCM-WA	.526	.005	.8	.132

Notes: All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors. Columns (1) - (3) report knowledge scores for baby danger signs, breastfeeding, and anemia. Column (4) combines the three outcomes from columns (1) - (3).

Adoption of Healthy Behaviours

Second, we assess the relative impact of different messaging models on respondents' adoption of healthy behaviours, which we measure through our pre-specified behaviour index. The index aggregates questions on respondents' self-reported behaviours related to nutrition, pregnancy preparedness, breastfeeding practice, and anaemia prevention. The results for the behaviour index are in columns 3 and 4 of **Table 7** above. As was the case for the knowledge index, the specification in column 3 controls for the randomisation strata and estimates heteroskedasticity-robust standard errors while the specification in column 4 also controls for demographic controls and enumerator fixed effects. **Figure 6** displays the results of the primary specification of the behaviour index (column 3) visually.

The mean of the behaviour index in the ARM arm is 11.4 (with 0 being the minimum and 16 the maximum attainable value), with a standard deviation of 2.1. **Both the RCM-WA (0.216) and RCM-BCM (0.239) arms have small positive impacts and are statistically significant at the 10% level** (the p-value for ARM = RCM-WA is 0.1 and the p-value for ARM = RCM-BCM is 0.073). Note that in column 4 in **Table 7,** where additional control variables are added, RCM-WA is not significant at the 10% level (p-value of 0.126) and is only significant at the 10% level in our primary specification. The coefficient sizes of 0.216 and 0.239 represent magnitudes of 0.1 and 0.11 standard deviations. The coefficient of the RCM-SMS arm is 0.054, implying a difference of 0.162 index points (0.07 standard deviations) between RCM-WA and RCM-SMS.[30] This difference is not statistically significant at conventional significance levels (p-value 0.233). The results are similar in terms of magnitude and statistical significance when controlling for demographic characteristics and enumerator fixed effects.

Next, we consider whether a specific topic area drives the modest increases in the behaviour index we find for the RCM-WA and RCM-BCM arms. **Table 9** shows the treatment effect estimates for the four index components separately. As column 1 shows, both the RCM-WA and RCM-BCM arms affect the nutrition score the most, with the p-values below or close to conventional significance levels (the p-value for ARM = RCM-WA is 0.033 and the p-value for ARM = RCM-BCM is 0.114). While the coefficients for the RCM-WA and RCM-BCM arms are positive in seven out of eight cases (the RCM-WA coefficient is negative for the pregnancy preparedness score, but close to 0 and insignificant), all but the ones for the nutrition score are small and insignificant at conventional levels. There is no statistically significant difference between the RCM-SMS and RCM-WA arms for any of the topic areas. In terms of magnitude, the diet and breastfeeding scores drive the overall positive effect of the RCM-WA arm on the behaviour index compared to the ARM arm. For the RCM-BCM arm, all four topic areas captured by the behaviour index contribute to the overall effect.

[30] The standard deviation of the behaviour index is 2.3 in the RCM-SMS arm.

Figure 7: Treatment Effects on Behaviours Index

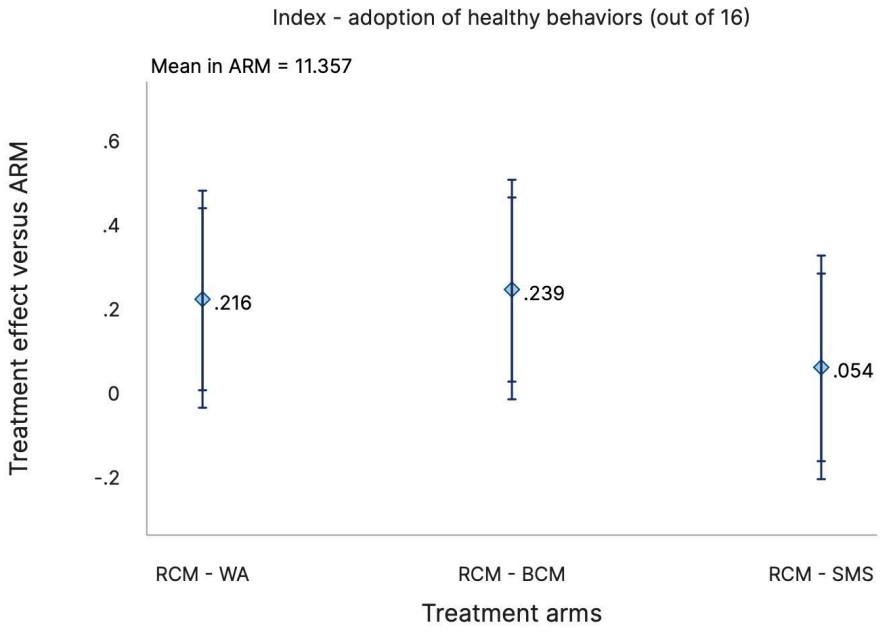


Figure 7 plots the difference in the index score for adoption of healthy behaviours between each treatment model and the ARM model (column 3 in Table 7). The coefficients are estimated from a regression of the outcome variable (the adoption index) on the treatment dummies, controlling for stratification variables and estimating robust standard errors. Bars show 90% and 95% Confidence Intervals, the narrower bars show 95% Confidence Intervals.

Table 9: Secondary Outcomes Table - Detailed Adoption of Healthy Maternal Health Behaviours

Dep. Var.	(1)	(2)	(3)	(4)	(5)
	Adoption of Healthy Behaviours				
	Diet (0-4)	Pregnancy prep. (0-7)	Breastfeeding (0-2)	Anemia treatment (0-4)	All Behavior (0-17)
RCM-SMS	.072	-.066	.016	-.012	.01
	(.055)	(.115)	(.043)	(.053)	(.148)
RCM-WA	.114	-.014	.051	.009	.16
	(.053)	(.112)	(.041)	(.051)	(.143)
RCM-BCM	.088	.077	.025	.058	.247
	(.055)	(.11)	(.043)	(.05)	(.141)
Observations	2178	2178	2178	2178	2178
Mean Dep. Var.	2.656	4.776	1.291	3.519	12.242
Mean Dep. Var. in ARM	2.588	4.77	1.269	3.506	12.132
Stratification Variable	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	No	No	No	No
Enumerator Fixed Effects	No	No	No	No	No
p-value ARM = RCM-WA	.033	.903	.214	.868	.264
p-value ARM = RCM-BCM	.114	.484	.568	.25	.081
p-value RCM-WA = RCM-BCM	.631	.417	.534	.348	.542
p-value RCM-SMS = RCM-WA	.44	.653	.406	.707	.317

Notes: All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors. Columns (1) - (4) report behaviour scores for diet, preganncy preparedness, breastfeeding, and anemia treatment. Column (5) combines the four outcomes from columns (1) - (4).

User Experience

Table 10 assesses the relative impact of the different messaging models on which particular component of MomConnect the participants find most useful. RCM-WA and RCM-BCM users are 6.3 and 5.5 percentage points less likely to find appointment reminders the most useful component, compared with ARM users. This is statistically significant at conventional levels for both with p-values of 0.028 and 0.062 and so RCM-WA is significant at 5% level, whereas RCM-BCM is statistically significant at 10% level. Conversely, RCM-WA and RCM-BCM users are 4.2 and 6.3 percentage points more likely to find baby information the most useful component of MomConnect, with p-values of 0.106 (so not quite significant at 10% level) and 0.021 (significant at 5% level). **The fact that users of RCM-WA and RCM-BCM are more likely to find the information content most useful, specifically baby information, as opposed to the appointment reminders, indicates that the additional information is positively viewed by participants.** It might be the case that this additional information is easier to access in the RCM-WA and RCM-BCM models, which could explain both this result and the improvement in the adoption of healthy behaviours in the two arms we found in the previous sub-section.

Tables 10 and 11 show a wider range of user experience questions and the relative impact of different messaging models on each of these. **Table 11** focusses on message preferences. A key finding here is that a slight majority of SMS users (62%) prefer to receive messages on SMS, rather than WhatsApp, while the vast majority of WhatsApp users (91%) prefer to receive messages on WhatsApp. **This suggests that for the majority of users, the messaging platform they were assigned to is a good experience, irrespective of which platform it was.** Together with the lack of strong differences in outcomes between RCM-WA and RCM-SMS arms, the results show that a well-designed SMS messaging model can achieve most of the impact of a WhatsApp messaging model. **Overall, it is clear that in our population of smartphone owners who have access to WhatsApp, WhatsApp is the preferred medium through which to receive messages for most users. There is also a significant increase in the share of respondents who report having read all MomConnect messages in full for both the RCM-WA (7.6 pp.) and the RCM-BCM (4.8 pp.) arms compared to the ARM arm, suggesting that the new messaging models improve user engagement.** Interestingly, the share is even higher in the RCM-SMS arm compared to the RCM-WA arm (3.8 pp.), suggesting that users might be more likely to check and fully read SMS messages compared to WhatsApp messages.[31] We find neither meaningful nor statistically significant differences in respondents' perception about the messaging frequency across messaging models. Overall, only 4% of respondents said they would have liked to receive messages at a higher frequency and this share is similar across arms.[32]

Table 12 contains several questions related to the user satisfaction of MomConnect users and across treatment arms. Generally, users are very positive about MomConnect: 93.3% are likely to recommend MomConnect to other mothers, 93% find it easy to find information on MomConnect, and 98% feel supported by MomConnect. Furthermore, 96% find appointment reminders helpful, 97% find baby growth information helpful, and 97% find baby health information helpful. Finally, 97% find MomConnect trustworthy and 76% used the internet to find health information during their pregnancy. While the RCM-WA and RCM-BCM arms lead to statistically significant improvements for some of the user satisfaction measures, the effects are small (within 1-3 pp.) and potentially not the most meaningful given the very high satisfaction rates in all four messaging models.

[31] Other possibilities are that 1) SMS messages are shorter and therefore quicker to read in full; and 2) respondents have a different understanding of what constitutes having "fully read" a SMS message compared to a WhatsApp message

[32] 73% of respondents said the messaging frequency was "just right" and 22% said they received messages "too often".

Table 10: User Experience - Most Useful MomConnect Component

	(1)	(2)	(3)	(4)	(5)
	Which part of MomConnect do you find most useful?				
Dep. Var.	Appoint-ment reminders	Pregnancy information	Baby information	Helpdesk	Other
RCM-SMS	-.023	-.022	.064	-.013	-.005
	(.03)	(.027)	(.027)	(.015)	(.013)
RCM-WA	-.063	-.016	.042	.039	-.002
	(.029)	(.027)	(.026)	(.017)	(.013)
RCM-BCM	-.055	-.029	.063	.017	.004
	(.03)	(.027)	(.027)	(.016)	(.014)
Observations	2155	2155	2155	2155	2155
Mean Dep. Var.	0.345	0.26	0.266	0.079	0.049
Mean Dep. Var. in ARM	0.377	0.281	0.226	0.067	0.049
Stratification Variable	Yes	Yes	Yes	Yes	Yes
p-value ARM = RCM-WA	.028	.557	.106	.026	.889
p-value ARM = RCM-BCM	.062	.291	.021	.296	.743
p-value RCM-WA = RCM-BCM	.79	.62	.451	.239	.643
p-value RCM-SMS = RCM-WA	.168	.809	.41	.002	.781

Notes: Respondents were asked to select which component of MomConnect they found most useful. Each outcome variable in columns (1) - (5) measures the proportion of respondents who reported finding that particular component the most useful. All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors.

Table 11: User Experience - Message preferences

Dep. Var.	(1) Prefers to receive messages on Whatsapp (1/0)	(2) Read all Mom Connect messages in full (1/0)	(3) Felt they did not receive messages frequently enough (1/0)
RCM-SMS	-.533	.114	-.007
	(.024)	(.024)	(.012)
RCM-WA	.018	.076	.003
	(.016)	(.024)	(.012)
RCM-BCM	.024	.048	-.01
	(.016)	(.026)	(.011)
Observations	2178	2157	2158
Mean Dep. Var.	0.791	0.816	0.04
Mean Dep. Var. in ARM	0.911	0.758	0.045
Stratification Variable	Yes	Yes	Yes
p-value ARM = RCM-WA	.259	.002	.807
p-value ARM = RCM-BCM	.138	.062	.359
p-value RCM-WA = RCM-BCM	.703	.23	.257
p-value RCM-SMS = RCM-WA	0	.086	.383

Notes: The binary outcome in column (1) measures whether respondents report that they would have preferred to receive MomConnect through WhatsApp. The binary outcome in column (2) measures whether respondents report to have read all MomConnect messages in full. The binary outcome in column (3) measures whether respondents report that they did not receive messages frequently enough. All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors.

Table 12: User Experience - User Satisfaction

Dep. Var.	(1) Is very likely to recom-mend Mom Connect to other mothers (1/0)	(2) Finds it very easy to find informat-ion on Mom Connect (1/0)	(3) Feels supported by Mom Connect (1/0)	(4) Finds appoint-ment reminders helpful (1/0)	(5) Finds baby growth infor-mation helpful (1/0)	(6) Finds baby health informa-tion helpful (1/0)	(7) Finds Mom Connect trust-worthy (1/0)	(8) Internet use during preg-nancy (1/0)
RCM-SMS	-.009	.005	.007	.002	.021	.014	.004	.014
	(.016)	(.016)	(.009)	(.012)	(.012)	(.012)	(.011)	(.026)
RCM-WA	.005	.014	-.002	-.001	.013	.018	.001	.028
	(.015)	(.016)	(.01)	(.013)	(.012)	(.012)	(.012)	(.025)
RCM-BCM	.004	-.001	.017	.002	.03	.024	.012	-.018
	(.015)	(.017)	(.009)	(.013)	(.011)	(.011)	(.011)	(.027)
Observations	2157	2157	2180	2180	2180	2178	2181	2182
Mean Dep. Var.	0.933	0.925	0.978	0.956	0.968	0.967	0.966	0.761
Mean Dep. Var. in ARM	0.935	0.922	0.973	0.955	0.953	0.955	0.962	0.753
Stratification Variable	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-value ARM = RCM-WA	.718	.378	.806	.938	.27	.114	.936	.261
p-value ARM = RCM-BCM	.776	.97	.053	.893	.004	.035	.269	.494
p-value RCM-WA = RCM-BCM	.945	.363	.029	.831	.061	.565	.3	.072
p-value RCM-SMS = RCM-WA	.352	.557	.304	.796	.46	.665	.79	.571

Notes: The binary indicators in columns (1) and (2) are derived from 5-point Likert scales and are coded as 1 if the mothers select the most positive ranking. The binary indicators in columns (3), (4), (5), (6), (7) are derived from 5-point Likert scales with options ranging from Not _, Slightly_, Moderately_, Very_, and Extremely_ and are coded as 1 if the mothers report "Very_" or "Extremely_". Column (8) measures whether respondents report that they used the internet to find information about their pregnancy, beyond MomConnect. All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors.

Engagement Outcomes

Table 13 assesses the impact that the different messaging models have on the number of messages sent and received by MomConnect. Columns 1 and 3 represent the frequency of messaging in number of days, with column 1 showing the number of days that a participant sent a message to MomConnect, and 3 the number of days that MomConnect sent a message to the participant. Columns 2 and 4 represent the total number of messages users sent to MomConnect and MomConnect sent to users.

All treatment groups sent messages to MomConnect on fewer days than in ARM: RCM-WA participants sent messages on 5.4 fewer days (p-value of 0.000) and RCM-BCM participants sent messages on 6.4 fewer days (p-value of 0.000), on average. Participants in RCM-SMS sent messages to MomConnect on 14.4 fewer days than participants in RCM-WA (p-value 0.000), on average. Participants sending messages on the largest number of days in ARM is most likely driven by the different structure of messaging, where for a user in the ARM group to receive the full set of

information per push reminder, they must send the message request for more on 3 consecutive days following the original push message. Whereas for any of the treatment models, they can send multiple messages on the same day and receive all relevant information. This is consistent with the results in column 3 which shows that MomConnect sent messages on 4 fewer days and 3.8 fewer days in RCM-WA and RCM-BCM, both with p-values of 0.000. MomConnect also sent 2.6 fewer messages, on average, in RCM-SMS compared to RCM-WA.

The first key result for engagement is that despite receiving messages on fewer days, the total messages users sent to MomConnect is highest for RCM-BCM with on average 6 more messages (p-value 0.017) sent than in ARM (column 2). This result is unique to RCM-BCM and does not apply to RCM-WA, for which the number of messages sent by users is not statistically different from ARM. This result is likely driven by the additional "browse" option available in RCM-BCM, which distinguishes RCM-BCM from RCM-WA. The attrition results in Table 14 for the midline could also be interpreted as a measure of engagement: participants in the RCM-WA and RCM-BCM arms were 5.4 pp. and 5.8 pp. more likely to respond to the midline survey than participants in ARM.

The second key result is that the number of messages sent by the RCM-SMS group to MomConnect is substantially less than all other groups. Importantly, the comparison with RCM-WA shows that RCM-SMS sent 67.2 fewer messages to MomConnect compared with RCM-WA (p-value 0.000), on average. This clearly distinguishes the impact that the WhatsApp platform has on engagement compared with a combination of SMS push messages and using USSD to respond.

We consider these results to be descriptive and potentially caused by the different structures of the messaging models, rather than an accurate measure of the impact of the different messaging models on engagement. Therefore, the results do not inform our recommendations. However, Reach will be conducting a more detailed analysis of engagement from the programmatic data collected on their platform that may contribute to decision-making.

Table 13: User Engagement

	(1) Number of days mother sends message(s) to MomConnect	(2) Number of messages sent to MomConnect	(3) Number of days MomConnect sends message(s) to mother	(4) Number of messages sent by MomConnect
RCM-SMS	-19.827	-66.335	-6.598	-71.037
	(.463)	(1.843)	(.488)	(2.296)
RCM-WA	-5.436	.868	-4.006	-1.623
	(.533)	(2.545)	(.53)	(3.088)
RCM-BCM	-6.446	6.023	-3.75	2.396
	(.528)	(2.522)	(.526)	(3.013)
Observations	8213	8213	8214	8214
Mean Dep. Var.	16.5	73.9	45.7	128.4
Mean Dep. Var. in ARM	24.4	88.6	49.3	145.8
Stratification Variable	Yes	Yes	Yes	Yes
p-value ARM = RCM-WA	0	.733	0	.599
p-value ARM = RCM-BCM	0	.017	0	.427
p-value RCM-WA = RCM-BCM	.013	.045	.479	.177
p-value RCM-SMS = RCM-WA	0	0	0	0

Notes: The outcomes in columns (1) and (2) measure the number of days users sent any message to MomConnect and the total number of messages users sent. The outcomes in columns (3) and (4) measure the number of days users received any message from MomConnect and the total number of messages users received. All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors.

5.3. Attrition

Since the response rate for our endline survey was far below 100%, we conduct tests for differential attrition across treatment arms in **Table 14**. Differential attrition could be caused by different messaging models leading to different levels of user engagement and more engaged users being more likely to participate in the survey. The presence of non-random differential attrition would imply that the treatment coefficients presented in sections 5.1 and 5.2 would be biased. We first estimate treatment effects on baseline survey completion. Column 1 shows that there are no large or statistically significant treatment effects on baseline survey completion, which is expected since the baseline survey took place prior to randomisation. Overall, 90.1% of study participants responded to the baseline survey. Next we estimate treatment effects on midline survey completion. Column 2 shows that there are large and statistically significant effects on midline survey completion: participants in the RCM-WA and RCM-BCM arms were 5.4 pp. (p-value 0.000) and 5.8 pp. (p.value 0.000) more likely to respond to the midline survey than

participants in the ARM arm. Participants in the RCM-SMS arm were 14.4 pp. less likely to respond to the midline survey than participants in the RCM-WA arm. We therefore do not analyse any outcomes measured at midline in this report. The overall midline completion rate is 54.4%. Finally, we estimate treatment effects on three measures of the endline survey response rate: column 3 shows that there are no large or statistically significant treatment effects on whether study participants are in the endline sample, including participants who lost their child due to a miscarriage or after giving birth.[33] Column 4 shows that there are also no large or statistically significant treatment effects on whether study participants are in the endline sample, excluding participants who lost their child (i.e., participants who actually responded to the endline survey). Finally, column 5 shows that there are no large or statistically significant treatment effects on whether study participants completed the long survey (which contained the questions on knowledge, behaviour, and user experience). **Overall, we find no evidence for differential attrition in our endline sample and therefore consider our treatment coefficients to be internally valid for our sample of study participants.**

The fact that only 63.7% of all study participants responded to our endline survey raises some questions about the external validity of the study. However, the descriptive statistics discussed in Section 4 suggest that at least for age, gestational age, and province, there are no meaningful differences between our study sample and our endline sample. Our post-stratification analysis addresses concerns regarding the external validity of our results with respect not just to our study sample, but to the overall population of MomConnect users, which differs somewhat from our study and endline samples as discussed in Section 4. Note that a comparison between this study's response rate and other phone survey response rate is difficult as phone survey response rates vary from greatly depending on factors such as whether you have numbers through a program they are enrolled in (e.g. MomConnect) or if it is a speculative call, whether you know the name of the participant, or what the level of incentivisation is. The fact that this is a mobile intervention where we had approximately 60% of the mothers' names makes it difficult to find a similar comparison against which to benchmark this response rate.

[33] We did not administer the endline survey to women who lost their child due to ethical considerations. Overall, the mortality rate in our endline sample was 3.8% (incl. miscarriages). We do not find statistically significant effects across messaging models on mortality and all estimated treatment coefficients are lower than 0.01.

Table 14: Attrition

Survey Round	(1) Baseline	(2) Midline	(3)	(4) Endline	(5)
Dep. Var.	Completed baseline survey = {1,0}	Completed midline survey = {1,0}	Completed endline survey = {1,0}	Completed endline call = {1,0}	Completed long endline survey = {1,0}
RCM-SMS	.005 (.009)	-.086 (.016)	-.003 (.015)	-.002 (.015)	-.001 (.015)
RCM-WA	-.004 (.009)	.054 (.015)	.004 (.015)	-.003 (.015)	0 (.015)
RCM-BCM	-.002 (.009)	.058 (.015)	.004 (.015)	.009 (.015)	-.003 (.015)
Observations	8232	8232	8232	8232	8232
Mean Dep. Var.	0.901	0.544	0.662	0.637	0.42
Mean Dep. Var. in ARM	0.902	0.537	0.662	0.636	0.421
Stratification Variable	Yes	Yes	Yes	Yes	Yes
p-value ARM = RCM-WA	.659	0	.794	.854	.984
p-value ARM = RCM-BCM	.824	0	.79	.536	.855
p-value RCM-WA = RCM-BCM	.827	.817	.995	.422	.839
p-value RCM-SMS = RCM-WA	.34	0	.648	.963	.938

Notes: The outcomes in this table are binary indicators for whether a study partcipant who enrolled in the study completed the respective survey described in the column heading. All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors.

5.4. Post-Stratification Weights

Since our sample statistically differs in terms of age, gestational age, and province from the population of MomConnect users in 2023 (see section 4), we apply post-stratification weights to our main results. This exercise strengthens the external validity of the analysis, trying to understand better what would happen if the different messaging models were scaled up for all MomConnect users and not just the type of users that were in our specific endline sample. We apply post-stratification weights to our two primary outcomes related to ANC visits and immunisations as well as the knowledge and behaviour indices.

Table 15 shows the treatment coefficients using post-stratification weights. Column 1 shows the coefficients for the ANC outcome. Comparing them with column 1 in Table 3, we see that the coefficients have a very similar magnitude and remain statistically insignificant at conventional levels. Column 2 shows the coefficients for the immunisation outcome. Comparing them with

column 1 in Table 5, we still find a similarly negative coefficient for the RCM-BCM arm (-0.029 without post-stratification weights and -0.032 with post-stratification weights), which remains statistically significant at the 5% level. As in column 1 in Table 5, we fail to reject any of the other three null hypotheses related to the immunisation outcome. Column 3 shows the coefficients for the knowledge index. Compared to column 1 in Table 7, the coefficient for the RCM-BCM arm is considerably higher (0.115 without post-stratification weights and 0.173 with post-stratification weights) and has become statistically significant at the 10% level. The other coefficients are similar to the unweighted ones and the difference between the RCM-WA and RCM-SMS arm has a similar magnitude and is still significant at the 5% level. Finally, column 4 shows the coefficients for the behaviour index. Compared to column 3 in Table 7, the coefficients of both the RCM-WA (0.216 without weights and 0.264 with weights) and RCM-BCM (0.239 without weights and 0.332 with weights) arms are considerably higher and their p-values are lower (for RCM-WA, the p-value reduces from 0.1 to 0.058 and for RCM-BCM the p-value reduces from 0.073 to 0.018). Overall, the results with post-stratification weights tell the same story as the results without post-stratification weights: there is a slight reduction in full immunisation coverage in RCM-BCM compared to ARM, and modest gains in secondary outcomes for both RCM-WA and RCM-BCM compared to ARM.

Table 15: Treatment Coefficiants of Core Primary and Secondary Outcomes using Post-Stratification Weights

	(1)	(2)	(3)	(4)
Dep. Var.	At least 8 ANC Visits	All 6 Vaccinations	Knowledge Index	Behaviour Index
RCM-SMS	.008	-.014	-.075	.048
	(.021)	(.012)	(.1)	(.145)
RCM-WA	.012	-.016	.118	.264
	(.021)	(.012)	(.099)	(.139)
RCM-BCM	-.013	-.032	.173	.332
	(.021)	(.013)	(.101)	(.14)
Observations	5142	4985	2178	2178
Mean Dep. Var.	0.436	0.894	5.222	11.485
Mean Dep. Var. in ARM	0.436	0.908	5.175	11.357
Stratification Variable	Yes	Yes	Yes	Yes
p-value ARM = RCM-WA	.566	.187	.237	.058
p-value ARM = RCM-BCM	.526	.013	.088	.018
p-value RCM-WA = RCM-BCM	.226	.248	.573	.613
p-value RCM-SMS = RCM-WA	.871	.879	.044	.122

Notes: All specifcations in this table apply post-stratification weights for the 54 randomisation strata by using the total number of 2023 MomConnect users in each stratum to construct weights. All specifications regress the outcome variable on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors.

5.5. Multiple Hypothesis Testing Adjustments

Table 16 reports sharpened q-values, which should be interpreted as adjusted p-values correcting for false discovery rates in the two primary outcome regressions as discussed in section 3.4. Using this more conservative approach to causal inference, we see that all sharpened q-values except for the RCM-BCM vs. ARM comparison for the immunisation outcome become 1. The sharpened q-value for the RCM-BCM vs. ARM comparison for the immunisation outcome is 0.16 whereas the unadjusted p-value is 0.018. Thus, the negative result of the RCM-BCM arm on immunisations is not robust to the multiple hypothesis testing correction we had prespecified. However, it is important to note that while relying on p-values when testing many different hypotheses leads to the over-rejection of null hypotheses, multiple hypothesis testing corrections tend to be too

conservative and lead to an under-rejection of null hypotheses.[34] This presents a trade-off: applying multiple hypothesis testing corrections reduces the likelihood of falsely rejecting a null hypothesis, but increases the likelihood of failing to reject a null hypothesis although it is not true. We prefer to take a conservative approach in how the different p-values inform our recommendation: we are ultimately not certain whether or not there is a real reduction in immunisation coverage resulting due to the RCM-BCM model, but there may be, so we think there would be too much risk to recommend scaling up the RCM-BCM model. Therefore, we make a soft recommendation for scaling up the RCM-WA model. Alternatively, if Reach had strong priors on the RCM-BCM model being preferable, it would make sense to conduct further research to understand the potentially negative impact of RCM-BCM on immunisations better.

Table 16: Comparison of primary outcomes' p-values and q-values with multiple hypothesis adjustments

Dependant Variable:	(1) At least 8 ANC visits	(2) All 6 Vaccinations
RCM-SMS Coefficient	.007	-.017
p-value RCM-SMS = RCM-WA	.986	.882
sharpened q-value RCM-SMS = RCM-WA	1	1
RCM-WA Coefficient	.007	-.015
p-value RCM-WA = ARM	.719	.194
sharpened q-value RCM-WA = ARM	1	1
RCM-BCM Coefficient	-.015	-.029
p-value RCM-BCM = ARM	.429	.018
sharpened q-value RCM-BCM = ARM	1	.169
p-value RCM-BCM = RCM-WA	.249	.289
sharpened q-value RCM-BCM = RCM-WA	1	1

Notes: The coefficients and p-values in this table are estimated by regressing the outcomes in the column headings on treatment indicators and stratification variables and use heteroskedasticity-robust standard errors. The sharpened q-values are then estimated through the procedure described in Anderson (2008), applied to the eight pre-specified main hypotheses of the study shown in the table.

[34] This post on the World Bank Development Impact Blog has an accessible explanation of Multiple Hypothesis Testing adjustment methods, including the one we are using.

6. Discussion

For ANC visits, we find no significant differences between the ARM, RCM-WA, and RCM-BCM models or the RCM-SMS and RCM-WA models and can rule out effects of more than +/- 5 pp. with 95% confidence. For immunisations, we find that the RCM-BCM model leads to a 3 percentage point decrease in full immunisation coverage compared to the status quo ARM model (significant at the 5% level, but not significant when applying our prespecified multiple hypothesis correction)**.** For knowledge, we find that both the RCM-WA and RCM-BCM models have positive effects of ~0.07 standard deviations compared to the ARM model, but both coefficients are insignificant at conventional levels (p-values 0.22 and 0.24). We do find a significant (at the 5% level) increase in knowledge of ~0.13 standard deviations for users of the RCM-WA model compared to the RCM-SMS model. For behaviour, we find that both the RCM-WA and RCM-BCM models have positive and significant (at the 10% level) effects of ~0.1 standard deviations compared to the ARM model. We also find that RCM-WA and RCM-BCM models lead to a higher perceived usefulness of the baby information component by 4.2 and 6.3 percentage points, respectively.

Since both the RCM-WA and RCM-BCM models had modestly positive effects on behaviour compared to the status quo ARM model, but the RCM-BCM model led to a 3 pp. decrease in full immunisation coverage, our overall recommendation is to scale up the RCM-WA model as the new default model for MomConnect. However, the effect sizes we find are generally modest. If switching costs from the ARM model to RCM-WA model were high, it might be more cost-effective to keep the ARM model. Since we find it difficult to explain the negative effect of the RCM-BCM model and the negative effect is not robust to applying our multiple hypothesis testing correcting, additional, potentially qualitative, data on users' experience with the RCM-BCM model could be helpful for understanding this result better. It is also important to note that while the difference in full immunisation rates was statistically significant between the ARM model and the RCM-BCM model, there was no statistically significant difference between the RCM-WA model and the RCM-BCM model. Therefore, and considering the overall modest differences across messaging models, our recommendation in favour of the RCM-WA model should be interpreted as a "soft" recommendation.

A key finding of the study is that delivering content via SMS is not substantially less effective than delivering content via WhatsApp as shown by the small and mostly non-statistically significant differences between the RCM-SMS and RCM-WA messaging models and the preference of a majority of SMS users to receive messages through SMS rather than WhatsApp. The study shows that the MomConnect model has the potential to work effectively in lower-resource settings where smartphone ownership levels are low and therefore messages could only be delivered via SMS. This is especially important because the marginal value of the maternal and infant health information delivered through MomConnect is likely highest in these lower-resource settings. However, the RCM-SMS arm is 4-6x more expensive than the RCM-WA arm and the current SMS Core delivered by MomConnect, making it infeasible to implement the same approach in real-world settings. **We believe that studying how to adapt the relevant content from this study to SMS is a fruitful area for further research given Reach's expansion plans for MomConnect, keeping in**

mind that sending more SMS messages entails a trade-off between more comprehensive content and higher cost which is not present to the same extent with the WhatsApp model.

It should be noted that the results of the study only capture the relative impact of the different messaging models, but not the impact of receiving MomConnect messages compared to receiving no messages. Given the difficulty of randomising access to MomConnect in South Africa because of MomConnect's scale – and because it is likely no longer ethical to deny access to the platform – an impact evaluation of MomConnect as a whole would need to take place in future expansion countries where rollout could be randomised, e.g. across health facilities. In order to get some suggestive evidence on the platform's overall impact, one might want to compare primary outcomes measured in this with the latest DHIS2 data from the NDOH. However, MomConnect users may be systematically different from non-MomConnect users, making such direct comparisons difficult.

The study has some limitations. First, while the response rate of 63.7% is similar, or slightly higher, than what we anticipated based on previous IDinsight phone surveys, it poses a potential threat to extrapolating the results from the specific study sample to the population of all MomConnect users. It is unclear whether attrition was random and therefore whether study participants we successfully reached at endline are comparable with those we did not reach. For example, it seems plausible that the participants who answered their phone or agreed to participate in the study conditional on reaching them were more engaged with MomConnect than study participants who did not participate in the endline survey. The relative effectiveness of the different messaging models may be different across these different subsets of participants. It is important to note that we did not find any differential attrition across treatment arms and that we can therefore be confident in the internal validity of our estimates, i.e., that our estimates are unbiased for our study sample.

Second, the recruitment strategy of the study led to a study sample that had a substantially higher gestational age at the time of registration (which often coincides with the timing of their first ANC visit), a slightly different average age, and some differences in the distribution of participants across provinces compared to the overall population of MomConnect users. Women who register for MomConnect (and have their first ANC visit) later may be different from women who register earlier in many ways which may affect the relative effectiveness of the different messaging models. The most obvious difference is that women who register later are exposed to the platform for a shorter time period and therefore differences across messaging models may be less likely to manifest. Similarly to the first limitation, differences in the study sample and the population potentially affect the external validity, but not the internal validity of the study. Similarly, women living in different provinces in South Africa likely differ in important ways from each other. To mitigate this concern, we report the robustness of our results to applying post-stratification weights in section 5.4 and reach the same conclusions as we do when not applying post-stratification weights.

Third, all our outcomes rely on self-reported survey data. Self-reporting can lead to random (i.e., uncorrelated to treatment) measurement error, e.g. if respondents face difficulties with recall of past events, which leads to attenuation bias and therefore an underestimation of treatment

effects. Our data allowed us to conduct sensitivity checks for immunisation outcomes, since we asked respondents whether they have access to their baby's vaccination cards during the endline interview. This enabled us to compare treatment effects in the sample of respondents who did and did not have access to their baby's vaccination card and we find minimal discrepancies between the estimated coefficients across the two groups of respondents. This suggests that respondent recall did not lead to significant measurement error compared to respondents who reported information based on the vaccination card. However, the lack of evidence for significant measurement error in the immunisation outcomes measured through recall does not imply that there is no measurement error in other outcomes measured through respondent recall. For example, self-reported ANC visits may suffer from higher measurement error than vaccinations because they happened further in the past at the time of the interview. We attempted to mitigate this concern by employing a standardised way of asking questions to respondents and frequently re-training our enumerators on how best to ask questions to enable accurate respondent recall.[35] Readers who are concerned about measurement error in some of our outcome variables may want to interpret estimated treatment coefficients as lower-bound estimates of true treatment effects.

In conclusion, this study finds moderate differences between the three new messaging models being tested on MomConnect. The core recommendations from the study are:

1) **Scale up RCM-WA**, as it has positive impact on adoption of healthy behaviours without any negative impacts on primary outcomes (although note that this is a "soft" recommendation due to the multiple hypothesis correction indicating the negative result of RCM-BCM on immunizations may be insignificant)
2) **Adapt the MomConnect SMS Core messaging model** to include some elements of the RCM model
3) **Consider conducting further qualitative investigations** to understand what may be driving some of the results, or lack of results, presented
4) **Consider conducting an impact evaluation of the MomConnect platform as a whole**, as it is rolled out to a new country where non-users of MomConnect can serve as a counterfactual control group
5) **Consider strengthening health facility adherence to registering women on MomConnect** during ANC visits in health facilities outside of Gauteng (as mentioned in Section 4)

[35] For example, for ANC visits mothers were asked at what month they found out they were pregnant, how many times they thought they visited the ANC clinic each month on average, and whether that increased at any stage of the pregnancy (e.g. the final month). The enumerator then did the calculation for this live with them on the interview and asked to confirm whether the final number seemed correct. Note that participant's allocation to enumerators was stratified and enumerator fixed-effects were included in regressions to control for different ways in which enumerators may have facilitated the questionnaire..

7. Bibliography

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

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

Gupta, Sandeep K. “Intention-to-treat concept: A review.” Perspectives in clinical research vol. 2,3 (2011): 109-12. doi:10.4103/2229-3485.83221

Hlongwane, Tsakane Mag et al. “Implementing antenatal care recommendations, South Africa.” Bulletin of the World Health Organization vol. 99,3 (2021): 220-227. doi:10.2471/BLT.20.278945

Mckenzie, David. 2021. “An updated overv iew of multiple hypothesis testing commands in Stata.” World Bank Blogs. https://blogs.worldbank.org/en/impactevaluations/updated-overview-multiple-hypothesis-testing-commands-stata

NDOH and ICF. 2019. South Africa Demographic and Health Survey 2016. Pretoria: National Department of Health - NDoH - ICF. Available at http://dhsprogram.com/pubs/pdf/FR337/FR337.pdf.

StatsSA. “P0305 - Recorded live births”. De;artment of Statistics South Africa. https://www.statssa.gov.za/?p=16902

8. Appendices

Appendix 1A: Baseline Survey Questionnaire

0. Introduction & Consent	
0.01	As part of the research study, we'd like to learn more about your pregnancy experience before you start getting messages from MomConnect. There are 13 quick questions to answer which will take about 5 minutes. When you're done, we'll send you a R5 airtime voucher. *Do you want to answer the questions?* *1* - Yes, answer now *2* - Answer later *3* - No, I don't want to answer
1. Healthy Behaviors	
1.01	Question 1 of 13 *Do you plan to breastfeed your baby after birth?* *1* - Yes *2* - No *0* - Skip this question
1.02	Question 2 of 13 *How long do you plan to give your baby only breastmilk before giving other foods and water?* *1* - Between 0 and 3 months *2* - Between 4 and 5 months *3* - For 6 months *4* - Longer than 6 months *5* - I don't want to only breastfeed *6* - I don't know *0* - Skip this question

1.03	Question 3 of 13 *What do you think about this statement?* "I think it is important to vaccinate my baby against severe diseases like measles, polio, and tetanus." *1* - I strongly agree *2* - I agree *3* - I don't agree or disagree *4* - I disagree *5* - I strongly disagree *0* - Skip this question
1.04	Question 4 of 13 *What do you think about this statement?* "The benefits of protecting my child against diseases like measles, tetanus, and polio outweigh the risks of my child developing a serious side effect from the vaccine" *1* - I strongly agree *2* - I agree *3* - I don't agree or disagree *4* - I disagree *5* - I strongly disagree *0* - Skip this question
1.05	Question 5 of 13 *How often do you plan to go to the clinic for a a check-up during this pregnancy?* *1* - More than once a month *2* - Once a month *3* - Once every 2 to 3 months *4* - Once every 4 to 5 months *5* - Once every 6 to 9 months *6* - Never *0* - Skip this question
1.06	Question 6 of 13 *Since becoming pregnant, do you eat vegetables at least once a week?* *1* - Yes *2* - No *3* - Skip this question

1.07	Question 7 of 13 *Since becoming pregnant, do you eat fruit at least once a week?* *1* - Yes *2* - No *0* - Skip this question
1.08	Question 8 of 13 *Since becoming pregnant, do you eat milk, maas, hard cheese or yogurt at least once a week?* *1* - Yes *2* - No *0* - Skip this question
1.09	Question 9 of 13 How frequently do you eat Liver? *1* - 2-3 times a week *2* - Once a week *3* - Once a month *4* - Less than once a month *5* - Never *0* - Skip this question
1.1	Question 10 of 13 *In your opinion, what is the biggest danger sign in pregnancy from this list?* *1* - Weight gain of 4-5 kilograms *2* - Vaginal bleeding *3* - Nose bleeds *0* - Skip this question
1.11	Question 11 of 13 *In your opinion, what is the biggest danger sign in pregnancy from this list?* *1* - Swollen feet and legs even after sleep *2* - Bloating *3* - Gas *0* - Skip this question
2. Demographics	

2.01	Question 12 of 13 *What is your marital status?* *1* - Never married/single *2* - Married *3* - Separated or divorced *4* - Widowed *5* - Have a partner or boyfriend *0* - Skip this question
2.02	Question 13 of 13 *Which answer best describes your highest level of education?* *1* - I didn't finish primary school *2* - I finished Grade 7 *3* - I finished Grade 12 *4* - I finished a certificate or degree after high school *0* - Skip this question
End	
	Thank you for answering these questions. Your R5 airtime will be sent to you within the next 48 hours. You'll get your first weekly message from MomConnect soon. Or you can start with any of these: *1 - FIND* information 🔎 *2 - UPDATE* your personal info ⚙️ *3 - ASK* a question 👩🏽 *4 - PRIVACY* policy 🔐 *5 - STOP* getting messages ⛔

Appendix 1B: Midline Survey Questionnaire

0. Introduction & Consent	
0.01	Hi [name] Thank you again for being part of the research study that will help us make MomConnect better. We have a few more questions to ask that won't take more than 3 minutes to answer. When you're done, we'll send you a R5 airtime voucher. *Do you want to answer the questions?* *1* - Yes, answer now *2* - Answer later *3* - No, I don't want to answer
1. Adoption & Knowledge of Healthy Behaviors - Generic	
1.01	Question 1 of 16 *Do you eat fruits at least once a day?* *1.* Yes *2.* No *0.* Skip this question
1.02	Question 2 of 16 *Do you eat vegetables at least once a day?* *1.* Yes *2.* No *0.* Skip this question
1.03	Question 3 of 16 *How frequently do you eat liver?* *1.* - Once a week *2.* - Once every 2 weeks *3.* - Once a month *4.* - Less frequently than once a month *5.* - Not at all *6.* Skip this question

1.04	Question 4 of 16 *Nuts, eggs, meat, fish, and green vegetables have a lot of what in them?* *1.* - Calcium *2.* - Vitamin C *3.* - Iron *4.* - Fibre *0.* - Skip this question
1.05	Question 5 of 16 *Since becoming pregnant, would you say the number of alcoholic drinks you have per week has:* *1.* - Stayed the same *2.* - Reduced *3.* - Increased *4.* - Stopped *5.* - I never drink *0.* - Skip this question
1.06	Question 6 of 16 *What can severe swollen feet even after a night's sleep be a symptom of?* *1.* - Urinary tract infection *2.* - Pre-eclampsia *3.* - Anemia *0.* - Skip this question
1.07	Question 7 of 16 What could a mix of feeling dizzy and weak/tired be a symptom of? *1.* - Urinary tract infection *2.* - Pre-eclampsia *3.* - Anemia *0.* - Skip this question

1.08	Question 8 of 16 *Do you think baby kicks should be felt every day in the third trimester of pregnancy?* *1.* Yes *2.* Maybe *3.* No *4.* Don't know *5.* Skip this question
1.09	Question 9 of 16 *In your opinion, what is the biggest danger sign in pregnancy from this list?* *1* - Weight gain of 4-5 kilograms *2* - Vaginal bleeding *3* - Nose bleeds *4* - Skip this question
2. Adoption & Knowledge of Healthy Behaviors - Breastfeeding	
2.01	Question 10 of 16 *Are you planning on breastfeeding your baby after he/she is born?* *1.* Yes *2.* No *0.* Skip this question
2.02	Question 11 of 16 *What is the biggest reason why you want to breastfeed your baby?* *1.* Breastmilk boosts my baby's immunity *2.* Breastmilk is tastier than formula *3.* Breastfeeding improves my health *4.* I was told to breastfeed *5.* Other *0.* Skip
2.03	Question 12 of 16 *When do you plan to start breastfeeding your baby?* *1.* Within 1 hour of birth *2.* After 1 hour post-delivery *3.* Day 2 & above

	4. Undecided *0.* Skip this question
2.04	Question 13 of 16 *How long do you plan to give your baby only breastmilk before giving other foods and water?* *1* - Between 0 and 3 months *2* - Between 4 and 5 months *3* - For 6 months *4* - Longer than 6 months *5* - I don't want to only breastfeed *6* - I don't know *0* - Skip this question
2.05	Question 11 of 16 *What is the biggest reason why you don't intend on breastfeeding your baby after he/she is born?* *1.* Breastmilk is not nutritious *2.* Low milk supply *3.* Sore nipples *4.* Takes too long *5.* Lack of information *6.* Lack of support *7.* Other 0 - Skip
3. Adoption & Knowledge of Healthy Behaviors - Vaccination	
3.01	Question 14 of 16 *What do you think about this statement?* "I think it is important to vaccinate my baby against severe diseases like measles, polio, and tetanus." *1* - I strongly agree *2* - I agree *3* - I don't agree or disagree *4* - I disagree *5* - I strongly disagree *0* - Skip this question

3.02	Question 15 of 16 *What do you think about this statement?* "The benefits of vaccinating my child outweighs the risks my child will develop side effects from them" *1* - I strongly agree *2* - I agree *3* - I don't agree or disagree *4* - I disagree *5* - I strongly disagree *0* - Skip this question
3.03	Question 16 of 16 *How likely are you to follow the recommended shot schedule for your child?* *1.* - Very likely *2.* - Unlikely *3.* - Not sure *4.* - Likely *5.* - Very likely *0.* - Skip this question
End	
	Thank you for answering these questions. Your R5 airtime will be sent to you within the next 24 hours. You will get your next MomConnect message soon. Have a lovely day 🌸

Appendix 1C: Endline Survey Questionnaire

reached > Speaking to Mother > Consenting Survey > ~			
reached > Speaking to Mother > Consenting Survey > ~ > ~			
A.46 I would first like to speak to you about your experience with antenatal care during your pregnancy			
A.47 Did you give birth naturally or via caesarian section?		1	Naturally
		2	Caesarian
		-888	Refuse
A.48 Did you attend any clinic visits for antenatal care during this pregnancy?		1	Yes
		2	No
		-888	Refuse
A.49 Do you have your pregnancy card easily available? *I am happy to wait if you need some time to look for it* *The pregnancy card has records of ANC visits, please ask mothers to use it if they have it.*		1	Yes
		2	No
		-888	Refuse
A.50 What day was your first antenatal clinic visit for this pregnancy? *Hint: You can check the date of your first ANC appointment on your pregnancy card*			
A.51 How many weeks pregnant were you when you first went for a antenatal care clinic visit for this pregnancy? *Hint: You can check the date of your first ANC appointment on your pregnancy card*		1	1-4 weeks (1 month)
		2	5-8 weeks (2 months)
		3	9-12 weeks (3 months)
		4	13-16 weeks (4 months)
		5	17-20 weeks (5 months)
		6	21-24 weeks (6 months)
		7	25-28 weeks (7 months)
		8	29-32 weeks (8 months)
		9	33-36 weeks (9 months)
		-999	Don't remember
		-888	Refuse
A.52 How many times did you visit the clinic for antenatal care during this pregnancy? *Hint: You can check the number of times your pregnancy card was filled out by counting the number of dates for which information has been filled out for*			
reached > Speaking to Mother > Consenting Survey > ~ > vac_healthy_behaviors			
A.53 Thanks! I will now ask you about your experience with vaccinations.			
A. 54 First, did your babies receive the same vaccinations on the same dates?		1	Yes
		2	No
		-999	Don't know/unsure
A. 54A OK great, so for the following questions about vaccinations, please just answer for the treatment received by both babies			
A. 54B OK, thanks for letting me know. For the following questions about vaccination, please answer for your first born baby			

A. 54C OK, thanks for letting me know. For the following questions about vaccination, please answer for your last born baby			
A.55 Do you have your Road to Health booklet or another document where your newborn baby's vaccinations are written down nearby? *I am happy to wait if you need some time to look for it* *If the mom says she has her RTH booklet at home and asks to call back later, save the survey and come back to it later after making an appointment. Hint: the vaccination status is on page 27 of the RTH booklet*		1	Yes
		2	No
		-888	Refuse
A.56 Was your baby vaccinated at birth? *Hint: the vaccination status is on page 27 of the RTH booklet*		1	Yes
		2	No
		-888	Refuse
A.57 Which vaccinations did your baby receive at birth?		1	BCG (tuberculosis vaccine)
		2	OPV (oral polio vaccine)
		-888	Refuse
A.60 Has your baby ever received a BCG (tuberculosis vaccine), that is an injection in the arm or shoulder that usually causes a scar, normally given at birth?		1	Yes
		2	No
		-999	Don't know/ don't remember
		-888	Refuse
A.61 Has your baby ever received a OPV (oral polio vaccine), that is, about two drops in the mouth to prevent polio, typically given at birth?		1	Yes
		2	No
		-999	Don't know/ don't remember
		-888	Refuse
A.62 Was your baby vaccinated at 6 weeks after birth?		1	Yes
		2	No
		-999	Don't know/ don't remember
		-888	Refuse
A.63 Which vaccinations did your baby receive at 6 weeks after birth?		1	OPV (oral polio vaccine)
		2	RV (rotavirus vaccine)
		3	Hexavalent / DTaP-IPV-Hib-HBV (diphtheria-tetanus-acellular pertussis-injectable polio-Haemophilus influenza b-Hepatitis B vaccine)
		4	PCV (pneumococcal conjugate vaccine)
		0	None of the above
		-888	Refuse
A.68 Has your baby ever received a OPV (oral polio vaccine), that is, about two drops in the mouth to prevent polio, typically given six weeks after birth?		1	Yes
		2	No
		-999	Don't know/ don't remember
		-888	Refuse
A.69 Has your baby ever received an RV (rotavirus vaccine), that is, syrup in the mouth to prevent diarrhea, typically given about six weeks after birth?		1	Yes
		2	No
		-999	Don't know/ don't remember
		-888	Refuse
A.70 Has your baby ever received a DTP-combination vaccination, also known as a Hexavalent vaccination? That is, an injection given in the left thigh or left arm to prevent		1	Yes
		2	No

	Diptheria, tetanus, and whooping cough, normally given about six weeks after birth?		-999	Don't know/ don't remember
			-888	Refuse
	A.71 Has your baby ever received a pneumococcal vaccination, that is, an injection in the right thigh to prevent pneumonia, normally given about six weeks after birth?		1	Yes
			2	No
			-999	Don't know/ don't remember
			-888	Refuse
	A.64 Were the 6 weeks vaccines all given on the same date?		1	Yes
			2	No
			-888	Refuse
	A.64A What date were these vaccines administered on?			
	A.64B What date was OPV (oral polio vaccine) administered on?			
	A.64C What date was RV (rotavirus vaccine) administered on?			
	A.64D What date was Hexavalent / DTaP-IPV-Hib-HBV (diphtheria-tetanus-acellular pertussis-injectable polio-Haemophilus influenza b-Hepatitis B vaccine) administered on?			
	A.64E What date was PCV (pneumococcal conjugate vaccine) administered on?			
	A.64F How many weeks old was your baby when these vaccines were administered?			
	A.64G How many weeks old was your baby when OPV (oral polio vaccine) was administered? *That is, about two drops in the mouth to prevent polio*			
	A.64H How many weeks old was your baby when RV (rotavirus vaccine) was administered? *That is, syrup in the mouth to prevent diarrhea*			
	A.64I How many weeks old was your baby when Hexavalent / DTaP-IPV-Hib-HBV (diphtheria-tetanus-acellular pertussis-injectable polio-Haemophilus influenza b-Hepatitis B vaccine) was administered? *That is, an injection given in the left thigh or left arm to prevent Diptheria, tetanus, and whooping cough*			
	A.65J How many weeks old was your baby when PCV (pneumococcal conjugate vaccine) was administered? *That is, an injection in the right thigh to prevent pneumonia*			
	A.66 Were the 6 weeks vaccines all given at the same location? *NOTE TO ENUMERATOR: By "location", we mean physical location such as a hospital, at home, a health clinic, etc*		1	Yes
			2	No
			-888	Refuse
	A.67 Where were these vaccinations administered?		1	Government hospital
			2	Government health center
			3	Government mobile clinic
			4	At home by government community health worker
			5	Private hospital
			6	Private clinic
			7	Private mobile clinic
			8	At home by private doctor
			9	Pharmacy
			10	NGO hospital
			11	NGO clinic
			12	Vaccination Campaign
			99	Other

			-888	Refuse
			-999	Don't Know
	A.67A Where was OPV (oral polio vaccine) administered?		1	Government hospital
			2	Government health center
			3	Government mobile clinic
			4	At home by government community health worker
			5	Private hospital
			6	Private clinic
			7	Private mobile clinic
			8	At home by private doctor
			9	Pharmacy
			10	NGO hospital
			11	NGO clinic
			12	Vaccination Campaign
			99	Other
			-888	Refuse
			-999	Don't Know
	A.67B Where was RV (rotavirus vaccine) administered?		1	Government hospital
			2	Government health center
			3	Government mobile clinic
			4	At home by government community health worker
			5	Private hospital
			6	Private clinic
			7	Private mobile clinic
			8	At home by private doctor
			9	Pharmacy
			10	NGO hospital
			11	NGO clinic
			12	Vaccination Campaign
			99	Other
			-888	Refuse
			-999	Don't Know
	A.67C Where was Hexavalent / DTaP-IPV-Hib-HBV (diphtheria-tetanus-acellular pertussis-injectable polio-Haemophilus influenza b-Hepatitis B vaccine) administered?		1	Government hospital
			2	Government health center
			3	Government mobile clinic
			4	At home by government community health worker
			5	Private hospital
			6	Private clinic
			7	Private mobile clinic
			8	At home by private doctor
			9	Pharmacy
			10	NGO hospital
			11	NGO clinic
			12	Vaccination Campaign
			99	Other
			-888	Refuse
			-999	Don't Know

A.67D Where was PCV (pneumococcal conjugate vaccine) administered?		1	Government hospital
		2	Government health center
		3	Government mobile clinic
		4	At home by government community health worker
		5	Private hospital
		6	Private clinic
		7	Private mobile clinic
		8	At home by private doctor
		9	Pharmacy
		10	NGO hospital
		11	NGO clinic
		12	Vaccination Campaign
		99	Other
		-888	Refuse
		-999	Don't Know
A.72 What is the main reason your child does not have all of these vaccines (BCG, OPV, RV, DTP-combination, PCV, Measles) *Hint to enumerator: please probe to ask what the most crucial reason is and select one*		1	Clinic is too far
		2	Clinic takes too long
		3	I want to do it when my child is older
		4	I don't think vaccines are safe
		5	I don't think vaccines are important
		6	My husband said not to
		7	My baby was unwell on the appointment date
		8	I was worried about COVID-19
		9	I am still not sure and deciding if I want to vaccinate my child
		10	I don't want to take my baby out of the house
		11	I have not been able to take my baby to a clinic because of transportation costs or work/ school schedules
		12	Family (including husband/ partner) said no
		99	Other
		-888	Refuse
A.72A OK great, and very quickly, can you please tell me - how did MomConnect send you messages?		1	SMS
		2	WhatsApp
		-888	Refuse
A.73 Thank you for answering those questions! I am now going to ask you about some of your experiences during your pregnancy/motherhood journey and some thoughts you might have.			
A.74 How often did you eat vegetables during pregnancy?		1	Once a day
		2	4-6 times a week
		3	2 - 3 times a week
		4	Once a week
		5	Once every two weeks
		6	Once a month
		7	Less frequently than once a month
		-888	Refuse
A.75 How often did you eat fruits during pregnancy?		1	Once a day

			2	4-6 times a week
			3	2 - 3 times a week
			4	Once a week
			5	Once every two weeks
			6	Once a month
			7	Less frequently than once a month
			-888	Refuse
	A.76 How often did you eat liver during pregnancy? *Hint : Enter the category that bes fits the mom's answer E.g "Never" would be recorded as "Less frequently than once a month"*		1	Once a day
			2	4-6 times a week
			3	2 - 3 times a week
			4	Once a week
			5	Once every two weeks
			6	Once a month
			7	Less frequently than once a month
			-888	Refuse
	A.77 How often did you drink alcohol during your pregnancy?		1	Never drank
			2	Once or more times
			3	At least once per month
			4	At least once per week
			-888	Refuse
	A.78 Before going to the hospital to give birth, did you pack a small bag with things you would need during your delivery?		1	Yes
			2	No
			-999	Don't know/ don't remember
			-888	Refuse
	A.79 What did you pack in your bag? *DO NOT READ ALOUD: You can select all options if the mother mentions at least one of the items in the option list*		1	Maternity record from the clinic
			2	A towel, facecloth, or toiletries
			3	Baby wipes, nappies, or bum cream
			4	Clean clothes for yourself
			5	Clean clothes for the baby
			6	A blanket, hat, and socks for the baby
			-999	Don't remember
			-888	Refuse
	A.80 Do you agree with the following statement: "It is safe to put a baby to sleep on their tummy."		1	Agree
			2	Disagree
			-999	Don't know/ not sure
			-888	Refuse
	A.81 Do you agree with the following statement: "It is safe to put a loose blanket over the baby to keep baby warm when sleeping."		1	Agree
			2	Disagree
			-999	Don't know/ not sure
			-888	Refuse
	A.82 What are some danger signs that would make you want to take your baby to see a doctor? Select all that apply *FOR ENUMERATORS: Please do not read out answer options*		1	Baby doesn't drink; has difficulty eating or is vomiting
			2	Baby has difficulty breathing
			3	Baby has a high fever
			4	Baby is not moving/ is very slow
			5	Baby shakes/ has seizures or fits
			6	Baby has yellow eyes (Jaundice)

	-999	Don't know	
	-888	Refuse	
A.83 Thanks for answering those questions! I now have a few questions to ask you about feeding your baby.			
A.84 What are you currently feeding your baby? *DO NOT READ ALOUD: If mom says that she is giving breastmilk plus allowing a few sips of water, that would still count as only breastfeeding.*		1	Breast milk only
		2	Combination of breast milk and formula
		3	Breast milk/ formula with solids
		4	Formula only
		99	Other
		-888	Refuse
A.85 How long after birth did you first attempt to breastfeed?		1	Within one hour of birth
		2	More than one 1 hour after birth, but less than 2 days
		3	On or after day 2
		4	Never attempted to breastfeed
		-999	Don't remember
		-888	Refuse
A.86 For how long do you intend to give your baby only breastmilk? This means no other food or water		1	Less than 6 months
		2	For 6 months
		3	Longer than 6 months
		4	I don't intend to exclusively breastfeed - I might breastfeed and feed the baby water or other food
		-999	Don't know
		-888	Refuse
A.87 When do you intend to introduce solid food to your baby? *If respondent answers "6 months" option 3 should be selected.*		1	Before 4 months
		2	4-6 months
		3	6 months or above
		-999	Don't know
		-888	Refuse
A.88 Do you agree with the following statement: "For as long as I am breastfeeding my baby, I cannot get pregnant."		1	Agree
		2	Disagree
		-999	Don't know/ not sure
		-888	Refuse
A.89 Thank you for your patience, we are more than three-quarters through the survey! I now have a few questions to ask you about anaemia. *FOR ENUMERATORS: If respondent does not know what anemia is then please explain "Anemia is a blood disorder in which the blood has a reduced ability to carry oxygen due to a lower-than-normal number of red blood cells."*			
A.90 Anaemia is when blood doesn't have enough red blood cells to carry oxygen			
A.91 Have you ever had your iron level tested during this pregnancy?		1	Yes
		2	No
		-999	Don't know/ don't remember
		-888	Refuse
A.92 What are some of the ways you can prevent anemia? *FOR ENUMERATORS: Please do not read out answer options. FOR ENUMERATORS: If respondent still does*		1	Eat foods with iron
		2	Take deworming tablets

not know what anemia is then you can now explain "Anemia is a condition relating to iron levels in your blood." If they want/need more information, you can reiterate that : "It is a blood disorder in which the blood has a reduced ability to carry oxygen due to a lower-than-normal number of red blood cells."	3	Take iron-folic-acid supplements
	4	Take other medicines
	0	I don't know what anemia is
	99	Other
	-999	Don't know
	-888	Refuse
A.93 (Specify other) Please type the respondent's "other" answer.		
A.94 What are some foods that have lots of iron? *FOR ENUMERATORS: Please do not read out answer options*	1	Cereals, Nuts, Raisins
	2	Meat, Fish, Liver
	3	Milk, Yogurt, Cheese
	4	Green leafy vegetables, Beetroot
	6	Fruits: Strawberries, Watermelon, Dates, Figs Prunes, Dried apricots, dried fruits
	5	Lentils, beans
	7	Other fruits
	-999	Don't know
	-888	Refuse
A.95 During your pregnancy, were you ever given or did you ever buy any iron-folic acid tablets or syrup?	1	Yes
	2	No
	-999	Don't know/ don't remember
	-888	Refuse
A.96 During your pregnancy, did you ever eat or drink any iron-folic acid tablets or syrup?	1	Yes
	2	No
	-999	Don't know/ don't remember
	-888	Refuse
A.97 In a typical month during pregnancy, how many days did you eat or drink the iron-folic tablets or syrup?	1	Every day/ almost every day
	2	Every other day/ 3-4 times a week
	3	Once or twice a week
	4	Once in 2 weeks
	5	Once a month
	6	I stopped taking iron folic tablets midway through my pregnancy
	-888	Refuse
A.98 We will now ask you a few questions about your experience with MomConnect. We are coming near the end of the survey, please give us a few more minutes of your time.		
A.99 Did you use the internet (including social media) during your recent pregnancy to find information about pregnancy, childbirth, and baby care?	1	Yes
	2	No
	-888	Refuse
A.100 In general, how trustworthy would you rate MomConnect as a source of information on pregnancy, childbirth, and newborn care?	1	Extremely trustworthy
	2	Very trustworthy
	3	Moderately trustworthy
	4	Slightly trustworthy
	5	Not at all trustworthy
	99	Don't know / Can't judge

			-888	Refuse
	A.101 What are some reasons why you trust MomConnect as a source of information on pregnancy, childbirth, and newborn care? *FOR ENUMERATORS: Please do not read answer choices.*		1	They are in partnership with the Department of Health
			2	They were recommended by my doctor/ nurse
			3	The information helped me have a safe pregnancy/ birth
			4	They helped me keep up with my clinic appointments
			5	They were quick to get back to me
			6	The information they provide is detailed
			-888	Refuse
	A. 102 Which part of MomConnect do you find most useful?		1	The appointment reminders
			2	The information about my pregnancy
			3	The information about my baby
			4	The helpdesk so I can ask questions
			5	The extra suggested information that I can read if I want to
			6	Other
			7	Nothing - MomConnect is not helpful
			-888	Refuse
			-999	Don't know / don't remember
	A.105 How helpful are the appointment reminders?		1	Extremely helpful
			2	Very helpful
			3	Moderately helpful
			4	Slightly helpful
			5	Not at all helpful
			0	I do not remember this part of the message
			-888	Refuse
	A.106 How helpful is the information about baby's growth?		1	Extremely helpful
			2	Very helpful
			3	Moderately helpful
			4	Slightly helpful
			5	Not at all helpful
			0	I do not remember this part of the message
			-888	Refuse
	A.107 How helpful is the information about how to keep you and your baby healthy?		1	Extremely helpful
			2	Very helpful
			3	Moderately helpful
			4	Slightly helpful
			5	Not at all helpful
			0	I do not remember this part of the message
			-888	Refuse
	A.108 Overall, how supported or not supported have you felt by MomConnect messages?		1	Extremely supported
			2	Very supported
			3	Moderately supported
			4	Slightly supported
			5	Not supported at all
			-999	Don't know

		-888	Refuse
A.109 Would you have preferred to have received messages over WhatsApp or SMS?		1	SMS
		2	WhatsApp
		-888	Refuse
A. 109A How easy or difficult is it to find accurate information using MomConnect?		1	Very easy
		2	Somewhat easy
		3	Neither easy not difficult
		4	Somewhat difficult
		5	Very difficult
		-888	Refuse
		-999	Don't know / don't remember
A. 109B How likely are you to recommend MomConnect to a friend or family member?		1	Very likely
		2	Somewhat likely
		3	Neither likely nor unlikely
		4	Somewhat unlikely
		5	Very unlikely
		-888	Refuse
		-999	Don't know / don't remember
A. 109C How many of the MomConnect messages did you read the whole way through?		1	All
		2	Most
		3	About half
		4	A few
		5	None
		-888	Refuse
		-999	Don't know / don't remember
A. 109D Do you feel that you got messages from MomConnect: Too often, Just Right or Not Often Enough?		1	Too often
		2	Just right
		3	Not often enough
		-888	Refuse
		-999	Don't know / don't remember
A. 109E Would you like to receive images and/or videos from MomConnect? e.g. showing how to breastfeed or how to bath your baby?		1	Yes
		2	No
		-999	Don't know / don't remember
		-888	Refuse
A.110 Thank you for answering those questions! For the final section, I will now ask some questions about you.			
A.111 What was your marital status at the time you learned about this pregnancy?		1	Never married/ single
		2	Married
		3	Separated or divorced
		4	Widowed
		5	Had a partner or boyfriend
		-888	Refuse
A.112 Which answer best describes your highest level of education completed at the time you learned about this pregnancy?		1	I didn't finish primary school
		2	I finished grade 7
		3	I finished grade 12
		4	I finished a certificate or degree after high school

		-888	Refuse
A.113 What was your work at the time you learned about this pregnancy?		1	I was unemployed
		2	I worked for a private person/ business
		3	I worked for the government
		4	I was self-employed
		5	Student
		-888	Refuse
A.114 Before you were pregnant, what was the total monthly income, in Rand, for your household normally? (Hint: This includes the take-home pay that members of your household brought home, and can be an estimate)			
A.115 Which population group do you consider yourself: black, white, colored, Indian, or something else?		1	Black African
		2	White
		3	Colored
		4	Indian/ Asian
		99	Other
		-888	Refuse
A.116 How many living children do you have in total? *Enter -888 for refuse*			
reached > Speaking to Mother > Consenting Survey > ~ > Baby was lost			

Appendix 2: Power Calculations

- Calculations conducted
- Assumptions

Due to resource restrictions, we do not ask all survey questions to each study participant. The target sample size for primary outcomes is 5,200, whereas for secondary outcomes it is 2,200. Meaning that 3,000 participants receive the "short survey" that asks questions pertaining to the primary outcomes, whereas 2,200 of them receive the "long survey" that asks all questions.

Primary Outcomes

Since the study is 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. Prior to the study, a sample size of 5,200 - 1,300 for each study arm - was decided. This 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 key parameters used in calculating this sample size are:

- **Level of statistical significance (Alpha):** 0.05 as is the research standard
- **Power**: 0.80 as is the research standard
- **Unit of Treated Assignment:** Individual
- **Predicted Mean of Outcome Variables:** 0.68. Based on previous data on the proportion of mothers on MomConnect that have provided their child with full infant vaccination, this is 68%

We 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.

Secondary Outcomes

We only measure our two secondary outcomes, the knowledge index and the behaviour index, 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.

Appendix 3: Survey Management

The phone survey protocol was as follows:

- Phone the phone number provided through the MomConnect registration, and if a respondent does not pick up, call them three times - this is classed as one attempt
- Attempt to contact all participants up to a maximum of seven times
- Only attempt to call someone a maximum of twice in one day
- Only one attempt per time slot (morning, early afternoon, and late afternoon)

Surveys are facilitated through SurveyCTO on the mobile app or a computer browser. The team included one supervisor and two team leads, who, along with the project manager, were responsible for conducting the following data quality assurance methods to ensure the integrity of the data:

1. High-frequency checks
2. Back checks for 12.1% of the sample (protocol was a minimum of 10%)
3. Weekly spot-checks with each enumerator

Appendix 4: Secondary Outcome Indicators

Our secondary outcomes include indices based on a number of questions related to both knowledge of healthy behaviours as well as actual adoption of healthy behaviours. 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 behaviour outcomes are 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 immunisations and ANC visits, no single topic is identified as more important or decision-relevant than others; second, compared to the alternative of analysing 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 immunisations six weeks after birth, the number of ANC visits pregnant women went to during their pregnancy, mothers' knowledge/attitudes, and mothers' behaviour.

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) is multiplied by a multiplier:

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

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

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

Outcome	Definition
Midline Outcomes	
Knowledge of/ Attitudes toward Healthy Behaviors (Max score = 11)	$((Diet\ score\ *\ \frac{11}{4})\ +\ (Pregnancy\ danger\ signs\ score\ *\ \frac{11}{16})\ +\ (Breastfeeding\ score\ *\ \frac{11}{12})\ +\ (Intention\ to\ vaccinate\ score\ *\ \frac{11}{12}))$ *where:* Diet score (0-1) = the sum of the following points: • 1 point if mothers correctly identify all of the foods mentioned in MomConnect messages that are rich in iron Pregnancy danger signs score (0-4) = the sum of the following points: • 1 point if mothers recognise the symptoms of preeclampsia • 1 point if mothers recognise the symptoms of anemia • 1 point if mothers recognise that vaginal bleeding is a danger sign • 1 point if mothers know the healthy frequency of baby kicks Breastfeeding score (0-3) = the sum of the following points: • 1 point if mothers intend on breastfeeding their baby • 1 point if mothers intend on introducing breastmilk to baby within 1 hour of birth • 1 point if mothers intend to exclusively breastfeed for 6 months Intention to vaccinate score (0-3) = the sum of the following points: • 1 point if mothers agree or strongly agree that it is important to vaccinate her baby • 1 point if mothers agree or strongly agree that the benefits of vaccines outweigh the risks of her baby developing side effects from them • 1 point if mothers are likely or very likely to follow the recommended vaccine schedule
Adoption of healthy behaviors (Max score = 4)	$(Diet\ score)$ *where:* Diet score (0-4) = the sum of the following points: • 1 point if mothers report eating fruits on a daily basis • 1 point if mothers report eating vegetables on a daily basis • 1 point if mothers report eating liver once or twice per month • 1 point if mothers report not drinking alcohol
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 recognise that her baby not moving is a danger sign • 1 point if mothers recognise that her baby having a seizure is a danger sign • 1 point if mothers recognise 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	$((Diet\ score) + (Pregnancy\ preparedness\ score\ * \frac{4}{7}) + (Breastfeeding\ score\ * 2) +$

behaviors

(Max score = 16)

$(Anemia\ treatment\ score))$

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:

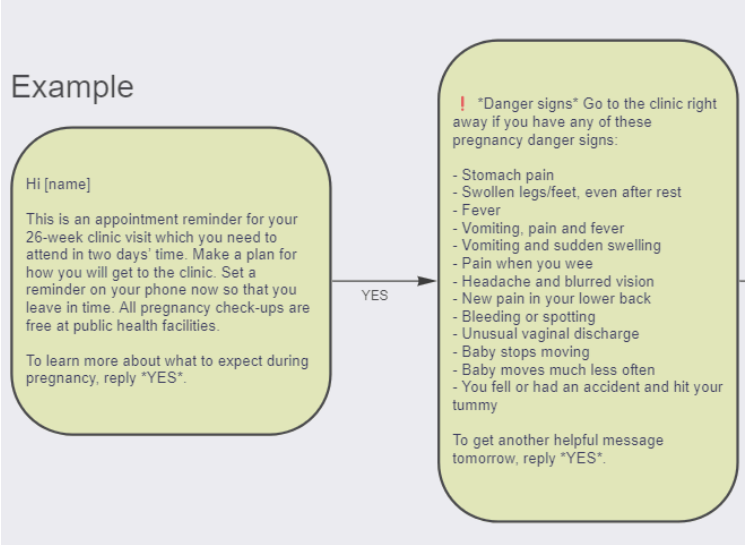
- 1 point if mothers are currently only feeding her baby breastmilk
- 1 point if mothers attempted to breastfeed their baby within the first hour of birth

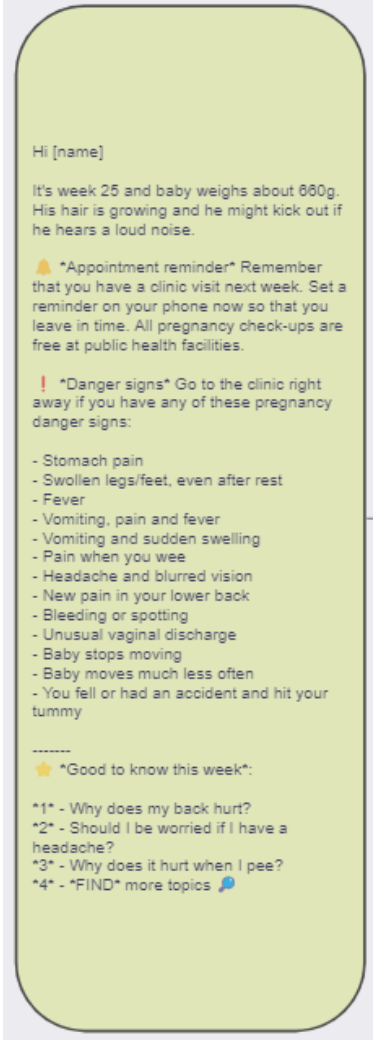
Anemia treatment score (0-4) = the sum of the following points:

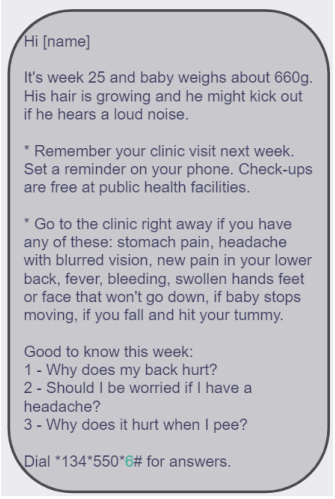
- 1 point if mothers were tested for anemia during her pregnancy
- 1 point if mothers possessed iron folic acid pills during her pregnancy
- 1 point if mothers consumed iron folic acid pills during her pregnancy
- 1 point if mothers consumed iron folic acid pills on a daily basis during her pregnancy

Appendix 5: Example messages of different treatment arms

ARM (control)	RCM-WA & RCM-BCM	RCM-SMS







Note that the only difference between RCM-WA and RCM-BCM is that RCM-BCM contains option 4 "*FIND* more topics" whereas RCM-WA does not