Trial Protocol: Chaco, Argentina – COVID-19 Chatbot

July 11, 2022

Contents

PICOS statement:

What is the problem?	Despite overwhelming medical evidence on the protection offered by second, third and fourth Covid-19 vaccine doses, uptake of these doses remain low despite available supply.
Comments	The problem likely stems from the fact there are currently no reminders in place in Chaco. It is therefore difficult for residents to know if they are eligible for the next dose of the COVID-19 Vaccine and where and when to access it.

Population: Who are the participants?	All adult citizens of the Province of Chaco, Argentina who have not completed the full Covid-19 vaccination schedule.
Comments	*People without phone numbers in administrative databases will be excluded from our sample given phone-based message delivery.*
Intervention: What are we doing?	Participants will receive interactive and personalised Whatsapp chatbot messages informing them they are eligible for the next dose, sharing practical information on where and when to get vaccinated, providing planning prompts and enabling them to set their own reminders to get their next dose of the Covid-19 vaccine on a convenient date.
Comments	Public health chatbot messaging intervention.
Comparison:	Randomised controlled trial. Comparison conditions: (1) Pure control, where participants do not receive a message; (2) Simple message control, where participants receive a single non-chatbot (i.e., not interactive) message encouraging vaccination.
Comments	Stratified randomization. Stratification variable: participants will be stratified on the number of vaccine doses that they have already received (2, 3, or 4). Unit of randomisation: Individual
Outcome(s):	**Primary outcome: '**Next dose' Covid-19 vaccination rate (binary – did the individual receive the latest dose for which they are eligible? – yes/no). This will be analyzed four weeks after the messages are sent. **Exploratory outcomes:** Click-through to chatbot messages (in chatbot treatment arms); sub-group analysis for specific doses (second, third, and fourth).
Comments	Some participants may not be able to receive a 'next dose' vaccination due to a recent Covid-19 infection; this will not be measured in the data.
Setting:	Province of Chaco, Argentina
Comments	Remote delivery.

1. Trial rationale & challenges

Rationale:

- Third and fourth dose vaccination against Covid-19 is effective in reducing infection, severe illness and death, especially for the Omicron variant.
- Despite the overwhelming evidence for the protection offered by boosters, uptake of third and fourth doses remain low.
- In collaboration with the Government of Argentina's Behavioural Insights Unit and the Province of El Chaco, the Behavioural Insights Team is working to conduct an evaluation of the impact of Whatsapp chatbot messages to increase uptake of second, third, and fourth Covid-19 vaccination doses.
- **Social Impact:** Booster vaccination is associated with lower risk of mortality due to COVID-19.[1] Increasing uptake of boosters can therefore save the lives of those receiving the booster. In addition, by reducing strain on healthcare providers, increased booster vaccination may have benefits to the health of others (i.e., not only those receiving the vaccine).

Identified barriers:

There is low uptake of third and fourth dose vaccines against Covid-19. Our desk-based research and fieldwork identified several behavioural barriers specific to the local context.

- **Confusion over how and when to get the vaccine**. A review of government websites and other official sources of information consistently showed out-of-date information about how to get vaccinated, including where and when to get a dose and eligibility criteria. For example, the purpose-built website for Covid-19 vaccination in Chaco (elijovacunarme.chaco.gob.ar) still shows a large pop-up message stressing that an appointment is needed to get vaccinated, and that upon attending that appointment it is necessary to show a medical certificate that verifies the patient has a condition that makes them eligible for vaccination (none of which is true any longer). Relatedly, we found evidence of confusion about how, when and where to get vaccinated for a first and additional doses among Chaco residents. We gathered this information from questions and reviews posted on Facebook and GoogleMaps.
- **Perceived low risk from Covid.** In focus groups and interviews with Chaco residents and healthcare workers, we identified very low perception of risk as a key barrier to vaccination. In early 2022 case numbers fell to very low levels in Argentina, overall vaccination rates were high and restrictions ended; according to interviewees this contributed to a feeling that the pandemic was ending and getting first or additional vaccinations wasn't necessary.

[1] Arbel, R., Hammerman, A., Sergienko, R., Friger, M., Peretz, A., Netzer, D., & Yaron, S. (2021). BNT162b2 vaccine booster and mortality due to Covid-19. *New England Journal of Medicine*, *385*(26), 2413-2420.

- **Lack of any reminder service.** Currently in Chaco, residents do not receive any communication or reminders from the local government or health services with information about when they are eligible for their next dose or prompting them to complete their set of vaccinations.
- **Low visibility of vaccination centres**. After the main mass-vaccination drive ended in Chaco, pop-up and vaccination-only sites were closed down, and vaccination became available primarily in hospitals and health clinics[2]. From observations carried out during fieldwork in Chaco, these vaccination locations do not advertise or make clear that they offer Covid-19 vaccination externally. This indicates that the visibility and salience of vaccination centres has decreased over time.

How our intervention addresses the barriers identified:

- **Confusion over how and when to get the vaccine:** The chatbot enables users to find their nearest vaccination centre(s). To do this, the chatbot has a location function where users can share their current location (either by "dropping a pin" in WhatsApp, or typing their home postcode), after which the chatbot will provide a list of their nearest centre(s), how far away they are, and directions for how to get there (in a later reminder message). This addresses the friction cost of having to go to an outdated government website and find the information yourself, and removes the risk that the information will be incorrect. The chatbot also provides opening hours for each centre.
- **Perceived low risk from Covid:** The chatbot uses behaviourally-informed messaging to motivate vaccination. One way it does this is using a loss aversion framing, which reminds users that they should get their next dose to avoid losing the immunity they currently have from previous vaccinations. It should be noted that this framing is also in the simple message control arm.[3]
- **Lack of any reminder service:** First, the chatbot provides personalised information to let users know which dose they are now eligible for, based on the dates of the doses of their previous vaccinations. Second, whilst the core chatbot messages themselves act as a prompt to get vaccinated, the chatbot also allows users to set their own reminder for a convenient date that they want to get the vaccine. This helps deal with an inattention barrier (i.e., getting vaccinated slips your mind). There is strong evidence demonstrating the use of mobile message reminders to increase uptake of vaccines, including the influenza vaccine and COVID-19 vaccine. A Cochrane review of 55 studies found that SMS messages increase relative uptake by 29% on average. In addition, the reminder we are using gives the user the agency to set a reminder for the day before a date that is convenient for them.

[2] Some pop-up vaccination sites have continued to be organised in Chaco for specific population drives, but increasingly these are primarily for other vaccinations, such as the flu vaccine

[3] In this project, we are interested in testing to what extent an interactive chatbot messaging service outperforms a simple message reminder. Since non-personalised motivational message frames can be incorporated into simple messages, and so does not constitute a chatbot "functionality", we have chosen to include it in both trial arms.

- **Low visibility of vaccination centres:** By combining a reminder service with an information service about locations of vaccine centres, the chatbot addresses the fact that residents of Chaco may not otherwise know where to get vaccination and receive no prompts from their environment.

To our knowledge, the effect of interactive chatbots on the uptake of vaccinations has not been causally evaluated using a behavioural outcome (actual vaccination rates rather than, e.g. stated intention). We believe that an interactive chatbot has the potential to be an even more effective reminder than an SMS, and we hope to be one of the first to evaluate this.

2. Roles and responsibilities in this trial

This section should summarise the roles and responsibilities of each organisation involved in the trial. These are intended as high-level summaries of critical tasks, ***not*** a detailed project plan. The actual project plan / Gantt chart should be linked to this section, but note that it is *not* the QA responsibility to review these documents.

Detailed project plan link: **[See here]**

Who is responsible for...	Organisation name	Person responsible
Data sharing agreements	BIT; ECOM	Adelaida Barrera
Collecting outcome data	ECOM	Lucas Ibañez
Randomisation	BIT	Pujen Shrestha
Delivering the intervention	ECOM	Lucas Ibañez
Analysis	BIT	Pujen Shrestha

3. Research Aims, question and hypotheses

What are the specific research questions for this trial?

This is a 3-arm individual-level randomised controlled trial (RCT) to test whether a COVID-19 vaccination messaging Chatbot in Chaco, Argentina can increase vaccine uptake amongst those who have not completed their full vaccination schedule (i.e., not received a second, third, or fourth dose of the vaccine). The social impact would be to improve the COVID-19 vaccine uptake.

What is the trial aiming to do?

To investigate whether a chatbot designed to make the process of getting a Covid-19 vaccination easier can improve COVID-19 vaccination rates.

There are several **research questions** for this trial:

1. Primary research question: Does sending Covid-19 vaccination reminders via Chatbot to Chaco residents who are due for a second, third, or fourth dose increase their vaccination, compared to those who do not receive a reminder or those who receive a simple non-chatbot reminder?
2. Exploratory research question: We will ask the same research questions, but separately for each sub-group of participants based on the number of existing doses they have had (one, two or three).

Primary analysis - effect on vaccination

Hypothesis 1a: Our chatbot intervention causes an increase in 'next dose' vaccination rate, compared to the pure control arm.

Hypothesis 1b: Our chatbot intervention causes an increase in 'next dose' vaccination rate, compared to the simple non-chatbot message arm.

Exploratory analysis – vaccine dose subgroups

Hypothesis 2a: Our intervention causes an increase in the vaccination rate within the sub-groups of participants whose 'next dose' is a second dose amongst the Treatment groups, compared to the pure control arm and simple non-chatbot message arm.

Hypothesis 2b: Our intervention causes an increase in the vaccination rate within the sub-groups of participants whose 'next dose' is a third dose amongst the Treatment groups, compared to the pure control arm and simple non-chatbot message arm.

Hypothesis 2c: Our intervention causes an increase in the vaccination rate within the sub-groups of participants whose 'next dose' is a fourth dose amongst the Treatment groups, compared to the pure control arm and simple non-chatbot message arm.

4. Interventions being tested in this trial

One line on what the intervention is:.

A Vaccine Messaging Chatbot delivered to residents of Chaco, Argentina who have not received a 'next dose' (i.e., a second, third, or fourth dose) of the COVID-19.

What resources are needed to deliver the intervention:

All materials and Chatbot functionalities are developed by BIT in collaboration with ECOM who deploy the Chatbot. The flow of the Chatbot can be found here.

What were the practical steps taken for intervention delivery?

BIT will conduct randomization. BIT will provide ECOM with a full list of participants that have been randomised with a variable that identifies which arm they are to be assigned to. ECOM will then send out messages to participants according to that list.

Shortly before launching the full trial, ECOM will send messages to a pilot sample of 200 people randomly selected from the eligible sample, to check that the chatbot is working as intended. This follows several rounds of internal pilot testing of the chatbot by ECOM and BIT team members.

Who delivered the intervention?.

The intervention is delivered by ECOM through a text message on Whatsapp.

How long did the intervention last / how long was each session?

The intervention will last four weeks starting at 1,000 messages a day ramping up to as many as 100,000 messages a day by the trial's end.

Where did it take place?

The intervention will take place in Chaco, Argentina.

Was there any planned adaptation/variation of the intervention?

There are no planned adaptations or variations of the intervention.

Control/business as usual (BAU):

There are currently no active text messaging reminders to be vaccinated in Chaco, Argentina.

Table for multi-arm trials

The table below describes the arms in this trial. We have included both the Spanish and English versions of the initial text message the participant receives.

Condition	Description
Control	Participants will not receive a message. Note: Following the trial period, if any messages are found to be effective at encouraging vaccination, these individuals would be sent the most effective message.
T1: Simple reminder	Participants will receive a single 'static' whatsapp message (i.e., a message they cannot interact with). Key features of this arm are: ● Motivational message leveraging loss aversion (don't lose your immunity) ● No personalized information, or information on how to access vaccines English translation of this message: *Remember it is important to complete your vaccination scheme against COVID-19 and get your booster shots. With time, your body's defenses against the virus decrease. Boost your vaccine and don't lose your protection against the virus!* *You can find more information on how to get your vaccine on the Ministry of Health official website.*
T2: Chatbot reminder / planning tool	Participants will receive an initial whatsapp message motivating them to get a next dose which they can interact with – if they consent to receiving future messages, they will receive further content via Chatbot functionality. Key features of this arm are: ● Motivational message leveraging loss aversion (don't lose your immunity) ● Chatbot functionalities:

	○ Personalised information (participants' name and previous dose information) ○ Information on closest vaccination site (after participant drops a pin or shares their postcode), and opening hours. ○ Prompt to plan when to receive the next dose (date and time). ○ Automated reminder one day ahead of selected date for next dose, including a link to Google maps directions to the vaccination centre. English translation of initial message: *Hi, <<personalized name>> 👋. This is a new service by the Ministry of Health to provide you information about your next free vaccination against COVID19.* 🦠 *According to the Ministry's registry, you received the <<number of dose>> on <<date of most recent vaccination>> and it is now time for you to get the <<next dose>> dose.* *Over time, your defences against the virus decrease. Strengthen your vaccine and don't lose your protection against the virus!* *In this chat we can help you learn when and where to get your next shot.* ***Would you like to continue?*** *(Find more information about this service on the Ministry of Health Chaco Facebook page. If you do not respond to this message, we will not contact you again).*

5. Design

Is this an RCT? Yes

Number of arms: Three arms - one Control group, and two treatments.

Unit of randomisation: Individual

What is the unit of randomisation? Chaco residents that have not received their 'next eligible dose' of the COVID-19 vaccination (i.e., a second, third, fourth, or fifth COVID-19 vaccination dose).

What is the unit of analysis?

We are using mobile phone numbers as the unit of analysis and randomisation (see Inclusion Criteria for further information on this). Importantly, we are using mobile phone numbers as a proxy for individuals. After applying our eligibility criteria there is a 1:1 match between the phone number and DNI for our sample, meaning that randomising by the phone number is equivalent to randomising by the DNI. Due to the administrative nature of the data we are using for analysis and the data sharing process, we don't have individuals' DNIs, which means the closest way for us to identify individuals is to use mobile phone numbers. We note that there are some limitations to this approach: For example, we exclude any individuals that do not have a mobile phone. However, we believe that these limitations are acceptable risks for our analysis.

Spillover / contamination:

		Potential impact on trial		
		Low	*Med.*	*High*
Risk of contamination	*Low*			
	Med.		X	
	High			

Please provide a brief **written justification** (<100 words) for the judgement you gave in the table:

The utilisation of a Chatbot as the intervention allows for a strict control of who receives the intervention allowing us to minimise the risk of contamination.

The main spillover risk comes from the fact that we are adding a page on the Ministry of Health (MoH) website and the MoH facebook group to explain that the MoH are piloting a new whatsapp

chatbot service to help people get a COVID-19 vaccination (so that people can verify the authenticity of our service). There is a chance that people in the control group observe this message.[4]

Whilst more than one individual can share the same phone number, we do not believe this poses a spillover risk for this study. By definition, individuals in the control group have their own phone number, and so are highly unlikely to be sharing a phone with an individual in the treatment group. It is possible that individuals in the treatment group know individuals in the control group and show them the chatbot service, but we believe this risk is low.

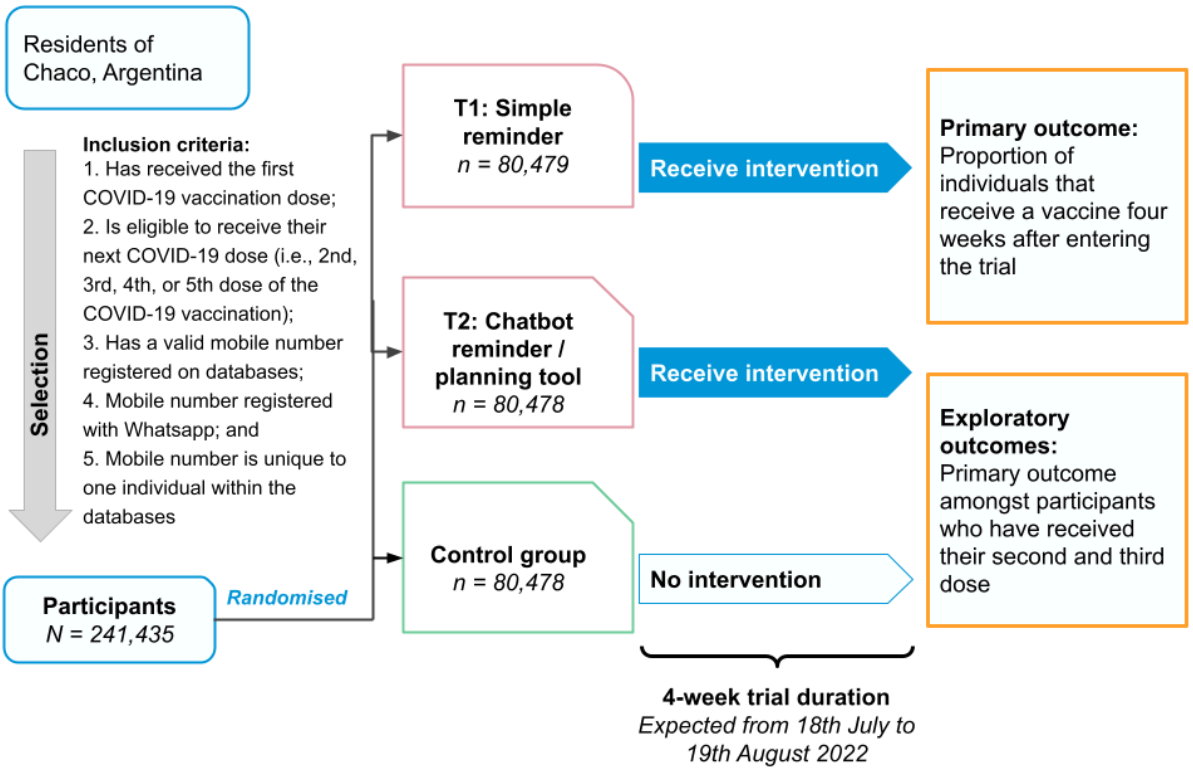
Please provide a brief description of how you will handle contamination in the trial design or account for it in analysis (and ensure that that you reflect analytical points in the analysis section below):

Contamination is handled in the intervention design by only allowing those with mobile numbers in the treatment groups to open the chatbot. If a participant in pure control or outside the intervention scope attempts to open the chat bot they will be presented with a default message which does not reference Covid-19 vaccination in order to not act as an inadvertent reminder. We are limiting the eligible participants to unique phone numbers, therefore, limiting the spillover where we can not be sure who is receiving the chatbot.

We note that there may be a certain amount of contamination from the sharing of screenshots of the chatbot, or people telling others about it verbally - which we cannot avoid. However, we believe that the extent of this contamination will be small, and very unlikely to occur in the first place. Importantly, this contamination would likely increase vaccination levels in the Control group, which would bias our estimated treatment effects towards zero, meaning that any effect we do estimate may be considered a lower bound for the true effect of the chatbot.

[4] We do not have a strong prior as to which direction this would bias our estimate. On the one hand, it could encourage individuals in the control group to get vaccinated if it served as a prompt about Covid-19 vaccinations. On the other hand, it could discourage them from getting the vaccination if they feel like they have “missed out” on a service to help make the process easier. However, given that an individual has to first be on the MoH website or facebook page to see this message, we think the risk that an individual in the control group is exposed is low.

Participant flow



6. Outcome Measures

Outcome Measure	Data to be collected	Point of Collection
Primary: Proportion of individuals that receive a vaccine four weeks after entering the trial	Participants' 'next dose' vaccination status (i.e., whether or not they received the next dose they were eligible for).	Throughout the trial period and for four weeks after the intervention completes.
Exploratory: Primary outcome applied to the sub-group of participants that have received one, two or three doses.	Participants' 'next dose' vaccination status (i.e., whether or not they received the next dose they were eligible for).	Throughout the trial period and for 4 weeks after the intervention completes (i.e., after the final introductory chatbot message is sent).

Sample Selection and Eligibility

Our dataset will use the combination of datasets: (i) data from the COVID-19 helpline; and (ii) Federal MoH (Nomivac); (iii) Pasaporte Chaco; (iv) SUMAR. These datasets comprise vaccinations that have been provided since May 15th, 2020 across Chaco Argentina.

We will exclude all observations that do not meet the following inclusion criteria:

1. Has received the first COVID-19 vaccination dose;
2. Is eligible to receive their next COVID-19 dose (i.e., 2nd, 3rd, 4th, or 5th dose of the COVID-19 vaccination)[5];
3. Is 18 years of age or older;
4. Has a valid mobile number registered on databases;
5. Mobile number registered with Whatsapp; and
6. Mobile number is unique to one individual within the database

Our dataset comprises a total sample of 1,027,125 observations. After exclusions, our sample comprises 241,435 mobile numbers.

Attrition

Do we have an estimate for expected attrition overall from the study?

No. This is a novel trial therefore we do not have an expected overall attrition rate for the study.

What is the overall expected attrition from the study, as a percent of units randomised?

We are using an administrative dataset on Covid-19 vaccinations for our outcome measure. Attrition in this context means that an individual has left Chaco province such that their vaccination status is no longer captured in the Nomivac dataset from Chaco that we have access to. We therefore believe that the rate of attrition will be extremely low, and do not believe there is any reason why the intervention itself would cause an individual to get vaccinated in another province. We therefore assume a low rate of attrition of 0-5% for both the control and treatment groups.

[5] Due to the complexity of the vaccination eligibility criteria we have included a full discussion of this in Appendix B.

Data Gathering table

Data	Collection Point	Source
Personal and contact details: DOB, sex, moblie phone number, indigenous ethnicity	At start of trial (used for sample selection and randomisation)	Administrative data collected routinely. This data comes from a combination of 4 different data sources (i) data from the COVID-19 helpline; (ii) Federal MoH (Nomivac); (iii) Pasaporte Chaco, and (iv) Padron Indigena.
Vaccination data: vaccine dose, vaccine dose date, vaccine dose brand	At start of trial (used for sample selection and randomisation); four weeks after the intervention has finished being deployed on the sample	Administrative data collected routinely.
Whether participants' click into the Chatbot	Collected automatically during trial. Send to BIT after trial	ECOM

Data Storage and Transmission

Data will be pseudo-anonymised by [ECOM] and stored in project folders with access restricted to the project team only. Data will not be transmitted to third parties, except where this is appropriate under the conditions of appropriate data sharing agreements.[6]

[6] Note: There is no data sharing agreement with BIT. Instead, the Government of Argentina's Behavioural Insights Unit has signed an agreement with the Ministry of Health in Chaco allowing them to share data with project partners.

7. Randomisation

Structure:

	Simple: do you know which units [people, places] will be in the trial before the trial starts with equal allocation (balanced sample sizes) to treatment and control?
	Pipeline: are you randomising as people/units enter into the trial? (AKA 'trickle trials'.
X	**Stratified:** are you making sure that the conditions are balanced in terms of a key variable e.g. gender?
	Wait-list: are you randomising participants/units to 'treatment now' vs 'treatment later'?
	Stepped wedge: are you randomising the order in which sites are allocated to conditions? (Note this is only used for cluster RCTs.)
	Unequal randomisation: are you randomising to have unequal sample sizes? What is the justification for this approach?

Blinding: Although participants will be aware of the motivational message they have been exposed to, and whether or not they have been exposed to the chatbot, prior to outcome data collection, we do not expect many to be aware that they are in a trial where different participants are exposed to different conditions. The Chatbot is set up in such a way that when participants, who are identified by their phone, re-engage with the Chatbot at a later time they will be presented with the same condition.

Allocation mechanism: Randomisation will be conducted by BIT. Participants will be allocated to a trial arm using a stratified randomisation at the individual-level over the unique identifier of the participants' mobile number[7] We will stratify participants on the number of vaccine doses for each mobile phone number (1, 2, 3, or 4).

Balance: We will conduct a balance check on the following baseline characteristics:

[7] There is a 1:1 match between the phone number and DNI in our sample, meaning randomising by phone number is identical to randomising by individual.

- Sex
- Age
- Indigenous ethnicity

We will conduct statistical tests to establish whether the experimental arms are balanced or not by these observables. We will not consider these tests as evidence of randomisation success or failure in terms of unobservables.

8. Trial Procedure

Trial implementation

The following steps will be undertaken by BIT and project partners to implement the trial.

Timing	Owner	Action
Pre launch	All project partners	Develop content for treatment arms
Pre launch	ECOM	Set up and internally test a chatbot to deliver messages for all treatments. Set up a mechanism to monitor message delivery and users' engagement with the chatbot.
Pre launch	BIT	Determine sample size and conduct power calculations based on a pre-trial cut of data.
Pre launch	BIT	Conduct randomization. BIT will provide ECOM with a full list of participants that have been randmomised with a variable that identifies which arm they are to be assigned to.
Pre launch	ECOM	Send out messages to a pilot sample of 200 people randomly selected from the eligible sample. Note: These individuals should then be excluded from the trial sample for launch.
Launch: Days 1-3	ECOM	Send out 1,000 business-initiated per day ("Tier 1" limit. See two examples of how moving through tiers may work in Figure X below). • On day 1: Send 1,000 Introductory messages, ensuring the number of introductory messages sent are balanced across treatment arms. • On day 2: Send any reminder messages (NIVEL 7 in chatbot flow) for people already in the trial who selected the next day as their vaccination date. Then, use any remaining messages (up to 1,000 limit) to send introductory text messages, ensuring the number of introductory messages sent are balanced across treatment arms. Keep a

		detailed spreadsheet log of how many messages are sent to each group. ● On day 3 (and potentially onwards): Repeat Day 2 procedures. Repeat until we reach “Tier 2” messaging limit.
Launch: Days 4-16	ECOM	Assuming the “Tier 2” limit is reached, send out 10,000 business-initiated messages per day. ● Each day, prioritise sending any reminder messages (NIVEL 7 in chatbot flow) for people already in the trial who selected the next day as their vaccination date. Then, use any remaining messages (up to 10,000 limit) to send introductory messages, ensuring the number of introductory messages sent are balanced across treatment arms. Keep a detailed spreadsheet log of how many messages are sent to each group. ● In order to reach the full sample our messaging account must be upgraded to Tier 2 by the 22nd day of the trial. It is not necessary for us to be upgraded to Tier 3.
Launch: Days 17 onwards		Assuming “Tier 3 limit is reached” send out up to 100,000 business-initiated messages per day. ● Each day, prioritize sending any reminder messages (NIVEL 7 in chatbot flow) for people already in the trial who selected the next day as their vaccination date. ● Then, use any remaining messages (up to 100,000 limit) to send introductory text messages, ensuring the number of introductory messages sent are balanced across treatment arms. Keep a detailed spreadsheet log of how many messages are sent to each group.
One month after sending final introductory message	ECOM	Conclude data collection and send BIT final dataset. ● If there are no delays with the message account graduations through the tiers we should expect all messages to be sent by the 12th day of the intervention. ● If we are not upgraded beyond Tier 1 we will continue messaging throughout the one month intervention period and conclude having sent ~37,000 intervention messages. This will affect the power of this trial (see: Appendix A). ● In order to reach the full sample our messaging account must be upgraded to Tier 2 by the 22nd day of the trial. It is not necessary for us to be upgraded to Tier 3.

Figure X. Two examples of moving through Chatbot messaging tiers.

	24 hours (Day 1)	24 hours (Day 2)	24 hours (Day 3)	24 hours (Day 4)	24 hours (Day 5)
Number of users messaged	1,000	1,000	1,000		
Total number of users messaged	1,000	2,000	3,000		
Messaging Limit Tier	1K	1K	10K		

Number of users messaged	500	500	500	500	500
Total number of users messaged	500	1,000	1,500	2,000	2,500
Messaging Limit Tier	1K	1K	1K	1K	10K

Additional notes on randomization and tier-based implementation strategy:

- BIT will randomise the dataset for ECOM who will send the maximum number of messages per 24 hour period that the tier cap allows.They will do this by on day 1 selecting participant 1 through x (the amount allowed by the tier cap) and then work their way down the list day to day (for each of the two treatment arms). They will record when each individual message goes out.
- This list will be designed to be balanced across the arms in blocks of 1,000 participants to ensure that we do not send significantly more messages to one trial arm than another.
- If a participant from the treatment group reaches out to our WhatsApp number before they are scheduled to be included in the trial, they will temporarily receive the default "this is a test service" message and then receive the treatment group chatbot message at a later date when they are scheduled to be included in the trial.
- However, there are a couple of risks with taking this automated "full list" approach, which we will navigate with the following strategies:
    1. We will ensure that people for whom a message failed to send are not lost the following day (i.e., at the point that we get an error because the cap is hit, we stop sending further messages and then start again the next day from that same point on the list of trial participants).
    2. We will ensure that people who are set to receive a reminder message on a given day take priority over new trial participants. To see why: suppose that on Day 2 of the intervention we are in Tier 1 (1,000 cap), and we are due to message 800 new trial participants, but there are also 500 existing participants who are meant to receive a reminder (because they previously selected tomorrow as their appointment date). We don't want to send the 800 new messages first, and then 200 of the existing participants, hit the cap, and

then fail to send a reminder that we promised to the remaining 300 existing participants.

- Given the rolling design of this trial, where participants receive the intervention on consecutive dates during the trial period, we have to divide the pure control group into a "pure control group for each day" to ensure the 4 week outcome window is consistent between treatment and control. To tie our hands, before trial launch and during randomisation we randomly assigned a number to each individual in the pure control group. At the conclusion of the trial period we will use that list to assign individuals from the pure control group to the “pure control group for each day”. E.g., if there are 500 individuals in the each of the chatbot and simple message arms on Day 1, and 400 in each of the chatbot and simple message arms on Day 2, we will take individuals in the pure control group who are randomly assigned the numbers 1-500 to become the “Day 1 pure control group”, take individuals numbered 501-900 to become the “Day 2 pure control group” and so on. The purpose of this strategy is to ensure that the pure control group is observed over the same time periods (in the rolling design) as the other two trial arms.

Data Quality Monitoring

The following steps will conduct the following steps for data quality monitoring:

Timing	Owner	Action
Pre-launch	BIT	BIT provides ECOM with a full list of participants in the trial indicating which arm they are assigned to.
3 day post launch	ECOM	ECOM provides BIT with chatbot dataset
4 days post launch	BIT	BIT reviews the chatbot outcome database to check randomization and message delivery measures, including clickthrough for chatbot arms.
8 days post launch	ECOM	ECOM provides BIT with chatbot dataset
9 days post launch	BIT	BIT reviews the chatbot database to check randomization and message delivery measures, including clickthrough for chatbot arms.
14 days post launch	ECOM	ECOM provides BIT with the full outcome database for review of implementation and message delivery measures, as well as vaccination outcomes.

16 days post launch	BIT	BIT reviews database to check randomization, message delivery measures, and primary measure (vaccination rate) data collection.
21 days post launch	ECOM	ECOM provides BIT with the full outcome database for review of implementation and message delivery measures, as well as vaccination outcomes.
23 days post launch	BIT	BIT reviews database to check randomization, message delivery measures, and primary measure (vaccination rate) data collection.

Stopping Rules

During the trial period, a trial arm will be halted if it is estimated to be causing a statistically significant decrease in vaccination rates equal or greater to 0.5 SDs. This will be assessed by BIT at 16 days post launch and 23 days post launch.

Stopping rule	Rule is that...	Who is responsible?
Rule 1	At 16 days post launch, assignment to a treatment arm yields a decrease in vaccination rate equal or greater to 0.5 SDs	BIT (Pujen Shrestha)
Rule 2	At 23 days post launch, assignment to a treatment arm yields a decrease in vaccination rate equal or greater to 0.5 SDs	BIT (Pujen Shrestha)

9. Power Calculations

Our power calculations are summarised below. We first determined the sample size based on data provided by ECOM and excluding all observations that do not meet our eligibility criteria. This resulted in a sample size of 241,435. Our analysis suggests that the minimum detectable effect for a sample of this size is 0.32pp, assuming 5% attrition.

To estimate our anticipated effect size, we first reviewed the literature investigating both SMS reminders for vaccines and the use of chatbots. Notably, while there is some indication that chatbots are being used in the public health response to Covid-19,[8] there is little empirical evidence that chatbots can increase vaccine

[8] Amiri, P., & Karahanna, E. (2022). Chatbot use cases in the Covid-19 public health response. *Journal of the American Medical Informatics Association, 29*(5), 1000–1010. https://doi.org/10.1093/jamia/ocac014

uptake. Consequently, we have used multiple sources to come to our estimated effect size:

- A meta-analysis of 13 empirical studies that tested digital-push interventions effectiveness at increasing vaccine uptake. The authors found that patients who received the digital-push interventions had 1.18 the odds of receiving vaccination (or series completion) compared to controls. An odds of 1.18 applied to the base rate of 3.8% in the Control group, our anticipated effect size is 0.7pp (equivalent to Cohen's h of 0.04).;[9]
- A meta-analysis of 10 studies that tested the effectiveness of text message reminders on childhood vaccination. The authors found a positive effect of text message reminders on childhood vaccination coverage (RR = 1.11);[10]
- One study found that the use of a chatbot designed to answer questions about COVID-19 vaccines significantly increases people's intentions to get vaccinated and has a positive impact on their attitudes toward COVID-19 vaccination;[11] and
- BIT undertook a project with a WhatsApp chatbot, delivering behaviorally-informed interventions aimed at helping South African adolescent girls and young women navigate unhealthy relationships. While the outcome was different to vaccination rates, the authors found that all chabots were effective in increasing attitudes and beliefs about power in relationships compared to a pure control (Cohen's *d* of 0.20-0.29). A Cohen's *d* of 0.20-0.29 applied to the base rate of 3.8% in the Control group translates to approximately a 4.7-7.4pp increase.

Overall, these data suggest that text messages can have a material impact on vaccination rates compared to a pure control group. Moverover, there is indicative evidence that chatbots could be more effective than text messages. Given the wide ranging (0.7-7.4pp increase), but directionally positive, estimates on effect size above, we conservatively assume that the intervention will approximately have an effect size of 1.0pp.

Summary of Power Calculation assumptions & inputs

[9] Atkinson, K. M., Wilson, K., Murphy, M. S. Q., El-Halabi, S., Kahale, L. A., Laflamme, L. L., & El-Khatib, Z. (2019). Effectiveness of digital technologies at improving vaccine uptake and series completion – A systematic review and meta-analysis of randomized controlled trials. Vaccine, 37(23), 3050–3060. https://doi.org/10.1016/j.vaccine.2019.03.063

[10] Mekonnen, Z. A., Gelaye, K. A., Were, M. C., Gashu, K. D., & Tilahun, B. C. (2019). Effect of mobile text message reminders on routine childhood vaccination: A systematic review and meta-analysis. Systematic Reviews, 8(1), 1–14. https://doi.org/10.1186/s13643-019-1054-0

[11] Altay, S., Hacquin, A. S., Chevallier, C., & Mercier, H. (2021). Information Delivered by a Chatbot Has a Positive Impact on COVID-19 Vaccines Attitudes and Intentions. Journal of Experimental Psychology: Applied. https://doi.org/10.1037/xap0000400

Alpha (significance level)	5%
Power	80%
Total planned sample size	241,435 mobile phone numbers. Due to the potential issues with delivering the intervention to the full sample due to the tiering constraints of Whatsapp discussed in the *Trial implementation* section, we have conducted further power calculations for potential tiering scenarios in Appendix A.
Clustered trial?	**No**
Number of trial arms	Three arms - One Control group and two Treatment groups
Base rate or SD	The base rate is the average proportion of individuals that receive a vaccination across any 4-week period in 6 months before the trial: 3.8%
Attrition	As described earlier in this document we do not have a confident estimate of the expected attrition rate in this trial. Therefore we have conducted the power calculations with an estimate of 5% attrition.
What is the planned MDES for this trial?	~0.31 ppts (0% attrition) to ~0.32 ppts 5% attrition)
Anticipated *statistical effect size* of the intervention?	An 1 percentage point difference in uptake between Treatment and Control
Anticipated *substantive* effect of the intervention?	An increase in vaccination of 1,609 individuals amongst the Treatment groups[12]
Is the planned MDES the same as or smaller than the anticipated effect of the intervention?	**The planned MDES is smaller than the anticipated effect of the intervention**
Have you corrected for multiple comparisons?	Yes, using Bonferroni multiple comparison adjustments to be conservative.

Table X: Power Calculations

Alpha	**5% (Bonferroni multiple comparison adjusted)**

[12] Calculated as (0.01 [Anticipated effect size] * 241,435 [Total sample] / 3 [Number of arms] * 2 [number of Treatment arms])

Power		80%					
Baseline	Attrition			Arms	Sample size	Cohen's h	MDES ppt
3.8%	0%	5%	20%	3			
✔	✔			✔	241,435	0.02	0.31
✔		✔		✔	229,363	0.02	0.32
✔			✔	✔	193,148	0.02	0.35

Notes: ppt = percentage points; MDES = minimum detectable effect size.

10. Analytical Strategy

Primary Outcome: Estimating the impact of the Chatbot on vaccination rates

We will use a logistic regression to estimate the Intention-To-Treat (ITT) effect of the Chatbot intervention on the binary primary outcome of vaccination rate. Random assignment of individuals to receive the Chatbot allows us to identify the effects.

Our primary analysis will use the following covariate-adjusted regression. We will include all complete cases and make a missing-at-random assumption:

$$Y_i \sim bernoulli(p_i)\ ;\ logit(p_i) = \beta_0 + \beta_1 TC_i + \beta_2 TS_i + X_i'\gamma$$

where the function $logit$ is defined as the log-odds ratio

$$logit(p) = log(\frac{p}{1-p})$$

and,

- Y_i is a binary indicator of whether the individual is receives a vaccination (1 if they do, 0 if not);
- p_i is the probability that the individual individual is receives a vaccination;
- TC_i is a dummy variable indicating whether individual i is assigned the chatbot (1 if they are, 0 if not);

- TS_i is a dummy variable indicating whether individual i is assigned the simple message reminder (1 if they are, 0 if not);
- X_i is a vector of pre-treatment covariates
    - Sex (binary)
    - Age as a categorical variable (18-29, 30-49, 50+)
    - Current vaccine dose (1st dose, 2nd dose, 3rd dose, 4th dose)
    - Length of time since previous dose
    - Date of initial intervention text message

We are interested in comparing each treatment arm (chatbot and simple reminder) to the pure control group, and we are interested in testing whether the chatbot performs significantly better than the simple message reminder. We are therefore making a total of 3 comparisons: (H_1: $\beta_1 = 0$,), (H_2: $\beta_2 = 0$,), (H_3: $\beta_1 = \beta_2$,) and so we will use the Benjamini-Hochberg Procedure for multiple comparison adjusted p-values.

Exploratory Outcome: Vaccination Dose Subgroups

We will use a logistic regression to estimate the Intention-To-Treat (ITT) effect of the Chatbot intervention on the vaccination rate separately within each sub-groups of participants whose 'next dose' is a second, third and fourth dose. Random assignment of individuals to receive the Chatbot allows us to identify the effects. We will estimate the following covariate adjusted regression, and as with the primary analysis we will include all complete cases using a missing-at-random assumption:

$$Y_i \sim bernoulli(p_i)\ ;\ logit(p_i) = \beta_0 + \beta_1 TC_i + \beta_2 TS_i + X_i'\gamma$$

where the function $logit$ is defined as the log-odds ratio

$$logit(p) = log(\frac{p}{1-p})$$

and,

- Y_i is a binary indicator of whether the individual is receives a vaccination (1 if they do, 0 if not);

- p_i is the probability that the individual individual is receives a vaccination;
- TC_i is a dummy variable indicating whether individual i is assigned the chatbot (1 if they are, 0 if not);
- TS_i is a dummy variable indicating whether individual i is assigned the simple message reminder (1 if they are, 0 if not);
- X_i is a vector of pre-treatment covariates
    - Sex (binary)
    - Age (as a categorical variable bins of 18-29, 30-49, 50+)
    - Current vaccine dose (1st dose, 2nd dose, 3rd dose, 4th dose)
    - Length of time since previous dose
    - Date of initial intervention text message

What steps will you take for assessing and dealing with missing data?

Missingness of covariates: We will assess if any covariate is missing for more than 5% of the observations. We will also check if the rate of missingness varies between arms, which we do not expect to be the case if randomisation was implemented as planned.

- If fewer than 5% of cases have some covariate data missing, and the missingness rate is within 2 percentage points across arms, we will simply exclude these cases (i.e. perform a complete case analysis).
- If more than 5% of cases have some covariate data missing, or the missingness rate differs by more than 2 percentage points across arms, we will add an additional 'missing' category to the relevant covariates (or in the case of a continuous covariate, imputing an arbitrary value and adding a 'missing' indicator variable) so that any effect associated with the data being missing can be modelled and accounted for in the treatment effect estimate.

Process evaluation

We are also conducting a process evaluation to better understand the implementation of the chatbot and how participants engaged with it. We will employ

descriptive statistics to analyse six research questions using the data recorded when participants use the chatbot. This data will be shared by ECOM with BIT.

Topic	Research question
Willingness to engage with MoH *(Understand to what extent there is any interest in reading a message sent from MoH)*	What proportion of the chatbot sample read the first message (i.e., double blue tick).
Interest in the chatbot *(Understand whether people have any interest in using a chatbot)*	Of the people in the chatbot arm for whom the message successfully delivered, what proportion responded Yes to the first message (as opposed to "No" or "Another time").
Vaccine intentions *(Understand the intention to get vaccinated within the chatbot arm)*	Of the people who wanted to use the chatbot (i.e., of those who responded "Yes" to the first message): what proportion got the full way through the bot to the point of setting a reminder message.
Ease of use of the chatbot *(Understand whether there are parts of the chatbot that people find more difficult to use)*	At what stage did the chatbot sample stop interacting with the chatbot (i.e., % of chatbot sample who dropped out at each of the messages).
Location functionalities *(Understand which location sharing functionality people prefer)*	Of the people who chose to submit their location, what proportion dropped a pin vs shared postcode.
Value of reminders *(Understand how valuable the self-initiated reminder within the chatbot is - e.g., if most people just choose to go the next day, then the reminder feature probably isn't particularly useful)*	Of people who choose a date, how far in advance do they choose to get vaccinated.

11. Ethical Issues & Review

This trial was self-assessed as being:	**Low**
The reason for assessment was...	Standard research methods; Recipients are non-vulnerable adults; Anonymous data; BIT has run a similar project in this domain before.
Link to completed ethics review form (if medium or high level risk)	Research risk assessment framework

What were the key ethical considerations for the project?

- The type of research methods used are standard - an RCT implemented remotely using routinely collected data.
- Coronavirus is a potentially contentious issue, however, the intervention is supporting health-promoting behaviours recommended by the government.
- We are not collecting Personal Data from respondents; ECOM will not share National ID numbers with BIT.

Did you seek informed consent from participants?

Obtaining informed consent for the purposes of this study could undermine the study results. As the intervention itself does not affect the treatment being provided to patients, we determined that not obtaining consent was justifiable.

Have you considered whether harms to participants might arise and how you will deal with them?

One risk is that the behaviourally-informed messages have an unintended backfire effect and reduce vaccination uptake in some groups. This risk has been minimised by developing the messages based on the evidence available from related research, as well as input from experts at BIT and consultations with the Government of Argentina's *Unidad de Ciencias del Comportamiento y Políticas Pública*.

To further mitigate this risk, we have conducted a focus group with local residents of Chaco, Argentina in which they tested a version of the BI-informed chatbot message flow in their own mobile phones.

12. Risks

Risk	Strategy to mitigate risk	Responsibility	Timeframe (if applicable)
Intervention may backfire and lead to worse outcomes	This risk has been minimised by developing the messages based on the evidence available from related research, as well as input from experts at BIT and consultations with the Government of Argentina's Unidad de Ciencias del Comportamiento y Políticas Pública. To further mitigate this risk, we have conducted a focus group with local residents of Chaco, Argentina in which they tested a version of the BI-informed chatbot message flow in their own mobile phones.	BIT	N/A
Randomisation failure	Ensure that ECOM are fully briefed on how to allocate participants to different conditions and carry out interim balance checks (if possible) to double check. Randomisation results can be shared in three separate and clearly labelled spreadsheets if ECOM prefers. BIT will ensure that no identifiable information is seen here by the BIT team. Any (potential) observed small imbalance (less than 10%) between treatment and control will be accounted for in the analysis by including the set of regressors available in the data.		Pre-trial; During trial; Analysis period

Spillover / contamination: (1) The main spillover risk comes from the fact that we are adding a page on the Ministry of Health (MoH) website and the MoH facebook group to explain that the MoH are piloting a new whatsapp chatbot service to help people get a COVID-19 vaccination. (2) There is a small risk of spillover as the chatbot will be sent to specific phone numbers so we will not be able to control whether multiple individuals share this number.	Contamination is handled in the intervention design by only allowing those with mobile numbers in the treatment groups to open the chatbot. If a participant in pure control or outside the invention scope attempts to open the chat bot they will be presented with a default message which does not reference Covid-19 vaccination in order to not act as an inadvertent reminder. We are limiting the eligible participants to unique phone numbers, therefore, limiting the spillover where we can not be sure who is receiving the chatbot.		During trial
Proportion of inaccurate contact numbers for participants higher than estimated	We have only include participants for which we have a unique mobile number and are verified Whatsapp users to mitigate this risk		Pre-trial
8 weeks (4 week message period plus 4 week outcome data collection period) is insufficient for caregivers to receive message, plan and attend appointment and for vaccination status to be updated on the system	Estimate the lag between booking and attendance more precisely based on existing records. Extend the time frame where appropriate if 4 week proves to be insufficient.	BIT, ECOM	During trial; Analysis period
Risk that due to the tiered Whatsapp message system our account never graduates from Tier 1 and we are not able to deliver the intervention to the full sample we intended to.	According to our power calculations in Appendix A we will still have a sufficient sample after 4 weeks as long as we can make max use of the Tier 1 messaging limit for that entire period.	BIT, ECOM	During trial; Analysis period
Risk from Covid-19, such as restrictions or pressure on healthcare services preventing routine care, such as vaccinations	Extend the trial period if the Covid-19 context changes before the start of the evaluation		During trial

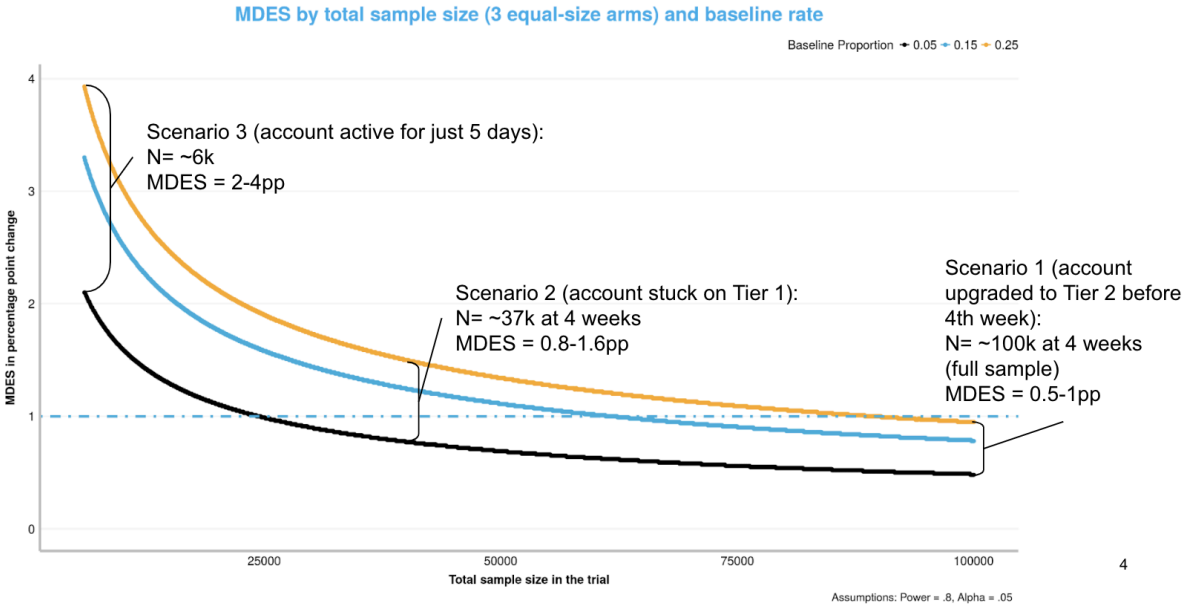
Appendix A: Power Calculations for WhatsApp messaging contingencies

In order to better understand the risks associated with the Whatsapp tiering system that dictates how many intervention messages we will be able to deliver we have conducted further power calculations using the same assumptions as presented in *Summary of Power Calculation assumptions & inputs*. These calculations are based on three scenarios, (i) No unexpected tier graduation failures, that is to say we are able to reach the full sample, (ii) we are limited to Tier 1 throughout implementation; and (iii) account is suspended after 5 days. Fundamentally these scenarios affect the power of our trial by reducing the sample size of our treatment arms. As table X and figure X indicates if our messaging account never graduates from Tier 1, we will still have a sufficient sample after 4 weeks to detect an effect size of 1 pp. However, the more problematic scenario where our messaging account is suspended after only 5 days of operation will result in an underpowered trial.

Table X: Power Calculations

Alpha				5% (Bonferroni multiple comparison adjusted)			
Power				80%			
Baseline	Tier Scenarios			Arms	Sample size	Cohen's h	MDES ppt
3.8%	(i) No unexpected tier graduation failures	(ii) Limited to Tier 1 throughout implementation	(iii) Account is suspended after 5 days	3			
✔	✔			✔	241,435	0.02	0.31
✔		✔		✔	37,333	0.04	0.82
✔			✔	✔	6,000	0.11	2.26

Notes: ppt = percentage points; MDES = minimum detectable effect size.



4

Appendix B: Vaccination Eligibility Criteria

In order to establish whether a participant is eligible to be included in our trial sample we used the most up-to-date vaccine eligibility criteria provided to us by Argentina's Ministry of Health. Due to the differences in eligibility based on various factors, primarily the initial vaccination they received, the time period they must wait till they are eligible for their next dose, and personal characteristics such as age and immunosuppressed status, we have described the specific eligibility, per initial dose brand, requirements below. It should also be noted that vaccines were used interchangeably, for example an individual may have received Moderna as their first dose, Pfizer as their second dose, and Moderna as their first booster. Therefore, in the eligibility criteria below a participant will enter their branch of eligibility based on the initial dose they received.

Moderna, Pfizer or Sputnik

If an individual's first dose was the Moderna, Pfizer or Sputnik vaccine they are only eligible for their second dose after 21 days. An individual has completed their full course once they've received their first two doses of these vaccines. Once completing their full course individuals are eligible to receive their first booster dose after 120 days. Once receiving their first booster individuals are eligible to receive their second booster dose after 120 days.

AstraZeneca

If an individual's first dose was the AstraZeneca vaccine they are only eligible for their second dose after 8 weeks. An individual has completed their full course once they've received their first two doses of these vaccines. Once completing their full course individuals are eligible to receive their first booster dose after 120 days. Once receiving their first booster individuals are eligible to receive their second booster dose after 120 days.

Sinopharm

If an individual's first dose was the Sinopharm vaccine they are only eligible for their second dose after 21 days. In most cases 2 doses of the Sinopharm is the full course (meaning an individual would be eligible for their first booster dose). However, there is an exception for people over the age of 50 or considered immunosuppressed. These individuals would be eligible for their third doses after 28 days, and only after receiving this dose have completed their full course. Once completing their full course individuals are eligible to receive their first booster dose after 120 days. Once receiving their first booster individuals are eligible to receive their second booster dose after 120 days.

Cansino

If an individual's first dose was the Cansino vaccine they are considered to have completed their full course of the vaccine. This vaccine was primarily administered to vulnerable populations, for example, the homeless. Once completing their full course, individuals are eligible to receive their first booster dose after 120 days. Once receiving their first booster individuals are eligible to receive their second booster dose after 120 days.

Ordering and counting of vaccination doses:
In the vaccination data shared with BIT occasionally an individual was described to have had a vaccination dose before they had the necessary previous dose (the labelling of which dose they received was inaccurate) or the previous dose was not recorded. Therefore we decided to use a “date approach”, where we count the number of doses in date order and to establish the number and date of the doses they received. We then use this information to check if they are eligible for their next dose. The date approach works in all cases except one type of instance where we know that some data wasn't recorded.[13]

[13] During our field visit we heard that some first doses had not been uploaded from paper records to the administrative dataset. Therefore in instances where an individual has had a later dose (e.g., dose 2) but not their first dose, we have assumed that they did have their first dose but it was simply not recorded on the system. For example, where 'dose 2' and 'booster' are recorded but there is no 'dose 1'. In this case we assume they've had 3 doses.