Facemasks to Reduce COVID-19 in Bangladesh

December 2020

Updated 6/25/21 where indicated
(prior to receipt of any test results)

NCT04630054

1 Introduction

1.1 Study Motivation

The COVID-19 pandemic has had health, economic, and social repercussions around the globe and, as of January 2021, has taken the lives of more than 1.8 million people. The use of face masks has become one of the most important health recommendations to slow the spread of the disease. Existing laboratory evidence reveals face masks can reduce exhaled viral load and thus the probability of transmitting the virus. However, the extent to which face masks effectively reduce transmission in the real world, where they may be worn imperfectly and inconsistently, is uncertain. Face masks could also give people a false sense of safety that encourages them to engage in risky behaviors. Further, there is little evidence on the best strategies to encourage correct face mask use. This research aims to fill these gaps by conducting a large-scale randomized evaluation on the impact of mask use on community and individual infection rates, measured through COVID-19 testing. The results of this evaluation will serve as global evidence for the fight against the pandemic.

As a densely populated country with 165 million people, Bangladesh needs to enacts effective COVID-19 prevention measures. However, we documented a troubling decline in self-reported and observed mask use since the beginning of the pandemic. We started are working in partnership with the Ministry of Health and Family Welfare, The Bangladesh Medical Research Council, and a2i, an information and data focused organization within the Bangladesh government, to address these concerns with rigorous research.

1.2 Research Questions

1. Can face mask distribution and promotion at mosques, markets, and homes successfully change community mask-wearing norms from the status quo to appropriate wearing of high-quality masks?
2. Can community mask-wearing reduce the transmission of symptomatic respiratory infections and symptomatic COVID-19 infection?
3. Does mask-wearing induce risk-compensation such as reduced physical distancing?
4. Can mask-wearing reduce the risk of COVID-19 infection for the individual mask wearer?

1.3 Intervention Overview

In Bangladesh, we are implementing a clustered randomized controlled trial to measure the impact of mask distribution and promotion on COVID-19 infection rates — measured by COVID-19 antibody testing — at the community and individual level. Other outcomes of interest include physical distancing behavior, mask use, and the impact of different types of mask use incentives. Researchers will also examine if cloth and surgical masks differ in their effectiveness in preventing COVID-19.

1.4 Community Intervention

Based on population data and in-person scoping by the IPA team 1,000 rural and peri-urban villages were randomly selected. Each village was assigned to a "Unit" based on its Upazila (500 administrative regions into which Bangladesh is divided) x (Above/Below median population within Upazila) x trajectory of cases per person of -1,0,1 based on (incomplete) case data from the preceding 6 weeks. Within units, we selected 600 pairs of villages (assigning pairs based on similarity of the most recent per capita case data. Within each pair, one village was randomly assigned to treatment and one to control. The intervention consists of the distribution and promotion of face masks. All villages within the sameTwo-thirds of the distributed masks are surgical masks and one-third are cloth masks. We believe surgical masks may be more effective at preventing COVID-19 transmission, but people may be more likely to consistently wear cloth masks. We will test the assumptions and the underlying health trade-off. There are also other encouragements and incentives cross-randomized to determine how to most effectively increase mask use.

Randomizations

Village-level Cross-randomizations Our intervention has four village-level cross-randomizations. All four randomizations are applied to each village. These randomizations are:

1. Randomization of treated villages to either cloth or surgical masks
2. Randomization of treated villages to either {no incentive, monetary incentive, non-monetary incentive}. These are rewards given to the village if a certain level of mask-use is met.
3. Randomization of treated villages to public commitment (asking households to place signage on doors that declares they are a mask-wearing household) or not
4. Randomization of treated villages to 0% or 100% of households receiving twice-weekly text reminders

Household-level Cross-randomizations We have three household-level cross-randomizations. One and only one randomization is applied to each village. While we do not observe mask-wearing for each household, the color of the masks distributed to the household is dependent on the randomization status of the household, and surveillance staff will record the color of our project's masks worn by community members.

The possible household-level randomizations are:

1. In some villages, households will be randomized to receive messages emphasizing altruism or messages that are focused primarily on self-protection

2. In other villages, households will be randomized to receive twice-weekly text reminders or not

3. In a third set of villages (chosen among those without the public signage commitment), households will be randomized to making a verbal commitment to be a mask-wearing household (all adults in the household promise to wear a mask when they are outside and around other people) or not.

Summary of two-stage randomization We conduct a two-stage randomization, randomizing treatments at both the village-level and the household-level.

1. VILLAGE RANDOMIZATION #1: We randomize 1/3 of intervention villages to have cloth masks and 2/3 surgical masks

2. VILLAGE RANDOMIZATION #2: We randomize $\frac{1}{2}$ of villages to have commitment signage on the door (if the household commits to mask wearing) and $\frac{1}{2}$ not to have signage.

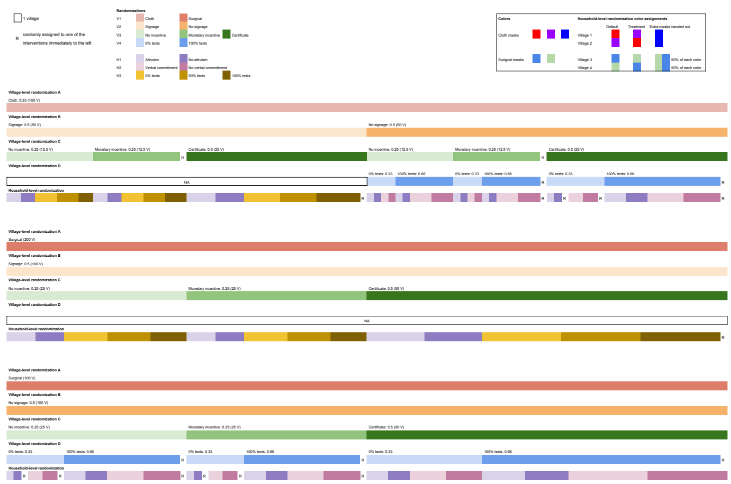
3. VILLAGE RANDOMIZATION #3: Within each of the first two randomizations, we randomize 1/4 to receive no incentive, 1/4 to receive a community-level reward of 190 USD, and 1/2 to a certificate of recognition from the Government of Bangladesh. The monetary reward and certificate will be given if village-level mask wearing among adults is >75% 8-weeks after the intervention has started.

4. VILLAGE RANDOMIZATION #4: Conditional on each of the first three assignments, in villages without signage (150 villages), we randomize 2/3 of these villages to receive mask wearing encouragement texts (100 villages) and 1/3 will not receive texts (50 villages).

5. HOUSEHOLD RANDOMIZATION #1: In the villages without signage (150 villages), we assign 1/3 of villages (50 villages) to have each household randomized to altruism or not and 2/3 of villages (100 villages) randomized to verbal commitment or not.

6. HOUSEHOLD RANDOMIZATION B: In the villages with signage (150 village), we randomize 1/3 of villages (50 villages) to altruism or not and we randomize the remaining 2/3 of villages (100 villages) to have either 100% of households receive texts (25 villages), 50% of households receive texts (50 villages), or 0% of households receive texts (25 villages)

Figure 1: Schematic of cross-randomizations



1.5 Individual Intervention

The goal of the individual experiment is to assess whether mask-wearing reduces the risk of COVID-19 infection for individuals wearing masks. We distinguish this from the community mask-wearing experiment, where the goal is to assess the joint impact of masks on the spread of the virus by preventing transmission.

In order to quantify the potential protective effect of a mask reducing the risk that the mask-wearer becomes infected, we enroll vendors at indoor markets and randomize 50% to the treatment group and 50% to the control group. We will collect blood from the highest-risk individual in each of the 2,500 shops. Individuals in the treatment group will receive masks and continue to be surrounded by most people who are not using masks (this is an assumption due to our past observations. We recorded only 20% of people wore masks in public locations). Through the intervention period of 12 weeks, we will observe whether or not the vendor is wearing a mask and conduct telephone surveys to ask about their respiratory symptoms. After 12 weeks we will collect blood from the same 2,500 vendors who provided blood at baseline and test theses samples for COVID-19 antibodies. We hypothesize that masks will protect the mask wearer from COVID-19 infection, so we expect to see fewer respiratory symptoms and COVID-19 infections in the treatment arm.

2 Primary Analyses

2.1 Community Intervention: Symptomatic Seropositivity

Our primary outcome is symptomatic seropositivity. We will construct a dataset with an observation for each surveyed household (i indexes households and j villages). In each village, define $Y_{ij} = 1$ if the highest risk individual in each household is 1) reporting either a. dry cough and fever and b. either fatigue, lack of taste/smell or shortness of breath in the last month at either the fifth week or ninth week telephone survey and 2) are seropositive in our blood test at endline. If either of these conditions fail to hold, $Y_{ij} = 0$. To assess seropositivity, we will test all individuals who are symptomatic in either our 5-week or 9-week household survey.

UPDATE 6/25/21 (prior to analysis of any test results): seropositivity will be assessed using a SARS-CoV-2 ELISA antibody test. Positivity is defined according to the manufacturer's pre-defined cut-off value for a positive sample determination based on the immunological status ratio (ISR). An ISR value of 1.1 or above is considered a positive test result.

Our goal will be to estimate the impact of the intervention on seropositivity, defined as: $\psi_0 = E_x[E(Y_{ij}|T_j = 1, x_j) - E(Y_{ij}|T_j = 0, x_j)]$ where T_j is an indicator for whether a village was treated and x_j are village-level covariates including baseline mask-use in each village (constructed as described below) and baseline influenza-like illness and COVID-19 based on reported symptoms, as well as indicators for each pair of villages from our pairwise stratification method. In an auxiliary specification, we will also include fixed effects for each surveillance staff member.

UPDATE 6/25/21 (prior to analysis of any test results): We will estimate this parameter by ordinary least squares, clustering at the village-level using the approach in (Correia, 2017). The dependent variable is Y_{ij}, the independent variable of interest is T_j, and controls will be included for the x_j covariates, including baseline mask-use and baseline respiratory symptom rates in each village.

To estimate the overall impact of masks on seroconversions, we have chosen a random cohort of 25,000 individuals from which we collected blood at baseline. We will test the baseline bloodspots for individuals who are symptomatic at endline in order to estimate baseline symptomatic seropositivity. By differencing the seropositivity rate among symptomatic individuals at baseline and endline, we can compute symptomatic seroconversions. We can determine the fraction of seroconversions prevented by dividing ψ_0 by the overall rate of symptomatic seroconversions.

2.2 Individual Intervention: Seroconversion

Our primary health outcome in the individual experiment will be seroconversion (seronegative at baseline and seropositive at endline). Note that this differs from symptomatic seropositivity, since we will observe a baseline serology test for all participants in the individual experiment and do not restrict attention to those who are symptomatic at endline for our primary outcome. Let C_i be an indicator for seroconversion for individual i. We will then estimate two specifications:

- An intent-to-treat specification, where we estimate:

$$E(C_i|D_i, \xi_m) = \beta_{ind} D_i + \xi_m \tag{1}$$

  by regressing C_i on the treatment indicator D_i and fixed effects for each indoor market within which the randomization is conducted.

- An instrumental variable (IV) specification, where we regress C_i on p_i instrumented with D_i and controlling for ξ_m.

3 Secondary Analyses

3.1 Community Intervention: Mask Wearing

We will create a dataset with an observation for each village j. We will define proper mask use as anyone wearing either a project mask or an alternative face-covering that covers their mouth and nose. We will consider two definitions of p_j, the proportion of surveilled individuals wearing masks. In our primary specification, we will define p_j using all surveilled individuals. Surveillance is conducted using a fixed protocol in each village where the surveillance staff rotate between public settings and observe the number of passersby, as well as their social distancing and mask use. In a secondary specification, we will consider individuals surveilled only in locations where we were not currently doing active mask-promotion (that is, where no staff were onsite distributing masks at the time of surveillance). The purpose of this second specification is to investigate separately whether the intervention increased mask-use in places where we did not have promoters on site. Our goal will be to estimate the impact of the intervention on the probability of mask-wearing, defined as: $\psi_1 = E_x[E(p_j|T_j = 1, x_j) - E(p_j|T_j = 0, x_j)]$ where T_j is an indicator for whether a village was treated and x_j are village-level covariates including baseline mask-use in each village (constructed analogously to p_j), baseline respiratory symptom rates, and indicators for each pair of villages from our pairwise stratification method.

We will estimate this equation at the village-level by ordinarily least squares, using analytic weights proportional to the number of surveilled individuals (the denominator of Δ_j) and heteroskedastic-robust standard errors. In this specification, the dependent variable is p_j, the independent variable of interest is T_j, and controls will be included for the x_j covariates. The same approach will be used to estimate all village-level specifications unless otherwise specified.

We will also run this analysis separately in mosques, markets and all other locations pooled (e.g. tea stalls and village entrances) to assess our ability to increase mask use in each location. We will also consider as an auxiliary regression the same specification with a minimal set of controls (i.e. only the village-level indicators).

3.2 Community Intervention: Village-level Cross-randomizations

We will analyze all four village level cross-randomizations jointly by estimating the conditional expectation:

$$E(p_j|T_j, x_j, D_k) = \beta T_j + \sum_k D_k \delta_k + x_j \gamma \tag{2}$$

where $D_k = 1$ if cross-randomization number k is turned on in that village (meaning that it is also a treatment village) and 0 otherwise. This specification will be identical to our estimating equation for the impact of treatment on mask-use, with the addition of the D_k terms.

The primary analysis assumes no interactions occur between cross-randomized treatments. We will evaluate potential interactions between cross-randomized treatments in a secondary analysis by estimating the following conditional expectation. We note that this analysis may have limited statistical power.

$$E(p_j|T_j, x_j, D_k) = \beta T_j + \sum_{k'} \sum_{k>k'} D_k D_{k'} \delta_{k,k'} + x_j \gamma \tag{3}$$

The same approach will be used to estimate all village-level specifications unless otherwise specified.

3.3 Community Intervention: Household-level Cross-randomizations

We will start by estimating, at the village-level, whether masks of the treated color in any given household randomization m are more commonplace than masks of the control color. In each village, we will compute Δ_j, the difference in the fraction of individuals wearing treated mask colors vs. control mask colors. We alternate across villages which color corresponds to treatment, so we can also control directly for indicators for whether specific colors are more popular (denote these by d_{jc}; $d_{jc} = 1$ if treated masks in village j are color c). Our estimate for each household randomization will be α_{0m}, given by:

$$E(\Delta_j|d_{jc}) = \alpha_{0m} + \sum_c \alpha_c d_{jc} + surgical_j \tag{4}$$

α_{0m} tells us how much more likely individuals are to wear masks of the treated color than masks of the control color. We will estimate this equation at the village-level by ordinarily least squares, using analytic weights proportional to the number of surveilled individuals (the denominator of Δ_j) and heteroskedastic-robust standard errors. We will additionally control for whether the village is a surgical or cloth mask village.

The above specification will be used to assess whether the household cross-randomizations impacted behavior. Under strong parametric assumptions, we can scale α_{0m} appropriately to recover the impact of the household level treatment on the probability of wearing a mask. Let Y_{ij} be an indicator variable which is 1 if the surveilled individual is wearing any type of mask (a mask distributed by the project or any other type of mask) over the nose and mouth and zero otherwise and let H_{ij} be 1 if the mask worn by individual i is the color assigned to the household-level randomization treatment condition in village j and zero otherwise. For each house-hold treatment m, we would ideally like to recover $\psi_m = E_{d,\xi}[E(Y_{ij}|I_{ijm} = 1, d_{jc}) - E(Y_{ij}|I_{ijm} = 0, d_{jc})]$ where I_{ijm} is an indicator for whether i's household received the treatment m. I_{ijm} is not observed, but if we assume that treated and non-treated households never wear the mask whose color is opposite their treatment status, and that any impact of the treatment is only via our project masks, then: $E(Y_{ij}|I_{ijm} = 1, d_{jc}) = 2E(Y_{ij}H_{ij}|d_{jc})$ (e.g. if 10% of people wear blue masks and half of people are treated, this implies 20% of treated people wear blue masks). Additionally: $E(Y_{ij}|I_{ijm} = 0, d_{jc}) = 2E(Y_{ij}(1 - H_{ij})|d_{jc})$. Finally, $E(\Delta_j|d_{jc}) = E(Y_{ij}H_{ij} - Y_{ij}(1 - H_{ij})|d_{jc})$, which implies that we can estimate ψ_m using $2\alpha_{0m}$.

3.4 Individual Intervention: Mask-Wearing

We will measure mask-wearing via direct surveillance. Since vendors will be surveilled three times in addition to baseline, we will compute p_i, the probability that a vendor is wearing a mask that covers the nose and mouth when surveilled (after baseline). We will then estimate:

$$E(p_i|D_i, \xi_m) = \beta_{ind} D_i + \xi_m \tag{5}$$

where D_i is an indicator for whether vendor i was treated and ξ_m are market fixed effects.

3.5 Community and Individual Intervention: Physical Distancing

Using analogous methods, we will estimate the impact of the intervention on the probability that wearing a mask influences physical distancing (being at least an arm's length away from all other people at the time of surveillance).

3.6 Community Intervention: Primary Outcome, IV

To estimate the impact of each percentile increase in mask use, we will also estimate an instrumental variable (IV) version of our primary-specification, where

Y_{ij} is regressed on p_j (the fraction of people wearing a mask over the nose and mouth), instrumenting for p_j with T_j and controlling as usual for x_j. We will again cluster our standard errors at the village-level. The coefficient on p_j will tell us the estimated reduction in symptomatic seropositivity that results from each percentage point increase in mask use. This is an individual-level regression, using the same sample as our primary regression.

3.7 Community Intervention: Other Auxiliary Outcomes

We will run both the intent-to-treat and IV regression for a number of auxiliary outcomes in addition to the primary outcome of symptomatic seropositivity. These auxiliary outcomes include:

- Seroconversions among a random cohort (25,000) tested at both baseline and endline, whether they are symptomatic or not.
- Proportion symptomatic for COVID-19 (based on self-reported symptoms; regardless of seropositivity)
- Proportion with influenza-like illness (based on self-reported symptoms)
- Fraction of pharmacy customers that purchase medicines for fever, cough/sore throat, headache/dizziness, muscle aches/fatigue, nasal congestion/runny nose, or antibiotics
- Hospitalizations and mortality (in both the village-level and individual-level experiments)

Following WHO definition, we define influenza-like illness is as fever and cough in the past 7 days. Following (alternative) WHO guidelines, we define symptomatic COVID-19 as any of the following

- Fever and cough in the past 7 days
- Any three or more of the following: fever, cough, general weakness/fatigue, headache, myalgia (muscle aches), sore throat, coryza (nasal congestion or runny nose), dyspnoea (shortness of breath/difficulty breathing), anorexia (loss of appetite)/nausea/vomiting, diarrhea, altered mental status
- Anosmia (loss of smell) or ageusia (loss of taste)

In defining COVID-19 in this way, we assume that:

- All people live or work in an area with high risk of transmission of virus (meeting the epidemiological criteria for suspected case of SARS-CoV-2)
- All people have been a contact of a probable or confirmed case of COVID-19 or are linked to a COVID-19 cluster (meeting the additional requirement for the presence of clinical criteria to result in a probable rather than suspected case)

4 Subgroup Analyses

We will also consider several subgroup analyses. We will run all of the above specifications in each of the following subgroups:

- Gender (restricting only to surveyed men or women)
- Age in each decade, 18-29, 30-39, 40-49, ... up to 90+, as well as 70+ (pooling all in this age group)

In the village-level experiment, we will also consider subgroups of villages based on:

- Population density, separating the village-level experiment by villages with above or below-median population density
- Baseline mask wearing (above/below-median)
- Distance of village to closest nearby city (population at least 100,000) (above/below-median)
- Estimated village-wealth based on administrative data for that upazila (above/below-median)

References

Correia, Sergio, "Linear models with high-dimensional fixed effects: An efficient and feasible estimator," *Unpublished manuscript, http://scorreia.com/research/hdfe. pdf (last accessed 25 October 2019)*, 2017.