Courant Research Centre
'Poverty, Equity and Growth in Developing and Transition Countries: Statistical Methods and Empirical Analysis'

Georg-August-Universität Göttingen
(founded in 1737)

Discussion Papers

No. 289

The Effect of Personalized Health Information on Preventive Behavior amongst Risk Groups: a Randomized Experiment in Pakistan during the COVID-19 Pandemic

Lisa Rogge, Sheraz Ahmad Khan, Zohaib Khan, Muhammad Jawad Noon, Andreas Landmann and Sebastian Vollmer

December 2021

Platz der Göttinger Sieben 5 · 37073 Goettingen · Germany
Phone: +49-(0)551-3921660 · Fax: +49-(0)551-3914059

Email: crc-peg@uni-goettingen.de Web: http://www.uni-goettingen.de/crc-peg

The Effect of Personalized Health Information on Preventive Behavior amongst Risk Groups: a Randomized Experiment in Pakistan during the COVID-19 Pandemic

Lisa Rogge[4,5], Sheraz Ahmad Khan[1], Zohaib Khan[2], Muhammad Jawad Noon[3], Andreas Landmann[4], Sebastian Vollmer[3]

[1]The University of Edinburgh, [2]Khyber Medical University, [3]Georg-August-Universität Göttingen, [4]Friedrich-Alexander Universität Erlangen-Nürnberg, [5]Leibniz University Hannover

Abstract

Preventing infections is crucial for population groups that are at higher risk to experience a complicated disease course and have limited access to healthcare. Our research with low-income households from Pakistan first documents gaps in knowledge and individual preventive practices in the context of the COVID-19 pandemic, despite pervasive public information campaigns. Second, using a randomized experiment, we evaluate whether a more targeted and personalized SMS information campaign exploiting administrative health records could contribute to narrowing this gap. We find that the intervention helped the at-risk population to adhere to higher levels of handwashing in the time between the first and second wave of infections, and all message recipients were more than twice as likely to use tele-medical services compared to the control group.

Keywords: COVID-19, health insurance, information campaign, risk group behavior

Study pre-registration: This study is registered in the AEA RCT Registry and the unique identifying number is: "AEARCTR-0006307"

1. Introduction

Preventing infectious diseases is particularly important for those who would suffer most from it, and for fragile health systems that are not equipped to attend to a high number of patients with a complicated disease course. These considerations have received considerable attention during the recent COVID-19 pandemic. To contribute to the adoption of individual preventive measures against COVID-19, the Government of Pakistan has issued detailed recommendations for preventive actions for its population and spread this information through various channels. Nevertheless, our telephone survey with low-income households in the province Khyber Pakhtunkhwa revealed that in early stages of the pandemic (April-June 2020) there were gaps in knowledge and uptake of COVID-19 preventive practices. Contrary to the expectation that individuals with a higher risk for a severe disease progression have higher returns to prevention, we did not find higher levels of preventive knowledge and practice in households with a household member who belongs to a group with elevated risk. According to official guidelines of the Government of Pakistan, main risk group indicators are age above 60 or a chronic pre-condition, i.e. cardio-vascular diseases, respiratory diseases, cancer, diabetes or hypertension (Government of Pakistan 2020).

Based on these findings, and in collaboration with the local public health service, we designed a text messaging campaign with the aim to reduce knowledge gaps and to increase preventive behavior in the population at risk. Implementing the intervention through the local public health insurance allowed a more targeted and personalized intervention that could be a viable complement to other information campaigns. The effectiveness of the intervention was tested via a randomized controlled trial. The intervention consisted of a set of six informative text messages, which were sent to a random subset of the health insurance beneficiaries over the course of five consecutive days in August and September 2020. We assess two main and two supplementary hypotheses: First, we test whether the intervention had an effect on the adoption of preventive practices (number of preventive practices, handwashing, wearing masks and using telemedicine) in the whole sample. Secondly, we consider the sub-samples of households with and without a member in the risk group separately to see whether the intervention is more effective in the risk group as they might become more aware of their higher individual return to adopting preventive measures. To further explain these main hypotheses, we test two secondary hypotheses: Within the risk group, we test whether making the individual risk more salient via risk personalization can make the messages more effective. Lastly, we examine whether the main effects are driven by improved knowledge about individual risk and prevention practices.

We find that the intervention increased the reported uptake of individual preventive practices. More specifically, it increased the uptake of handwashing by 6 percentage points, which is an

18% increase relative to the control group uptake of 47%, and telemedicine usage in case of a health need by 5 percentage points, compared to 2% usage in the control group. No such effect could be detected for wearing masks. The effect on the number of preventive practices and handwashing is driven by households with a member who belongs to the COVID-19 risk group. In the risk group alone, handwashing uptake increased by 9 percentage points, while no effect can be detected in the non-risk group. We find evidence for higher telemedicine treatment effects among risk group households who have received messages with a risk personalization. As we do not detect changes in knowledge after the intervention, this does not seem to be the main channel for the observed impact. Apart from the experimental outcomes, we show descriptively the potential of scaling up the intervention using the enrollment and claims data of the health insurance program.

The role of information and awareness has long been acknowledged in the uptake of preventive health behavior, which remained widely under-used in LMICs before the pandemic (Kremer et al. 2019). Information provision has the potential to boost it either by providing new information and updating beliefs (e.g. Dupas (2011), Brown et al. (2017), Madajewicz et al. (2007)), or making existing information more salient via reminders (e.g. Busso et al. (2015), Pop-Eleches et al. (2011)). With increasing mobile-phone coverage, phone-based interventions (Aker 2017) and text messages in particular have been widely used as means to provide both functions. Systematic reviews like Hall et al. (2015) on health behavior in general, Armanasco et al. (2017) on preventive health or a multi-arm study on vaccination uptake in the US (Milkman et al. 2021) show overall small, but meaningful effects of text messaging interventions and provide best practices for message design.

As mobile-phone based interventions are low-cost tools, they have been widely adopted by governments and NGOs during the COVID-19 pandemic, which also led to an upsurge in experimental impact evaluations that are related to ours. Early in the pandemic, Banerjee et al. (2020) found that broadcasted SMS with links to celebrity-endorsed videos increased the uptake of handwashing and reporting of COVID-19 symptoms. An increase in handwashing was also detected for a more generic prevention information intervention via SMS in Peru in June 2020 (Boruchowicz et al. 2020). None of these interventions had an effect on social distancing. Another messaging intervention during the peak of the first wave in the Indian state of Bihar did not lead to more handwashing either (Bahety et al. 2021).

Our study has two main contributions to the literature: First, we contribute to the literature on solutions to shield COVID-19 risk groups. Such evidence remains scarce for LMICs, which have lower health system capacities and rarely have the opportunities to target risk groups directly like for instance in the United Kingdom or other high-income countries (Burd and Coleman 2020). By including both age-based and precondition-based risk factors, we take a population-based perspective in contrast to other studies that focus on specific disease groups such as people with diabetes (Dizon-Ross et al. 2020). Moreover, we extend the list of preventive practices by the use of telemedicine which is particularly relevant for the risk group. Secondly, we contribute to the text messaging literature more broadly as relying on health insurance enrollment and claims data allows us to combine scalability and personalization of the intervention. On the one hand, scalability is possible via broadcasting of telecommunication providers like in (Banerjee et al. 2020), but in a non-emergency context when the information is not relevant to everyone this poses the risk of an overflow of information and less attention to relevant messages. On the other hand, personalization has been found to enhance the effectiveness of such interventions (Head et al. 2013), which was in the past often done via a pre-intervention contact, and makes the intervention more costly and less scalable. Targeting and personalizing messages through sparse but potentially sufficient information in administrative data could therefore combine the strengths of both approaches. Birth registers as used for contacting, but not personalizing in India (Bahety et al. 2021), city records in Peru (Boruchowicz et al. 2020) or NGO records in Bangladesh (Siddique et al. 2020) have similar advantages, but cover more geographically limited areas or specific population groups. With increasingly digitized health systems additional applications beyond COVID-19 are likely to emerge.

2. Context

2.1. Policy and societal context

Our study is set in the Pakistani province Khyber Pakhtunkhwa (KP). Even before the pandemic, Pakistan's and particularly KP's health systems were fragile (Asian Development Bank 2019). More recently, the provincial government of KP has implemented several reforms such as the Social Health Protection Initiative providing free inpatient health insurance (Government of KPK 2010). Nevertheless, a review of the provision of higher level inpatient care conducted in fall 2019 flagged substantial gaps in the availability of material, trained staff and management capacities (Asian Development Bank 2019), which are essential in caring for patients with a complicated COVID-19 disease course. From the onset of the pandemic, KP recorded high infection rates and a high case fatality rate (Anser et al. 2020). The high case fatality rate could be a consequence of a larger share of undetected cases or worse treatment of detected cases.

In response to the outbreak of the pandemic, the federal government of Pakistan issued the *National Action Plan for The Corona Virus Disease* (Ministry of Health 2020). Cell-broadcasting of text and voice messages on preventive measures and symptoms was an integral part of this strategy. The initial plan did not include risk-group specific policies, but paved the way for targeted recommendations, which were published shortly after (Government of Pakistan 2020). For the elderly and people with certain preconditions (cardio-vascular, respiratory diseases, cancer, diabetes and hypertension), recommendations stressed the adherence to common preventive practices to avoid an infection. Special attention was given to care for preconditions as poorly controlled preconditions intensify the risk of a complicated disease course (Coppelli et al. 2020) and the use of tele-medical services where possible. The recommendations also stressed that caregivers and other household members should apply more caution.

Considering Pakistan’s demographic situation with a rather low life expectancy at birth (67 years) and only 4.3% of the population over the age of 65 (World Bank 2021), the share of the population in the risk group for a complicated COVID-19 disease course is estimated to be lower than the global average (around 17% according to Clark et al. (2020)). Considering the low health system capacity to cater even to a low number of severe cases in combination with a culture of large multi-generational households aggravates the burden. Our survey data with KP’s low-income population shows that around two thirds of households have at least one member that is in the risk group (Table A 19). 60% of households have a member over the age of 60, and 26% of households have a member with one of the five preconditions. Hypertension is the most common precondition in our sample, followed by diabetes and other cardio-vascular conditions, while respiratory diseases were only reported in 2% and cancer in 1% of households (Table A 19).

2.2. State of COIVD-19 knowledge, practice and information campaign exposure

As depicted in Figure 1, the trial is embedded in a larger study, was preceded by three survey waves during an earlier stage of the pandemic and is itself set in the time between the first and second wave of the COVID-19 pandemic in Pakistan.

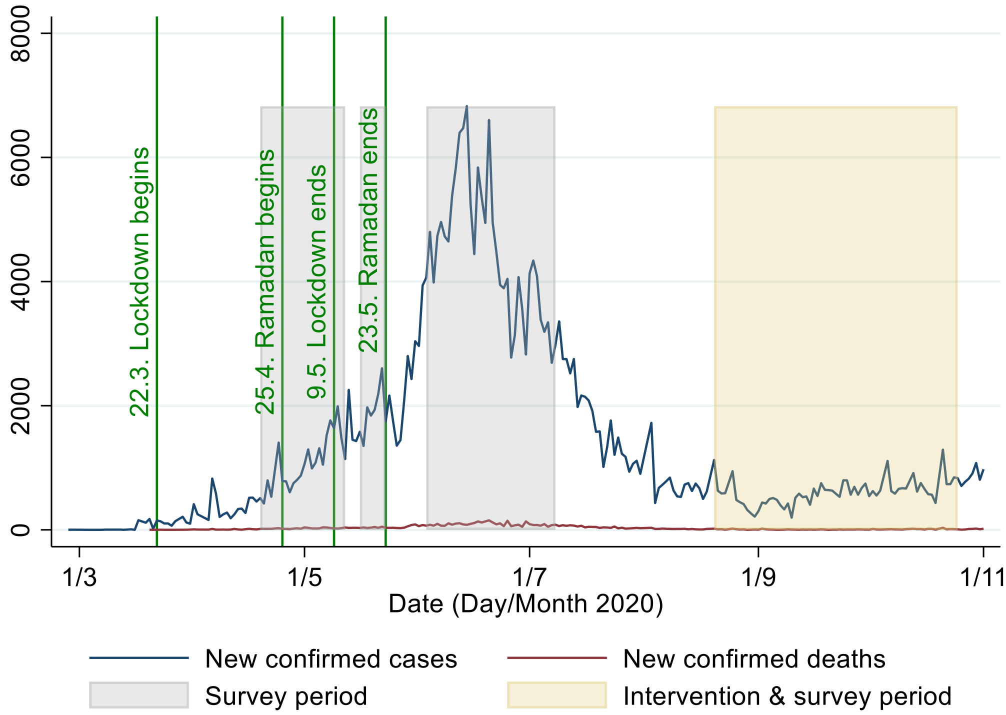


Figure 1 Study timeline including Pakistan-wide daily new COVID-19 cases and deaths as well as major events (case and death data from (Hale et al. 2020)), see Table A 1 for details.

The rapid response survey generated initial insights into the target population's knowledge, attitudes and behavior pertaining to the COVID-19 pandemic[1]. We found that from the beginning of the survey period, there was a high general awareness of COVID-19 and its severity in our study region, but also substantial knowledge gaps about specific preventive practices that did not narrow over time. Only half of the respondents could name both fever and cough as symptoms of COVID-19, 80% knew that SARS-Cov-2 can be transmitted through physical contact but only 40% knew that it can also be transmitted via air droplets (Figure A 9). Social distancing was widely known as preventive method, wearing masks was initially only known by about half of respondents and also hygiene measures such as handwashing were only named by less than half of the respondents. Around 60% of respondents were aware of old age being a risk factor, only 20% mentioned any precondition as a risk factor and over 30% falsely mentioned children as a risk group. It stands out that respondents with at-risk household members do not display substantially different knowledge or preventive practice compared to respondents without a household member in the risk group (Table A 20).

The majority of respondents relied on other people and television for information on COVID-19. Internet and newspaper only play a considerable role among people with higher education (Figure A 7). Around 75% of the 250 interviewees in the last weeks of pre-intervention data collection confirmed that they had received some information on COVID-19 through their

[1] See appendix 7.2.2 for a description of the data collection, which is very similar to the post-intervention survey and appendix 7.2.3 for more detailed results.

mobile phone. Out of those, around half report to have received information on a daily basis. Around the same number of respondents report to have received government SMS, but with a lower frequency (Figure A 8).

3. Experimental details

3.1. Experimental set-up

We test as primary hypotheses whether the intervention had any effect on the uptake of preventive practices (hypothesis 1) and whether this effect was larger in the risk group (hypothesis 2). As secondary hypotheses, we test the effect of personalized messages within the risk group (hypothesis 3), and whether effects work through an increase in knowledge (hypothesis 4). These main analyses follow the registered pre-analysis plan (Khan et al. 2020).

The experimental design is depicted in Figure 2. The sample can be divided into households with and without a household member in the risk group. Two thirds of the sample received an intervention and one third did not receive any intervention. In the risk group, there are two treatment arms: half of the treated households received a risk-personalized intervention and half received generic messages. In the non-risk group, all treated households received generic messages.

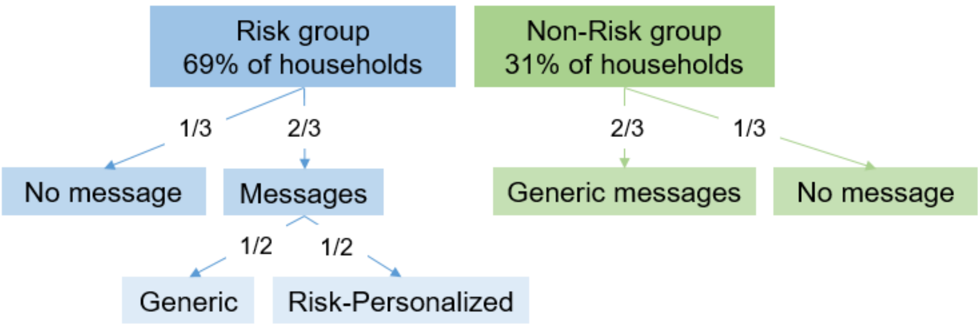


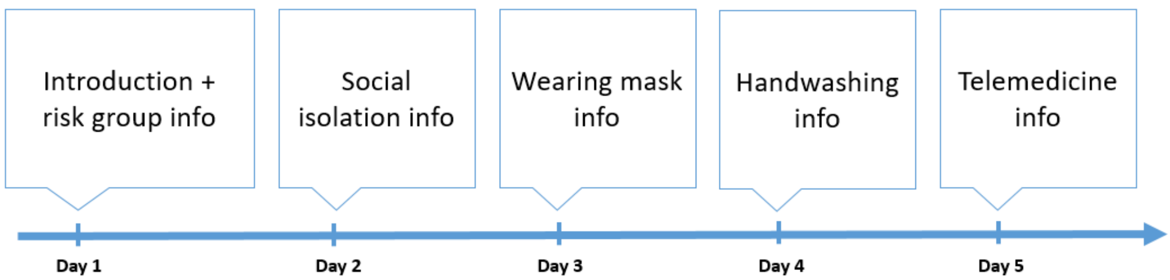
Figure 2 Experimental design

3.2. Intervention

The intervention was an information campaign of the Social Health Protection Initiative (SHPI) in the province of Khyber Pakhtunkhwa. It consisted of a set of six informative text messages that were sent to the selected recipients over the course of five days. The information content reinforces the government of Pakistan's specific recommendations for COVID-19 risk groups (Government of Pakistan 2020), which reflected the current state of knowledge on COVID-19 risk groups and prevention. As depicted in Figure 3, on the first day, an introductory message was sent before the first information message on risk groups. On each of the following days, one information message on either social distancing, wearing masks, handwashing or using telemedicine before visiting a doctor was sent. In addition to the main information content, each message contained elements that have previously been found to enhance the effectiveness of such interventions. First, the sender mask of the message was “Sehat Insaf Card”, which is the name of the SHPI’s health insurance program that all recipients are beneficiaries of and can therefore be considered a well-trusted sender of health-related messages. Second, the main cardholder, who is likely also the main decision-maker in the household, was addressed by name, which is a second trust-building element as well as a means to increase relevance. Third, on every day, the recipient was provided with a telephone number that s/he could call in case of further questions. On most days, this was the number of a helpline that normally consults (potential) health insurance beneficiaries on enrollment and card usage related queries and would either provide the caller with basic information or re-direct him/her to a telemedicine helpline in case of a medical query. The information message on telemedicine directly contained the telemedicine helpline number. All messages were sent in Urdu language with Latin script (as listed in Table A 2) as the majority of the study population was either literate in Urdu language themselves or had another family member who could read the message to them (see Table A 4). Each message also contained the call to “tell your family” about this message as it was directed towards the main cardholder, but relevant for all household members.

In addition to this general message specification, a subgroup of those with at least one household member in the risk group received a more personalized version of the risk group information message. Personalization was reached by listing risks first that were known to be present within the household. All risk groups that were not specific to the respective family were then listed with decreasing frequency. The distribution of messages in the respective order was then also applied to the groups that received a generic risk message to ensure comparability of the groups except for the personalization.

Figure 3 Intervention timeline



The messages were sent through the Telenor bulk messaging portal by the helpline company ICU healthcare, which provides an infrastructure to launch awareness campaigns for the SHPI and is part of the research team. Selection into the sample and treatment allocation was only known to the research team. As there was no explicit baseline data collection, participants were fully blinded to treatment assignment prior to the intervention and were unaware of the existence of a treatment and a control group. Those who were interviewed before, consented to being contacted by us again, but did not receive any specific information on text messages. The interviewers of the post-intervention survey were also unaware of treatment allocation and posed the same questions to all respondents.

We see that around 40-50% of treatment group respondents remember receiving our messages, and examine barriers to receiving and reading the messages in section 4.3.

3.3. Data

Sample selection

The sampling frame for the trial consists of the list of households that were enrolled in the Sehat Sahulat Program up until 2019 as provided by the SHPI. Eligibility to enrollment for the program is restricted to the poorest 69% of the population based on the household poverty score that was collected as part of the PMT census in 2010. Between 2015 and 2019, 1.5 million out of the 2.4 million eligible households have been enrolled in the insurance. Appendix Table A 3 displays that the enrolled households (or their designated main cardholders) are on average less wealthy, slightly older, to a higher proportion male and married than the general eligible population. As this study contains a mobile-phone based intervention, the sampling frame was restricted to the almost 0.6 million households for whom there is a unique phone number in the records[2]. Within the household, we aimed to interview the main insurance card holder, which was successful in over 75% of the interviews (Table A 4). A household was excluded if the main cardholder was not member of the household anymore.

The intervention sample is derived from a combination of households that were already interviewed during a previous survey round and an additional draw from the sampling frame (see Figure A 1 for a graphic display of the composition of the sample from the sampling frame until the estimation sample). The previous interview as well as the additional intervention sample were selected following the same procedure. The main part of the sample was then drawn from the list of households with unique phone numbers, stratified by district to ensure representativeness of all regions of the province based on the proportion of their enrolled population. Furthermore, households with previous insurance claims that are likely to indicate an increased risk for a complicated COVID-19 infection were over-sampled to gain sufficient observations from this population group that usually only comprises of 5% of households.

[2] For the majority of the remaining (66%), there is either no phone number in the records or a clearly wrong number (e.g. not sufficient or too many digits).

Finally, 1,769 households with a previous interview and complete information on self-reported risk were included in the intervention sample, as well as 27,229 households without a previous interview.

Randomization

Treatment assignment was done just before the launch of the intervention by the authors at the individual level by assigning random numbers with the function *runiform* in Stata 15. Treatment was first assigned in the previously interviewed sample. In the risk group, one third of households were randomly allocated to the control group, one third to the personalized and one third to the generic treatment arm. The distribution of exact risk messages (the order of the mentioned risk groups) was determined by the prevalence of risk groups in the personalized treatment arm and then applied to the generic treatment group in both samples. In the additional sample, we randomly allocated two thirds to the treatment group again mirroring the distribution of risk messages in the personalized group.

Analysis sample, balance and attrition

The sample in the post-intervention survey comprises of 2,382 respondents, among which 306 respondents are from the previous interview sample and 2,077 from the additional sample (Figure A 1). The sample characteristics are displayed in Table A 4. Treatment and control group characteristics were balanced at randomization (Table A 5), and among post-intervention survey respondents except for a slightly higher age in the control group (Table A 6). As displayed in appendix Table A 7 to Table A 11, there is no differential attrition between treatment and control group. It needs to be mentioned that there are detectable differences along the sparse administrative data characteristics between the attrited and the interviewed (Table A 8) in the additional sample, but not among previously interviewed respondents (Table A 9).

Conducting a survey during a pandemic made some deviations from the survey protocol necessary. We present the results from a restricted sample that is closer to the intended protocol. This restricted sample includes all respondents who were interviewed up to one week after the intended interview date (for the treatment group, this is two weeks after the end of the intervention) and excludes the last week of data collection. As outlined in appendix section 7.2.2, we had intended to interview message recipients around one week after receiving the last intervention message. The second deviation was an extension of the data collection period. We had intended to complete the data collection within one month to keep contextual factors such as the progression of the pandemic rather constant, but only reached the stopping rule (reaching the intended sample size) after two months of data collection. It stands out that in the last week of interviews, the sample characteristics are not as clearly balanced between treatment and control group as in the remaining survey period (Figure A 2, Figure A 3). In addition, the end of the data collection period falls into a time of the beginning of the second pandemic wave.

3.4. Estimation strategy

Following the experimental design with random treatment assignment, we use OLS regression models to compare the outcomes of treatment and control households in an intention-to-treat analysis. The outcome measures, risk group indicators, treatment and control variables are defined as follows, and as specified in the pre-analysis plan (Khan et al. 2020).

Outcome measurement

The main outcome is uptake of preventive practices, which is measured in two ways. First, a count index captures the number of different practices that were mentioned in the messages (physical distancing, handwashing, wearing masks, telemedicine usage). Secondly, we use the individual binary indicators for the uptake of handwashing, wearing masks and telemedicine usage. The uptake of handwashing and wearing masks is self-reported in an unaided recall question. Though measured in the same way, the individual physical distancing indicator is not included as uptake was already very high in the pre-intervention survey. Telemedicine usage is also measured as a survey-based indicator derived from a question about calling a doctor or telemedicine helpline for a health need in the family during the previous month. Consequently, the sample for the telemedicine usage outcome only includes the 21% of the sample who reported to have had any health need in the household during the previous month. Additionally, we pre-registered an alternative measure of telemedicine usage capturing the number of calls to the telemedicine helpline during two months after the intervention as derived from the helpline's call records. As too few calls from the study population could be identified in the records[3], these are only studied descriptively.

The secondary outcome is knowledge about risk groups and preventive practices. Both indicators are measured using unaided recall questions of which COVID-19 risk groups and preventive practices the respondent can name. For both risk groups and preventive practices, we derive a count index, which captures the number of correctly named elements that were part of the messages (0-2 for risk groups and 0-4 for preventive practices).

[3] As the helpline did not record each caller's national identification number for privacy reasons, only households that called the helpline with the same telephone number that is noted in their health insurance enrollment data could be identified.

Risk group definition

Every household that reports[4] to have at least one member over the age of 60 and/ or a member with a relevant precondition (cardiovascular or respiratory disease, cancer, diabetes or hypertension) is defined as a risk group household. As only one member of the household is the respondent, this information is collected from him/ her representing the household.

Treatment

Treatment is measured by assignment. For hypotheses 1 (prevention uptake), 2 (risk group heterogeneity) and 4 (knowledge) it takes value 1 if we sent the intervention to the household and 0 otherwise. For hypothesis 3 (personalization), the treatment variable takes value 1 if the risk group household was sent a personalized risk message and value 0 if it was sent a generic risk message.

Control variables

The main specification does not include any covariates. In an alternative specification, we add respondent's age in years, an indicator for being female, three categories of completed education (up to primary as reference, secondary and tertiary) as reported in the survey. As a measure of wealth, we use the proxy means test (pmt) score, which is a continuous wealth measure that was calculated for each household in a census in 2010, and reported in the insurance data as the poverty line for health insurance eligibility is also based on this score.

Regression specification

We estimate the intention-to-treat effect on practice and knowledge outcomes (hypotheses 1 and 4) using the following framework:

$$(1)\ Y_i = \alpha + \beta treat_i + \varepsilon_i \ \ (+\theta\ C_i)$$

Y_i is the respective outcome variable (i.e. preventive practice index, binary indicators of mask wearing, handwashing and telemedicine usage, risk and prevention knowledge indices) for household *i*. In addition to the basic specification that regresses the respective outcome on a treatment dummy $treat_i$ (assigned to receive the intervention), we also estimate a second specification that includes basic control variables C_i (age, gender, education, wealth):

$$(2)\ Y_i = \alpha + \beta treat_i + \gamma risk_i + \delta(treat_i * risk_i) + \varepsilon_i \ \ \ (+\ \theta C_i)$$

[4] For the subset of households that was already interviewed in a pre-intervention survey wave, we use the risk group information from the first interview in case it differs at endline as this influenced the randomization. Results do not change when using the endline risk group information.

To test for the difference in the treatment effect between risk and non-risk group (hypothesis 2), equation 2 is used to estimate the interaction effect between the binary risk group indicator $risk_i$ and the same treatment indicator as above.

$$(3)\ Y_i = \alpha + \beta personalized_i + \varepsilon_i$$

The treatment effect of the personalized messages (hypothesis 3) is estimated using equation 3 to compare the outcomes of personalized against generic message recipients among the treated in the risk group. The probability of receiving a personalized message by assignment in the previously interviewed sample differs from receiving a personalized message by chance in the additional sample. Therefore, the estimates of each risk group are re-weighted using a propensity score that reflects the likelihood of receiving a personalized message depending on the kind of risk group and whether the households was part of the previously interviewed or the additional sample.

As a robustness check, p-values of the primary hypotheses (H1 and H2) are adjusted for multiple hypotheses testing using the Benjamini-Hochberg method (Benjamini and Hochberg 1995).

4. Results

4.1. Treatment effect on the uptake of preventive practices

Main results

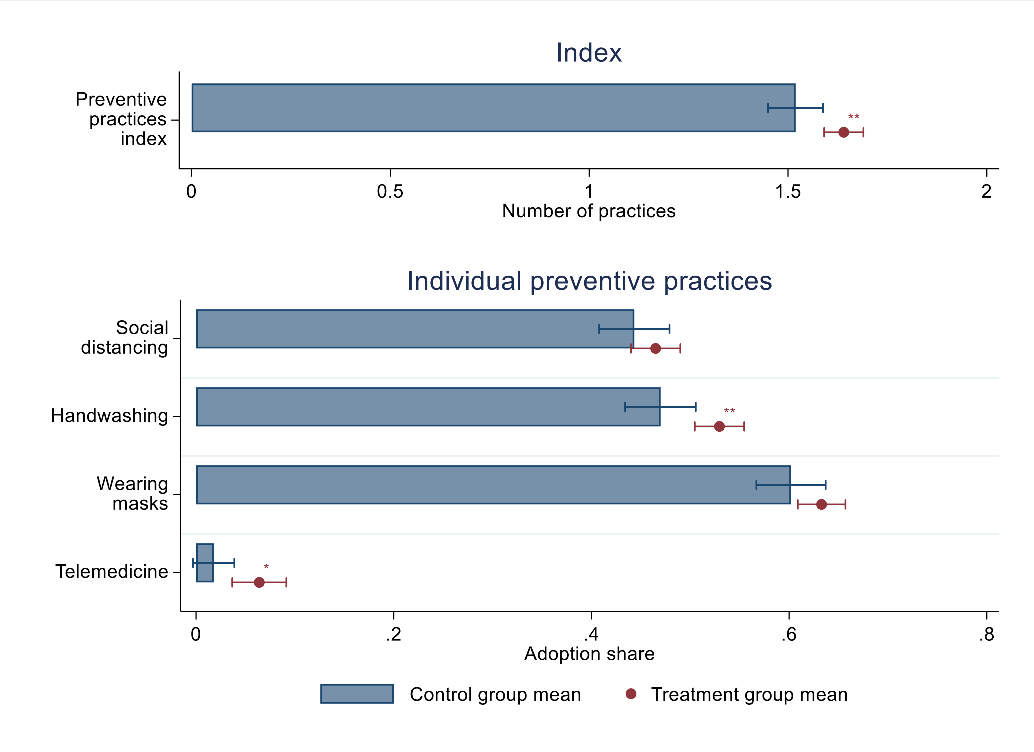
We find that the intervention increased the uptake of preventive practices. Figure 4 shows that control group respondents practice on average 1.5 out of the four preventive practices that were mentioned in the messages, which increases by around 0.12 (8%) in the treatment group. Out of the individual practices, most control group respondents (60%) report to adhere to regular mask-wearing, 47% to handwashing, slightly less continue to practice social distancing, and out of those who had an illness in the household in the previous month only 2% made use of telemedicine. The average treatment group uptake is higher in all index elements, but it is only significantly distinguishable from control group uptake for handwashing[5] and telemedicine usage. For handwashing, the treatment effect is around 6 percentage points (13% increase relative to the control group), and a 5 percentage points increase in telemedicine usage when ill almost triples the control group uptake (Table A 12). The main effect on the preventive practices index and handwashing uptake is robust to multiple hypothesis testing adjustments,

[5] When looking at the uptake of preventive practices over time, we see that on average the number of preventive practices adopted by the control group, and handwashing and social distancing in particular, decreased in the intervention period (Figure A 5). Consequently, the treatment effect on handwashing uptake comes from the treated maintaining higher levels of prevention.

but the adjusted q-value of the treatment effect on telemedicine usage increases to above 0.1 (Table A 17).

Considering the call record of the telemedicine helpline that was mentioned in the messages, only 23 calls can be attributed to the intervention sample, and all are from the treatment group. Most callers ask about the messages, and two of them ask for advice regarding a specific health complaint. All these calls were made directly after receiving the intervention, and we do not detect any longer-term effects over the following three months (until November 2020). This shows that the effect on the uptake of telemedicine is not driven by calling the helpline, but rather by calling health workers or health facilities directly.

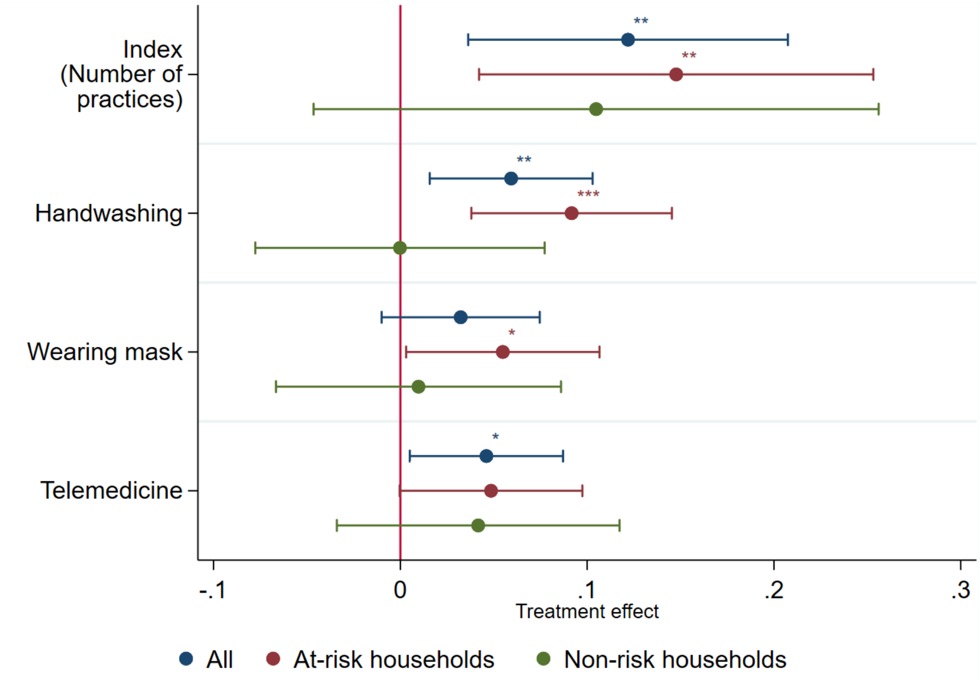
Figure 4 Treatment effect on preventive practice uptake



*Control group (bars) and treatment group (dots) means of main outcomes with 90% confidence intervals. Stars indicate the p-value of regressing the respective practice indicator on the binary treatment indicator following equation 1, * p < 0.1, ** p < 0.05, *** p < 0. (see Table A 12 for tabular display)*

Estimating the heterogeneous treatment effect for households with and without a member in the risk group shows that the effect on handwashing uptake is significantly higher in risk group households. This difference is not robust to multiple hypothesis testing (Table A 18) and the specification with control variables (Table A 13). Nevertheless, Figure 5 and Table A 14 show that respondents from risk group households (red) drive the whole sample treatment effect for the prevention index and handwashing, while no treatment effect can be detected in the non-risk sample (green) alone. In the risk group only, regular handwashing to prevent the spread of COVID-19 increased by over 9 percentage points and mask wearing increased by around 5 percentage points. The uptake of telemedicine is similar across risk and non-risk group.

Figure 5 Heterogeneous treatment effect across at-risk and non-risk households



*Treatment coefficients from estimating equation 1 in the complete sample, and for at-risk and non-risk households separately. Stars indicate the p-value of regressing the respective practice indicator on the binary treatment indicator following equation 1, * p < 0.1, ** p < 0.05, *** p < 0. See Table A 12 and Table A 14 for the tabular display.*

4.2. Mechanisms

Personalization

We find that receiving a personalized message rather than a generic one increases the likelihood to call a doctor or telemedicine helpline before visiting a health facility by 7 percentage points (Table 1). For the other preventive practices, we cannot detect a significant difference in the treatment effect between recipients of personalized and generic messages.

Table 1 Personalization treatment effects on preventive practices (hypothesis 3, treated risk group only)

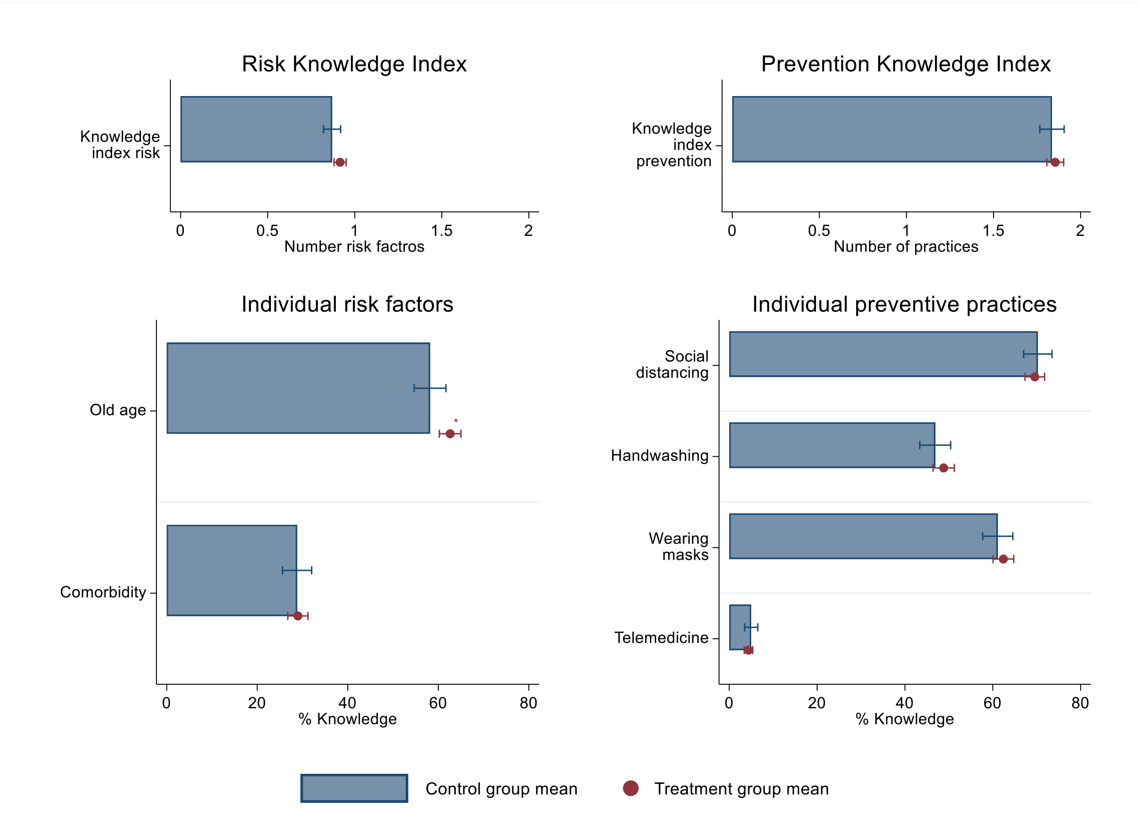
	Preventive practices index	Handwashing	Wearing masks	Telemedicine if sick
Personalized vs generic treatment	-0.0023	0.0460	-0.0354	0.0728*
	(0.0787)	(0.0401)	(0.0382)	(0.0392)
Observations	706	706	706	168

*Estimation results of equation 3 with the respective preventive practice on the binary personalization indicator, re-weighted taking personalization probability into account. Sample is restricted to the treatment group with at least one risk group member in the household. Personalization entails listing the household-specific risk factor first in the risk message. The preventive practices index is a count-index ranging from 0 to 4 counting whether the respondent mentioned to practice handwashing, social distancing or mask wearing on a regular basis and/ or using telemedicine in case of a health need. Standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01.*

Knowledge

We do not find that the increases in the uptake of preventive practices are mainly driven by closing information gaps in risk and prevention knowledge (hypothesis 4). As depicted in Figure 6, around 58% of the control group respondents knew that old age implies a higher risk and only 29% knew a relevant precondition. Different from the patterns in prevention uptake, the best known preventive practice was social distancing (70%), followed by wearing masks (61%) and handwashing (47%). We see neither an effect of the intervention on the aggregate risk and prevention knowledge indices nor on the individual items of the prevention knowledge index. Due to the precision of these estimates and the robustness of the null effect across specifications (Table A 15, Table A 16), we rule out that the intervention increased prevention and risk knowledge sufficiently to explain the effect on prevention uptake.

Figure 6 Treatment effect on risk and prevention knowledge



*Control group (bars) and treatment group (dots) means of main outcomes with 90% confidence intervals. Stars indicate the p-value of regressing the respective knowledge indicator on the binary treatment indicator following equation 1, * p < 0.1, ** p < 0.05, *** p < 0.01 (see Table A 15, Table A 16 for tabular display).*

4.3. Scale-up potential

As we find the intervention to be effective, we further examine its scale-up potential. The intervention could be scaled up in the same way as it was implemented for this study, except for the risk group targeting and personalization, for which we relied on survey-based information. As we see that targeting and potentially also risk-personalization are effective components of the intervention, we first assess the scale-up potential of the intervention through administrative data alone. Secondly, it is inherent to low-touch text messaging interventions that only a subset of assigned message recipients is exposed to the intervention. Therefore, we examine descriptively factors that hinder message recipients from receiving, reading and understanding the messages to understand how the effectiveness of the intervention can be improved.

Risk group identification in the administrative data

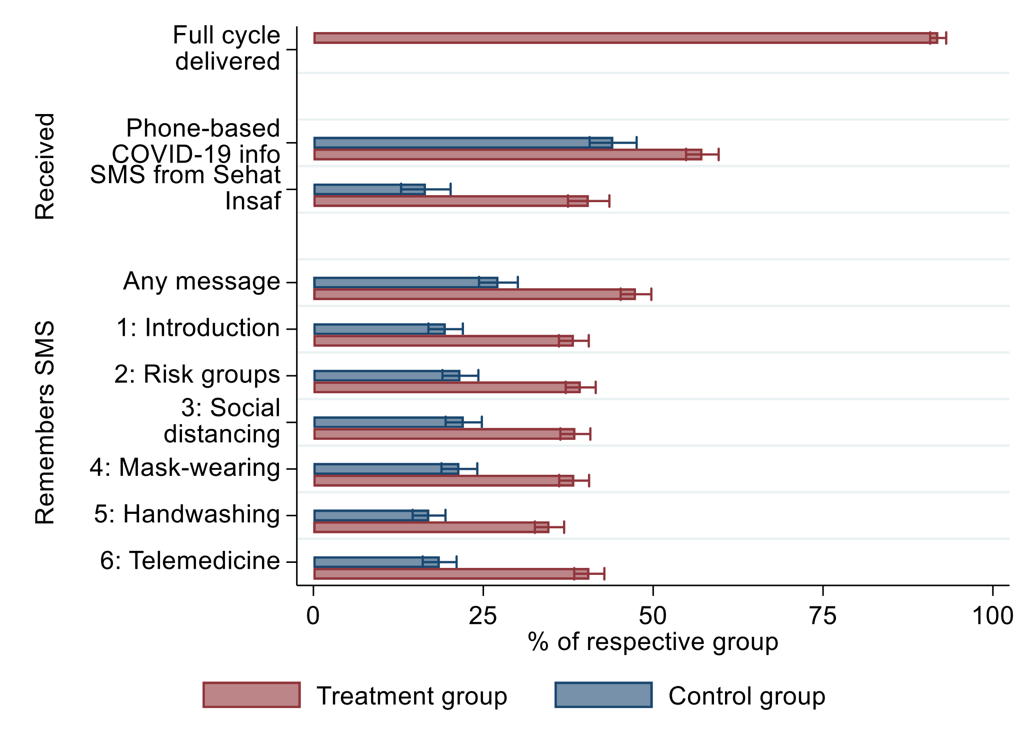
Age and precondition risk group information can also be derived from the health insurance enrollment and claims data. Age can be identified in the enrollment data via the age of the main cardholder, but the age of all family members would only be available through another national database. Preconditions can be derived from the insurance claims on the main cardholder's ID. Each of these claims is recorded with one of 829 treatment categories, which allow to identify previous treatments that point towards a precondition that might increase the risk for a complicated COVID-19 progression. By design of the health insurance scheme, these treatments only include inpatient or maternity related care and therefore do not include all potentially relevant preconditions.

Comparing this risk group identification to the information from the survey data shows that age-based risk status was correctly identified for almost 70% of survey households, but precondition-based risk is, with the data that is available to us, subject to substantial error (see Figure A 6 for details).

Intervention exposure

Around 90% of messages were correctly delivered to interview respondents (Figure 7). This share could be further improved if databases were updated more frequently (less invalid numbers) and network coverage improved (less failed deliveries to correct numbers). 40 to 50% of respondents report to remember specific messages that we sent. As we know that the government, telecommunication companies, NGOs and others sent out more general messages throughout the pandemic, it is not surprising that also some people in the control group report to remember specific messages. The share in the treatment group is around twice as high for all messages. As the intervention was sent out in a geographically sparse manner throughout the province, it is very unlikely that this is an indication of spillovers, but rather that the respondent cannot distinguish messages that s/he has received from other sources from our messages.

Figure 7 Different measures of message receipt shares in the interviewed sample



Treatment and control group means of each binary message receipt indicator with 90% confidence intervals; "Full cycle delivered" is taken from the provider's delivery reports; the remaining indicators are survey-based: "Phone-based COVID-19 info" and "SMS from Sehat Insaf" are based on closed-ended questions whether the respondent remembers receiving any SMS on COVID-19 during the previous 3 weeks and who sent this message; the indicators in the "Remembers SMS" block are based on whether the respondent remembers to have seen the message after being read out loud by the interviewer.

The ability to understand the messages poses an (expected) barrier to being exposed to and reacting to the messages. Less than half of respondents report to be literate in Urdu themselves, but almost 80% of households have at least one member who can read Urdu. Additionally, around 23% of respondents report that they did not even understand the message when it was read to them by an enumerator.

It is possible that messages were not trusted or the recipient did not pay attention to them. From the qualitative interviews after the pilot intervention (appendix section 7.2.3), we know that messages from the health ministry or the health insurance program are generally perceived as trustworthy and mentioning the name of the recipient was also named as a reason for paying attention to the messages. Respondents further stated that the messages were rather perceived as official messages than advertisement.

5. Discussion & Conclusion

In settings with fragile health systems and limited vaccination rollout, such as our study region in Pakistan, avoiding a COVID-19 infection with individual preventive practices remains key, especially for those at higher risk of experiencing a severe disease course. Our randomized experiment shows that a personalized and targeted text messaging campaign delivered to health insurance beneficiaries can be a complimentary measure to increase uptake of preventive practices. The intervention increased the uptake of handwashing by 6 percentage points, an 18% increase relative to the control group. This effect is driven by the at-risk population for whom handwashing uptake increased by 9 percentage points. In the whole sample, the intervention increased the uptake of tele-medical services by 5 percentage points, which almost tripled control group uptake of 2%. Within the risk group, making the household-specific risk more salient via light risk-personalization in the messages makes the intervention more effective for telemedicine usage. The treatment effect on handwashing uptake is comparable in size and direction to Banerjee et al. (2020)'s celebrity-endorsed text and video message broadcasting intervention at an earlier stage of the pandemic in India. Such an effect could not be replicated by Bahety et al. (2021) with a generic, pure text messaging intervention during the peak of the first wave in India.

We do not find evidence for increased knowledge as channel for these effects. This is in line with other text-messaging interventions that do not detect any updates in knowledge early in the pandemic (Banerjee et al. 2020) or during the peak of the first wave in India (Bahety et al. 2021). We suspect that our text message campaign rather narrows the knowledge-action gap by making existing information more salient and helping message recipients to form habits. Message personalization possibly facilitated this.

Our study is subject to several limitations. First, we rely on self-reported outcomes that are susceptible to social desirability. We address this concern by using only unaided recall questions for outcome measurement and ensuring blinding of interviewers and participants to treatment allocation. To keep the interview as short as possible, the response rate high and retain comparability to the pre-intervention survey, we opted to not include additional measurements. Comparing our results to Bahety et al. (2021) gives confidence in the validity of the outcome. They also use the open question as main specification, but test in addition a direct elicitation, asking about the community rather than the individual and a list experiment. Second, it is inherent to telephone surveys that the sample that is reached and willing to give an interview is different from the general population, especially in low-income settings. As opposed to random digit dialing or mere lists of telephone numbers, we can leverage household characteristics from the insurance database to get an idea of what sections of the population our sample adequately represents. Finally, deviations from the planed survey protocol due to pandemic conditions led to a sample size that was lower than we intended based on power calculations. Since the primary treatment effect is robust in the pooled sample as well as in the risk group alone, this is mostly an issue for the effects on risk personalization.

All in all, personalized and targeted SMS campaigns can be effective complements to efforts to shield COVID-19 protect risk groups. It increased the uptake of individual preventive practices, particularly in households that have a member that is at higher risk of experiencing a severe disease course. The intervention was successful in making existing knowledge more salient and encouraged continued adoption. There is potential in scaling up text messaging interventions that make use of the sparse individual information in health insurance records.

6. List of references

Aker, Jenny (2017): Using Digital Technology for Public Service Provision in Developing Countries. In Sanjeev Gupta, Michael Keen, Alpa Shah, Geneviève Verdier (Eds.): Digital revolutions in public finance. Washington DC: International Monetary Fund.

Anser, Muhammad Khalid; Yousaf, Zahid; Khan, Muhammad Azhar; Nassani, Abdelmohsen A.; Qazi Abro, Muhammad Moinuddin; Hinh Vo, Xuan; Zaman, Khalid (2020): Social and administrative issues related to the COVID-19 pandemic in Pakistan: better late than never. In *Environ Sci Pollut Res* 27 (27), pp. 34567–34573. DOI: 10.1007/s11356-020-10008-7.

Armanasco, Ashleigh A.; Miller, Yvette D.; Fjeldsoe, Brianna S.; Marshall, Alison L. (2017): Preventive Health Behavior Change Text Message Interventions: A Meta-analysis. In *American Journal of Preventive Medicine* 52 (3), pp. 391–402. DOI: 10.1016/j.amepre.2016.10.042.

Asian Development Bank (2019): Khyber Pakhtunkhwa Health Sector Review: Hospital Care.

Bahety, Girija; Bauhoff, Sebastian; Patel, Dev; Potter, James (2021): Texts Don't Nudge: An Adaptive Trial to Prevent the Spread of COVID-19 in India. Center for Global Development, Working Paper 585.

Banerjee, Abhijit; Alsan, Marcella; Breza, Emily; Chandrasekhar, Arun G.; Chowdhury, Abhijit; Duflo, Esther et al. (2020): Messages on COVID-19 Prevention in India Increased Symptoms Reporting and Adherence to Preventive Behaviors Among 25 Million Recipients with Similar Effects on Non-recipient Members of Their Communities. In *National Bureau of Economic Research*. Available online at https://www.nber.org/papers/w27496.

Benjamini, Yoav; Hochberg, Yosef (1995): Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. In *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1), pp. 289–300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.

Boruchowicz, Cynthia; López Bóo, Florencia; Finamor Pfeifer, Flora; Russo, Guilherme A.; Souza Pacheco, Tainá (2020): Are Behaviorally Informed Text Messages Effective in Promoting Compliance with COVID-19 Preventive Measures Evidence from an RCT in the City of Sao Paulo. Inter-Americal Development Bank (Technical Note, IDB-TN-2021).

Brown, Joe; Hamoudi, Amar; Jeuland, Marc; Turrini, Gina (2017): Seeing, believing, and behaving: Heterogeneous effects of an information intervention on household water treatment. In *Journal of Environmental Economics and Management* 86, pp. 141–159. DOI: 10.1016/j.jeem.2016.08.005.

Burd, Hannah; Coleman, Cathy (2020): Using behavioural insights to create a Covid-19 text service for the NHS.

Busso, Matias; Cristia, Julian; Humpage, Sarah (2015): Did you get your shots? Experimental evidence on the role of reminders. In *Journal of health economics* 44, pp. 226–237. DOI: 10.1016/j.jhealeco.2015.08.005.

Clark, Andrew; Jit, Mark; Warren-Gash, Charlotte; Guthrie, Bruce; Wang, Harry H. X.; Mercer, Stewart W. et al. (2020): Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. In *The Lancet Global Health*. DOI: 10.1016/S2214-109X(20)30264-3.

Coppelli, Alberto; Giannarelli, Rosa; Aragona, Michele; Penno, Giuseppe; Falcone, Marco; Tiseo, Giusy et al. (2020): Hyperglycemia at Hospital Admission Is Associated With Severity of the Prognosis in Patients Hospitalized for COVID-19: The Pisa COVID-19 Study. In *Diabetes care* 43 (10), pp. 2345–2348. DOI: 10.2337/dc20-1380.

Dizon-Ross, Rebecca; Jayachandran, Seema; Zucker, Ariel (2020): De-biasing people's over-optimism about their health risk. In *AEA RCT Registry*. DOI: 10.1257/rct.5951-1.0.

Dupas, Pascaline (2011): Health Behavior in Developing Countries. In *Annu. Rev. Econ.* 3 (1), pp. 425–449. DOI: 10.1146/annurev-economics-111809-125029.

Government of KPK (2010): Khyber Pakhtunkhwa Health Sector Strategy 2010–2017. Government of Khyber Pakhtunkhwa. Available online at http://dghskp.gov.pk/downloads/Policies/HEALTH-SECTOR-STRETEGY.pdf.

Government of Pakistan (2020): Guidelines: Care of Old Patients in the Wake of COVID-19. 35-01. Available online at http://covid.gov.pk/guideline, checked on 8/19/2020.

Hale, Thomas; Angrist, Noam; Cameron-Blake, Emily; Hallas, Laura; Kira, Beatriz; Majumdar, Saptarshi et al. (2020): Oxford COVID-19 Government Response Tracker. Blavatnik School of Government. Available online at https://covidtracker.bsg.ox.ac.uk/.

Hall, Amanda K.; Cole-Lewis, Heather; Bernhardt, Jay M. (2015): Mobile text messaging for health: a systematic review of reviews. In *Annual review of public health* 36, pp. 393–415. DOI: 10.1146/annurev-publhealth-031914-122855.

Head, Katharine J.; Noar, Seth M.; Iannarino, Nicholas T.; Grant Harrington, Nancy (2013): Efficacy of text messaging-based interventions for health promotion: a meta-analysis. In *Social science & medicine (1982)* 97, pp. 41–48. DOI: 10.1016/j.socscimed.2013.08.003.

Khan, Sheraz; Khan, Zohaib; Landmann, Andreas; Noon, Jawad M.; Rogge, Lisa; Vollmer, Sebastian (2020): Information Constraints and Preventive Behavior amongst COVID-19 Risk Groups in Khyber Pakhtunkhwa, Pakistan. In *AEA RCT Registry*. DOI: 10.1257/rct.6307-1.1.

Kremer, Michael; Rao, Gautam; Schilbach, Frank (2019): Behavioral development economics. In : Handbook of Behavioral Economics - Foundations and Applications 2, vol. 2: Elsevier (Handbook of Behavioral Economics: Applications and Foundations 1), pp. 345–458.

Madajewicz, Malgosia; Pfaff, Alexander; van Geen, Alexander; Graziano, Joseph; Hussein, Iftikhar; Momotaj, Hasina et al. (2007): Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh. In *Journal of Development Economics* 84 (2), pp. 731–754. DOI: 10.1016/j.jdeveco.2006.12.002.

Milkman, Katherine L.; Patel, Mitesh S.; Gandhi, Linnea; Graci, Heather N.; Gromet, Dena M.; Ho, Hung et al. (2021): A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment. In *Proceedings of the National Academy of Sciences of the United States of America* 118 (20). DOI: 10.1073/pnas.2101165118.

Ministry of Health (2020): National Action Plan for Corona Virus Disease (COVID-19) Pakistan. Government of Pakistan. Available online at https://www.nih.org.pk/wp-content/uploads/2020/03/COVID-19-NAP-V2-13-March-2020.pdf, checked on 8/31/2020.

Pop-Eleches, Cristian; Thirumurthy, Harsha; Habyarimana, James P.; Zivin, Joshua G.; Goldstein, Markus P.; Walque, Damien de et al. (2011): Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. In *AIDS* 25 (6), pp. 825–834. DOI: 10.1097/QAD.0b013e32834380c1.

Siddique, Abu; Rahman, Tabassum; Pakrashi, Debayan; Islam, Asad; Ahmed, Firoz (2020): Raising COVID-19 Awareness in Rural Communities: A Randomized Experiment in Bangladesh and India. Munich. Munich Papers in Political Economy, Working Paper No. 9/2020.

World Bank (2021): World Development Indicators | DataBank. Available online at https://databank.worldbank.org/reports.aspx?source=world-development-indicators, updated on 9/21/2021, checked on 9/21/2021.

7. Appendices

7.1. Tables and figures

Table A 1 Data collection and intervention timeline

		2020																															
		April				May				June				July				August				Sept.				October				Nov.			
Week	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Survey design and piloting																																	
Pre-intervention survey (19.4.-11.5.; 16.5.-22.5.; 3.5.-8.7.))																																	
Intervention piloting																																	
Intervention sample selection, randomization & enrollment																																	
Intervention (20.8.-20.9.)																																	
Post-intervention survey (3.9.-25.10.)																																	
Helpline outcome measurement																																	

Table A 2 Wording of messages in English and Urdu

Message	English	Urdu
Introduction	Dear [name of main cardholder], Your health is important to us. We will be sending you important information by SMS regarding protection from corona virus. For more information contact Sehat Card helpline [insert number].	Mohtram [name], Apki sehat hamaray liye ahem hai. Hum apko SMS ke zariye Coronavirus se bachao ke mutaliq ahem malomat faraham kren ge. Mazeed maloomat ke liye Sehat Insaf Card helpline [insert numner] pe rabta kren.
Risk group information	Dear [name of main cardholder], Coronavirus disease is more dangerous and complicated in people above 60 years of age. Also, in people suffering from chronic conditions like diabetes, hypertension, cancer, heart or lung diseases. Practicing prevention can help protect your family.	Mohtram [name], Coronavirus 60 saal se zaid umar k logon mai ziada khatarnak aur paicheeda sabit hota hai. Aur un logon mai bhi jo kisi aur daimi bemari jese k sugar, blood pressure, cancer, dil ya phaipharon k marz mai mubtala hain. Ehtyat krne se ap k ghar walo ki hifazat mumkin ha.
Social distancing information	Dear [name of main cardholder], Staying at home and keeping a distance from others reduces the risk of contracting coronavirus, also in those people who are at higher risk of complicated disease. Provide this information to your household members as well. For more information contact Sehat Card helpline [insert number].	Mohtram [name], Ghar pe rehnay aur doosron se fasla rakhnae se coronavirus lagne k imkanat kum ho jatay hain, un logon mai bhi jin mai paicheeda bemari ka khadsha ziada hai. Apne ghar walo ko bhi ye maloomat dain. Mazeed maloomat ke liye Sehat Card helpline [insert number] pe rabta kren.
Wearing mask information	Dear [name of main cardholder], Keeping a 2-meter distance from others and wearing a mask outside home reduces the risk of contracting coronavirus, also in those people who are at higher risk of complicated disease. Tell this to your household members as well. For more information contact our helpline [insert number].	Mohtram [name], Doosron se 2 meter ka fasla rakhne aur ghar se bahir mask pahen'nay se coronavirus lagne k imkanat kum ho jatay hain, un logon mai bhi jin mai paicheeda bemari ka khadsha ziada hai. Apne ghar walo ko bhi ye btayen. Mazeed maloomat ke liye hamari helpline [insert number] pe rabta kren.
Handwashing information	Dear [name of main cardholder], Regular hand washing for at least 20 seconds reduces the risk of contracting coronavirus disease, also in those people who are at higher risk of complicated disease. Tell this to your household members as well. For more information contact Sehat Card helpline [insert number].	Mohtram [name], Sabun se kum az kum 20 seconds k liye baqaidgi se hath dhonay se coronavirus lagne k imkanat kum hojatay hain, un logon mai bhi jin mai paicheeda bemari ka khadsha ziada hai. Apne ghar walo ko bhi ye btayen. Mazeed maloomat k liye Sehat Card helpline [insert numner] pe rabta kren.
Telemedicine information	Dear [name of main cardholder], In case of a health need, it is possible to have a free telephonic medical consultation before visiting a doctor. This can further protect you from contracting coronavirus. For free medical consultation, call PHA helpline [insert number]. Provide this information to your household members as well.	Mohtram [name], Tabiat kharab honay ki surat mai doctor k pass janay se pehle phone pe muft tibi mashwara liya ja skta hai. Is tarah ap coronavirus se mazeed bach sktay hain. Muft tibi mashwaray k liye ap PHA telemedicine helpline [insert numner] pe rabta kren. Apne ghar walo ko bhi ye maloomat dain.

Table A 3 Differences in means of cardholder characteristics across datasets

	Entitlement	Enrollment	Unique phone
Wealth (PMT score)	18.8819	16.3941	16.6146
	(7.5178)	(6.1754)	(6.1575)
Age	58.6863	59.1794	59.3395
	(13.5752)	(12.9920)	(12.5503)
Female	0.2016	0.1880	0.1676
	(0.4012)	(0.3907)	(0.3735)
Married	0.9448	0.9533	0.9572
	(0.2283)	(0.2109)	(0.2025)
Region			
- North		0.3670	0.3892
		(0.4820)	(0.4876)
- Central		0.2860	0.2160
		(0.4519)	(0.4115)
- Hindko		0.1268	0.1588
		(0.3328)	(0.3655)
- South		0.2182	0.2360
		(0.4130)	(0.4247)
Claim history			
- Any		0.0519	0.0501
		(0.2219)	(0.2181)
- Covid risk		0.0481	0.0465
		(0.2140)	(0.2106)
N	2,371,685	1,480,841	585,657

Standard deviations below mean; a higher PMT score indicates more wealth; all differences between entitlement and enrollment as well as enrollment and households with a unique phone number are statistically significantly different from zero based on a ttest.

Figure A 1 Steps towards final estimation sample from previously interviewed and additionally sampled households

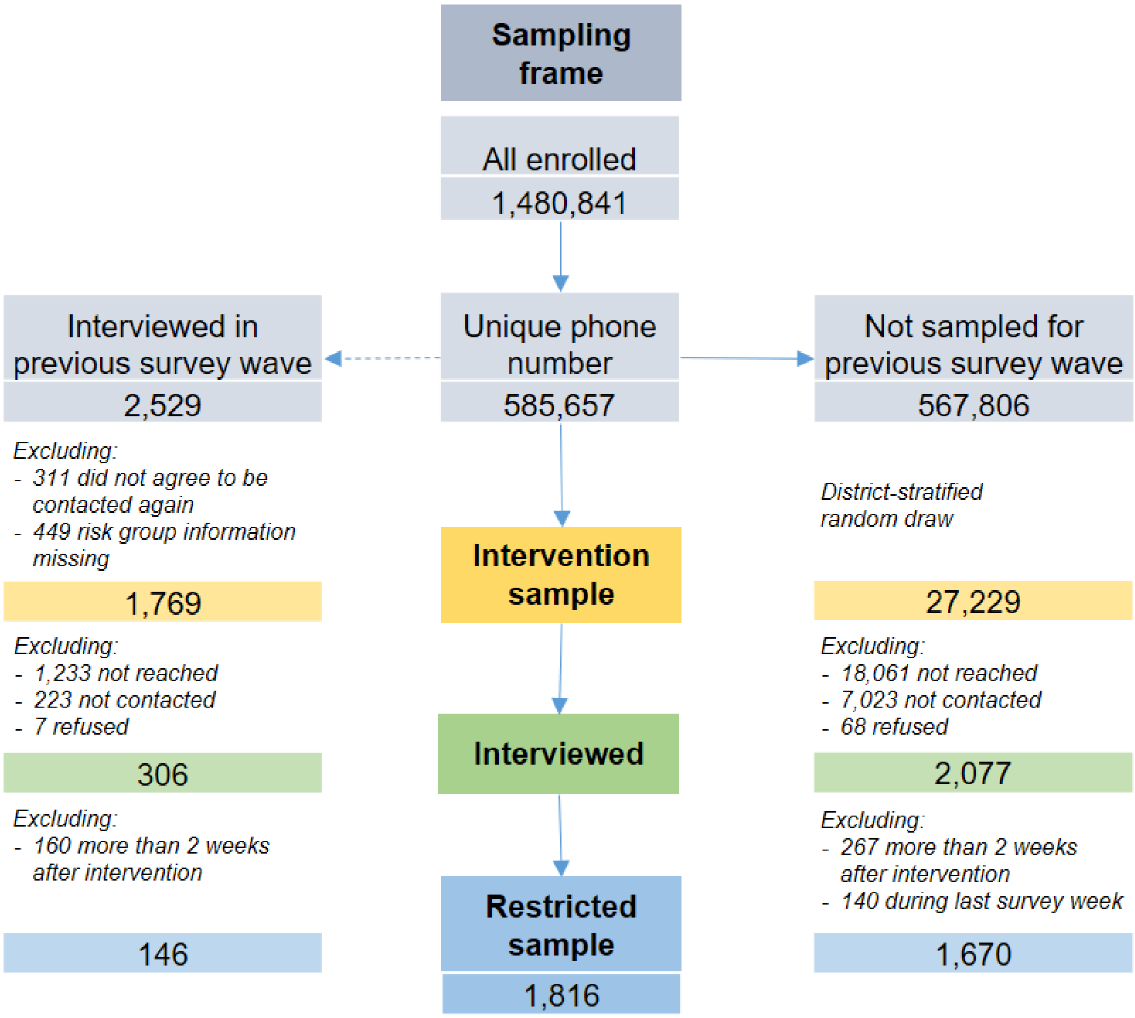


Table A 4 Complete sample characteristics

	Mean	SD	Min	Max	N
Intervention sample					
Cardholder age	49.59	12.57	4	111	29,182
Cardholder female	0.17	0.38	0	1	29,181
Wealth (PMT score)	16.75	6.14	0	27	29,182
Region					
- North	0.29	0.45	0	1	29,183
- Central	0.22	0.41	0	1	29,183
- Hindko	0.09	0.29	0	1	29,183
- South	0.14	0.35	0	1	29,183
Any claim	0.15	0.36	0	1	29,183
Interviewed					
Respondent age	47.57	14.87	18	86	2,408
Any member >60	0.60	0.49	0	1	2,243
Number members >60	1.58	2.15	1	60	1,065
Wealth (PMT score)	17.33	5.97	0	27	2,395
Female	0.06	0.24	0	1	2,413
Respondent literate	0.44	0.50	0	1	2,394
Household literacy	0.78	0.41	0	1	2,105
Respondent					
- Cardholder	0.77	0.42	0	1	2,414
- Household head	0.01	0.12	0	1	2,414
- Spouse	0.03	0.16	0	1	2,414
- Child	0.15	0.36	0	1	2,414
- Other family	0.04	0.19	0	1	2,414
Education					
- Up to primary	0.46	0.50	0	1	2,385
- Secondary	0.16	0.36	0	1	2,385
- Tertiary	0.38	0.49	0	1	2,385
Occupation					
- Civil servant	0.10	0.30	0	1	2,106
- Private employee	0.13	0.33	0	1	2,106
- Self-employed	0.23	0.42	0	1	2,106
- Daily wage laborer	0.33	0.47	0	1	2,106
- Unemployed	0.11	0.31	0	1	2,106
- Other	0.11	0.31	0	1	2,106

Intervention sample refers to all households who were included in the treatment randomization; Interviewed refers to all respondents of the post-intervention survey; a higher PMT score indicates more wealth.

Table A 5 Intervention sample balance based on administrative data characteristics

	Treatment group mean	Control group mean	p-value
Cardholder age	49.49	49.64	0.32
	(12.54)	(12.59)	
Cardholder gender	0.18	0.17	0.23
	(0.38)	(0.38)	
PMT score	16.73	16.76	0.68
	(6.15)	(6.14)	
Claim history	0.15	0.15	0.49
	(0.36)	(0.36)	
N	19,400	9,783	

Intervention sample refers to all households who were included in the treatment randomization; p-values for the test of difference between treatment and control group mean are based on t-tests. Standard deviations in parentheses * *p < 0.1,* ** *p < 0.05,* *** *p < 0.01.*

Table A 6 Interviewed sample balance

	Treatment group mean	Control group mean	p-value
Cardholder age	48.43	49.19	0.12
	(10.94)	(11.49)	
Cardholder female	0.13	0.14	0.62
	(0.34)	(0.35)	
PMT score	17.40	17.29	0.66
	(5.88)	(6.02)	
Claim history	0.18	0.16	0.12
	(0.39)	(0.36)	
Respondent age	46.78	47.96*	0.07
	(14.79)	(14.89)	
Respondent female	0.06	0.06	0.52
	(0.23)	(0.25)	
Respondent education			0.551
- Primary or less	0.49	0.45	
	(0.50)	(0.50)	
- Secondary	0.14	0.16	
	(0.35)	(0.37)	
- Tertiary	0.37	0.38	
	(0.48)	(0.49)	
N	1,622	792	

Interviewed refers to all respondents of the post-intervention survey; p-values for the test of difference between treatment and control group mean are based on t-tests for all binary characteristics and on the Pearson chi-squared test for the categorical education variable. Standard deviations in parentheses * *p < 0.1,* ** *p < 0.05,* *** *p < 0.01.*

Attrition analysis

We test for differential attrition from the whole intervention sample to respondents of the post-intervention survey in the following three ways. We display all tests for the complete intervention sample using characteristics from the administrative records and display additionally more detailed characteristics for the subset that was interviewed in a previous survey wave. As displayed in Figure A 1, the majority (77.5%) of households are lost to follow up because they were contacted at least 3 times according to the calling schedule, but not reached or interviewed. An additional 22.5% of sampled households was not contacted before the stopping rule applied.

First, we test whether there is differential attrition between treatment and control group by regressing a binary attrition indicator on the binary treatment indicator:

$$Attrit_i = \alpha + \beta Treat_i + \varepsilon_i \qquad \textit{(A1)}$$

Secondly, we test whether there is differential attrition based on observable characteristics y_i. For the whole sample, this is restricted to the administrative health insurance data: age of main cardholder, gender of main cardholder, poverty score, region of residence, and any previous claim experience. For the interview sample, this can be extended to survey-based respondent characteristics: age, gender, education and occupation.

$$y_i = \alpha + \beta Attrit_i + \varepsilon_i \qquad \textit{(A2)}$$

Finally, we examine whether these characteristics are significantly different among attrited treatment and control households by restricting the sample to attrited households only:

$$(y_i|Attrit = 1) = \alpha + \beta Treat_i + \varepsilon_i \qquad \textit{(A3)}$$

Table A 7 Attrition 1: Test for differential attrition between treatment and control group

	Intervention sample	Interviewed in previous survey wave
	Attrited	Attrited
Treatment group	-0.00171	0.00455
	(0.00341)	(0.0193)
N	29150	1769

Regression of a binary attrition indicator (sampled, but not (re-)interviewed) on a binary treatment indicator following equation A1. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 8 Attrittion 2: Test for differential attrition based on observable characteristics in intervention sample

	Cardholder's age	Cardholder female	Wealth (PMT score)	At least one health insurance claim
Attrited	0.421	0.0363***	-0.603***	-0.0144*
	(0.260)	(0.00806)	(0.131)	(0.00765)
Observations	28965	29148	29117	29150

Separate regressions of each characteristic on the binary attrition indicator (sampled, but not interviewed) using equation A2. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 9 Attrition 2: Test for differential attrition based on observable characteristics in households interviewed in a previous survey wave

	Cardholder age	Cardholder female	Wealth (PMT score)	At least one health insurance claim	Respondent age	Respondent female
Attrited	0.373	0.00103	0.121	-0.00984	-0.519	-0.00343
	(0.701)	(0.0229)	(0.363)	(0.0229)	(0.924)	(0.0146)
N	1765	1769	1768	1769	1743	1739

Separate regressions of each characteristic on the binary attrition indicator (sampled, but not re-interviewed) using equation A2 in the sub-sample that was interviewed in a previous survey wave. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 10 Attrition 3: Test for difference along observable characteristics between treatment and control group among the attrited in the complete sample

	Cardholder's age	Cardholder female	Wealth (PMT score)	At least one health insurance claim
Treatment group	-0.0124	-0.00651	0.0440	0.00601
	(0.159)	(0.00492)	(0.0794)	(0.00463)
N	26579	26755	26727	26757

Separate regressions of each characteristic on the binary treatment indicator (sampled, but not interviewed) using equation A3. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 11 Attrition 3: Test for difference along observable characteristics between treatment and control group among the attrited in the previously interviewed sample

	Cardholder age	Cardholder female	Wealth (PMT score)	At least one health insurance claim	Respondent age	Respondent female
Treatment group	1.185^{*}	-0.0176	-0.0317	0.0118	0.541	-0.0281^{**}
	(0.628)	(0.0205)	(0.325)	(0.0206)	(0.841)	(0.0131)
N	1,477	1,481	1,480	1,481	1430	1427

*Separate regressions of each characteristic on the binary treatment indicator (sampled, but not re-interviewed) using equation A3 in the sub-sample that was interviewed in a previous survey wave. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Sample restriction

Figure A 2 Sample balance in terms of household characteristics from the administrative data across treatment and control group in each data collection week (only for those interviewed up to one week after the intended interview date)

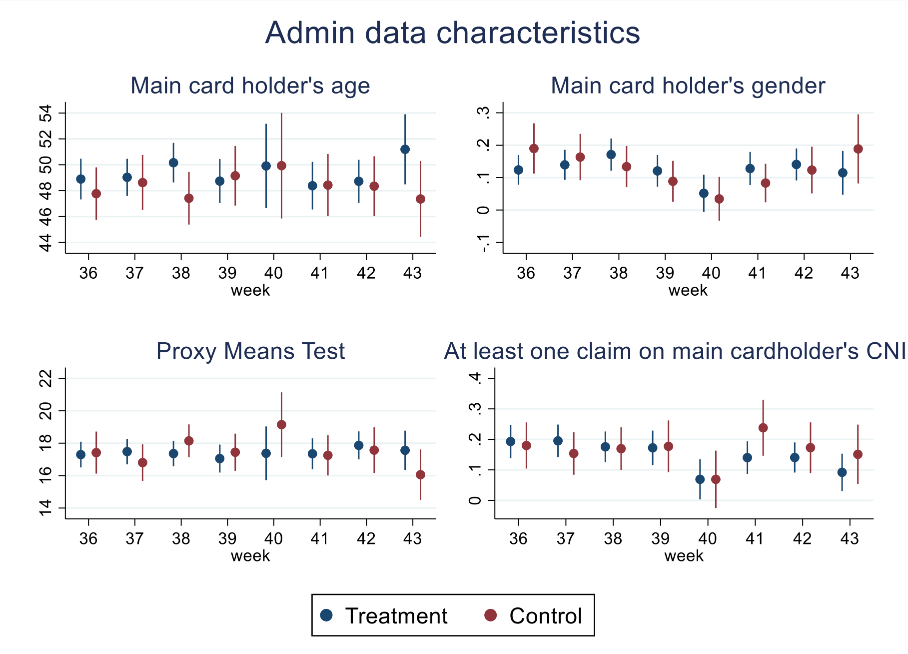


Figure A 3 Sample balance in terms of respondent characteristics from the administrative data across treatment and control group in each data collection week (only for those interviewed up to one week after the intended interview date)

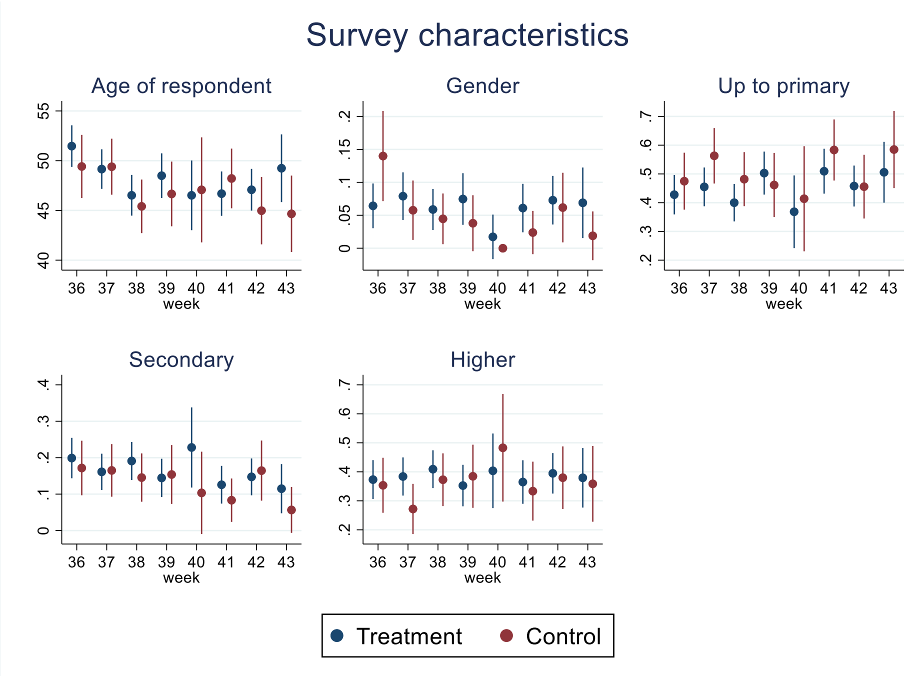


Table A 12 Treatment effects on preventive practices (hypothesis 1): specification with and without control variables

	Preventive practices index		Handwashing		Wearing masks		Telemedicine if sick	
Treated	0.116**	0.101*	0.0582**	0.0525**	0.0295	0.0252	0.0457*	0.0517**
	(0.0521)	(0.0514)	(0.0266)	(0.0263)	(0.0258)	(0.0255)	(0.0251)	(0.0250)
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
N	1605	1605	1605	1605	1605	1605	329	329
Control group mean	1.524	1.524	0.470	0.470	0.604	0.604	0.0182	0.0182

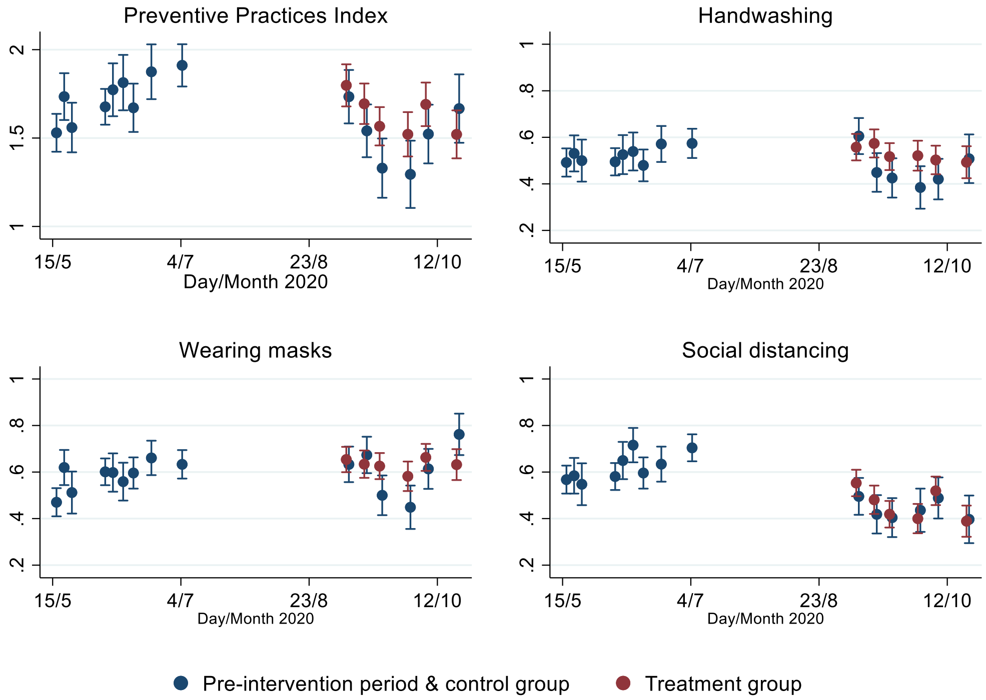
*Estimation results of equation 1 with the respective preventive practice on the binary treatment indicator. The sample is restricted to being interviewed within one week of the intended interview date and excluding the last week of data collection. The preventive practices index is a count-index ranging from 0 to 4 counting whether the respondent mentioned to practice handwashing, social distancing or mask wearing on a regular basis and/ or using telemedicine in case of a health need. The sample for telemedicine usage is restricted to households who has a health need during the previous month. Control variables: respondent age, gender, education, household wealth. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table A 13 Treatment effects on preventive practices among risk and non-risk group (hypothesis 2)

	Preventive practices index		Handwashing		Wearing masks		Telemedicine if sick	
Treated	0.1003	0.0800	-0.0017	-0.0045	0.0080	-0.0010	0.0417	0.0439
	(0.0919)	(0.0909)	(0.0468)	(0.0466)	(0.0455)	(0.0451)	(0.0575)	(0.0575)
Risk group	0.0655	-0.0014	-0.0562	-0.0689	0.0218	-0.0019	0.0225	0.0153
	(0.0918)	(0.0917)	(0.0468)	(0.0470)	(0.0454)	(0.0455)	(0.0534)	(0.0534)
Treatment x Risk group	0.0489	0.0437	0.0947*	0.0862	0.0457	0.0492	0.0069	0.0098
	(0.1120)	(0.1107)	(0.0571)	(0.0568)	(0.0554)	(0.0550)	(0.0641)	(0.0643)
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
N	1576	1568	1576	1568	1576	1568	326	324

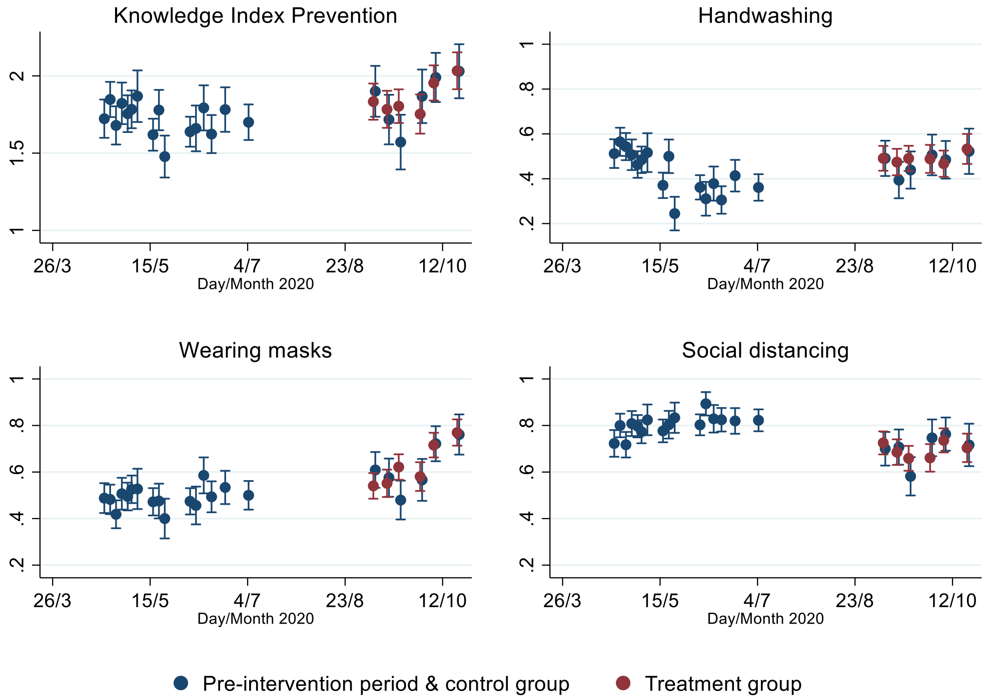
*Estimation results of equation 2 with the respective preventive practice on the binary treatment indicator. The preventive practices index is a count-index ranging from 0 to 4 counting whether the respondent mentioned to practice handwashing, social distancing or mask wearing on a regular basis and/ or using telemedicine in case of a health need. Risk group means having at least one household member above the age of 60 and/ or with a precondition that increases the risk for a severe COVID-19 infection. The sample for telemedicine usage is restricted to households who has a health need during the previous month. Control variables: respondent age, gender, education, household wealth. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Figure A 4 Preventive practice outcomes over time and across treatment and control group



Plotted mean estimates over time intervals in pre-intervention period (blue), and by treatment (red) and control group (blue) in the intervention period; with 90% confidence intervals; each displayed estimate is based on 218 observations on average (min: 108, max: 358) that were collected over the course of 1-6 days (in five cases 7, 9, 10 and 17 days).

Figure A 5 Prevention knowledge outcomes over time and across treatment and control group



Plotted mean estimates over time intervals in pre-intervention period (blue), and by treatment (red) and control group (blue) in the intervention period; with 90% confidence intervals; each displayed estimate is based on 218 observations on average (min: 108, max: 358) that were collected over the course of 1-6 days (in five cases 7, 9, 10 and 17 days).

Table A 14 Treatment effects on preventive practices in risk and non-risk group separately

	Preventive practices index		Handwashing		Wearing masks		Telemedicine if sick	
Group	Risk	Non-risk	Risk	Non-risk	Risk	Non-risk	Risk	Non-risk
Treated	0.1447**	0.0887	0.0900***	-0.0018	0.0542*	0.0011	0.0480	0.0417
	(0.0642)	(0.0922)	(0.0327)	(0.0472)	(0.0315)	(0.0465)	(0.0300)	(0.0454)
Control variables	No	No	No	No	No	No	No	No
N	1054	514	1054	514	1054	514	256	68
Control group mean	1.5347	1.4790	0.4538	0.5090	0.6012	0.5868	0.0230	0.0000

*Estimation results of equation 1 with the respective preventive practice on the binary treatment indicator, separately in the sub-sample of households with a risk group member and no household member in the risk group. The preventive practices index is a count-index ranging from 0 to 4 counting whether the respondent mentioned to practice handwashing, social distancing or mask wearing on a regular basis and/ or using telemedicine in case of a health need. Standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A 15 Treatment effects on risk knowledge (hypothesis 4): specification with and without control variables

	Knowledge index risk		Old age		Precondition	
Treated	0.0460	0.0324	0.0443*	0.0372	0.0017	-0.0048
	(0.0368)	(0.0358)	(0.0256)	(0.0252)	(0.0238)	(0.0235)
Control variables	No	Yes	No	Yes	No	Yes
N	1656	1656	1656	1656	1656	1656
Control group mean	0.8699	0.8699	0.5818	0.5818	0.2881	0.2881

*Estimation results of equation 1 with the respective knowledge outcome on the binary treatment indicator. The sample is restricted to being interviewed within one week of the intended interview date and excluding the last week of data collection. The risk knowledge index is a count-index ranging from 0 to 2 counting whether the respondent mentioned age or a correct precondition as risk factors for a complicated COVID-19 disease course. Control variables: respondent age, gender, education, household wealth. Standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A 16 Treatment effects on prevention knowledge (hypothesis 4): with and without control variables

	Knowledge index prevention		Social distancing		Handwashing		Wearing masks		Telemedicine	
Treated	0.0186	0.0058	-0.0068	-0.0114	0.0193	0.0156	0.0129	0.0089	-0.0062	-0.0070
	(0.0516)	(0.0505)	(0.0239)	(0.0236)	(0.0260)	(0.0259)	(0.0252)	(0.0249)	(0.0106)	(0.0106)
Control variables	Yes	No	Yes	No	No	Yes	No	Yes	Yes	No
N	1695	1695	1695	1695	1695	1695	1695	1695	1791	1791
Control group mean	1.8358	1.8358	0.7026	0.7026	0.4690	0.4690	0.6113	0.6113	0.0500	0.0500

*Estimation results of equation 1 with the respective knowledge outcome on the binary treatment indicator. The sample is restricted to being interviewed within one week of the intended interview date and excluding the last week of data collection. The prevention knowledge index ranges from 0 to 4 counting whether the respondent mentioned handwashing, social distancing or mask wearing on a regular basis and/ or using telemedicine in case of a health need as prevention methods against a COVID-19 infection. Control variables: respondent age, gender, education, household wealth. Standard errors in parentheses * p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A 17 Adjustment for multiple hypothesis testing for hypothesis group 1

	Preventive practices index	Handwashing	Wearing masks	Telemedicine if sick
Treated	0.1214	0.0597	0.0310	0.0461
	(0.0197)**	(0.0242)**	(0.2285)	(0.0648)*
	[0.0969]*	[0.0969]*	[0.3656]	[0.1727]
Observations	1,614	1,614	1,614	331

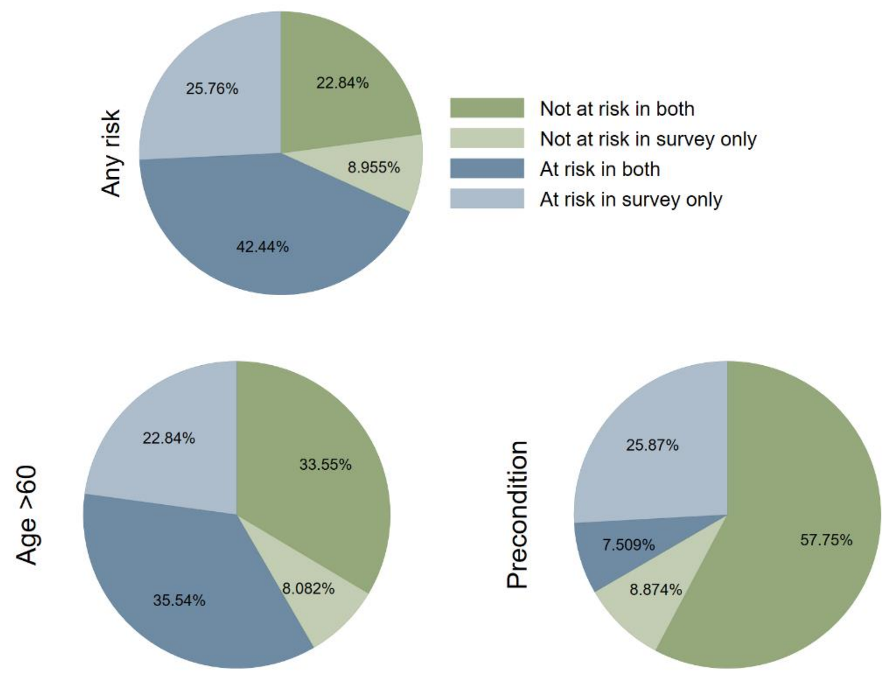
*Estimation results of equation 1 with the respective preventive practice on the binary treatment indicator. The preventive practices index is a count-index ranging from 0 to 4 counting whether the respondent mentioned to practice handwashing, social distancing or mask wearing on a regular basis and/ or using telemedicine in case of a health need. Unadjusted p-values in parentheses, q-values adjusted for multiple hypothesis testing following Benjamini and Hochberg (1995); * p/q < 0.1, ** p/q < 0.05, *** p/q < 0.01.*

Table A 18 Adjustment for multiple hypothesis testing for hypothesis group 2

	Preventive practices index	Handwashing	Wearing masks	Telemedicine if sick
Treated x Risk group	0.0489	0.0947	0.0457	0.0069
	(0.6623)	(0.0974)*	(0.4100)	(0.9147)
	[0.7569]	[0.1948]	[0.5466]	[0.9147]
Observations	1,576	1,576	1,576	326

*Estimation results of interaction coefficient in equation 2 with the respective preventive practice on the binary treatment indicator times the binary risk group indicator. The preventive practices index is a count-index ranging from 0 to 4 counting whether the respondent mentioned to practice handwashing, social distancing or mask wearing on a regular basis and/ or using telemedicine in case of a health need. Unadjusted p-values in parentheses, q-values adjusted for multiple hypothesis testing following Benjamini and Hochberg (1995); * p/q < 0.1, ** p/q < 0.05, *** p/q < 0.01.f*

Figure A 6 Identification of COVID-19 risk factors in the interview and administrative data



Shares of all interviewed households (during pre-intervention and intervention period) who are identified as having at least one at-risk member in the survey and/ or administrative data. "Any risk" captures all households who have at least one member over the age of 60 and/ or with a relevant precondition.

7.2. Supplementary information

7.2.1. Power calculation

To determine the sample size that is needed to detect an effect on the main outcomes, we first derived the expected minimal detectable effect size (MDE) using the pre-intervention interview data. Based on the theory of change, we expected the effect of reported practices to be driven by an increase in knowledge of risk groups and preventive practices. The MDE was therefore determined following formula A1 as the sum of the difference in adoption of each preventive practice *i* that might be associated with closing the gap of knowing the respective practice and at least one of the major risk factors of a complicated COVID-19 infection. We expected that around 1/3 of endline respondents in the treatment group would have received and read the messages and therefore had a chance to improve their knowledge on risk and practice information that is conveyed in the messages. This ratio is reasonable as 34% of respondents reported to have received our SMS announcing the interview. A similar share of pilot message recipients could recall receiving and reading the messages. The needed sample sizes were calculated following formula A2 for the whole sample and separately for the group of households with and without at least one member that is in the risk group *r*. As both knowledge and adoption of the aggregate measure of social distancing were already high and not as related, sample sizes of above 7000 in each sample would be required to detect a 2 percentage points (pp) increase. As this was not logistically not feasible, we focus the outcome measurement on handwashing, wearing mask and using telemedicine. To detect a difference of 6.4-7.7pp for both handwashing and wearing masks, a sample size of slightly above 1000 is needed in both the risk and the non-risk group. The effect sizes in the risk group can be expected to be slightly larger if the respondent becomes aware of the specific risk that the household is exposed to rather than any risk factor in general. Hence, to detect an effect in the total population will require a sample size of around 1000 and a sample size of 2500, where 2/3 are allocated to the treatment group should ensure being able to detect a difference of 6-7pp between risk and non-risk group and personalized versus generic messages.

$$(A1)\; mde_{ir} = \left((1 - know.prac_{ir}) * \frac{1}{3}\right) * (prac_{ir} * know.prac_{ir} - prac_{ir} * not.know.prac_{ir}) + \left((1 - know.risk_{r}) * \frac{1}{3}\right) * (prac_{ir} * know.risk_{r} - prac_{ir} * not.know.prac_{ir})$$

$$(A2)\;\; n_{ir} = \frac{prac_{ir}}{\left(\frac{2}{3} * mde_{ir}^{2}\right)} * \frac{-prac_{ir} + 1}{\frac{1}{3}} * (-1.28 - 0.84)^{2}$$

The sample size for the outcome of telemedicine usage could not be calculated in the same way ex-ante as telemedicine usage in case of a health need was only included in the survey at a later stage, so that there is not a sufficient number of observations. Out of the 200 respondents, around 18% reported a health need not related to COVID-19 in the family during the previous month, but only one reported to have called a doctor, and none reported the use of a telemedicine helpline. Around half reported self-treatment and the other half visited a doctor or hospital. One reason for this might be a low awareness of the recommended use of telemedicine before visiting a doctor – when asked about what is currently recommended to do in case someone has a health need that is not related to COVID-19, only two mentioned telemedicine and three mentioned calling a doctor. Hence, any change in knowledge and practice with respect to telemedicine should be detectable in the sample that can detect changes in the other practices.

In order to account for non-response in the post-intervention survey, it was necessary to draw a substantially larger intervention sample. Based on the experiences from a previous follow-up survey wave, it was expected that 30% of those who were previously interviewed would be reached again for follow-up and 10% of the previously uncontacted households will be reached. To reach an interview sample of 2,500, it will therefore be necessary to contact around 21,500 numbers, the intervention will have been sent to two thirds of these (around 14,300).

7.2.2. Data collection and processing

For this study, we used a research infrastructure that was established with the helpline firm ICU healthcare and the Social Health Protection Initiative before the onset of the COVID-19 pandemic. This infrastructure builds on a data sharing portal as well as a platform for web-based questionnaires that ensure protection of the personal data.

The pre- intervention survey was collected over the phone in three waves between April and July 2020 as a rapid response to document attitudes, knowledge and actions of the sample population over the course of the COVID-19 outbreak. As depicted in Figure 1 in the main text, the first survey wave started on April 19, a few weeks into the nationwide lockdown and a few days before the onset of Ramadan. Hence, most of the first wave took place during Ramadan and a follow-up wave with the same respondents was conducted within the last week of Ramadan, but after the strictest lockdown regulations were lifted. For the third wave after Ramadan, a new sample was selected following the same sampling procedure as before. The main areas covered by the survey were: Socio-economic characteristics of the respondent and the household; attitudes such as trust in different groups and organizations regarding their message on corona and the government's reaction; knowledge of coronavirus (symptoms, transmission, prevention, treatment, risk groups); prevention practices as well as actions in the case of a suspected corona case in the family / community; disease perceptions (severity and likelihood) and capacity for self-isolation. Where possible, the survey instruments were taken from previously validated and standardized questionnaires. The questionnaire was translated into Pashto, Urdu and Hindko language by the local research partners and administered in the language that the interviewee was fluent in.

In order to follow the social distancing recommendations, the trained interviewers conducted the interviews over the phone from their homes. They were using SIM cards with an official and uniform number from ICU healthcare. Informed consent was taken verbally, and the interview data was entered into a web-based form, which is part of the secure data framework. The research team is only able to download anonymized datasets, but can still link all datasets using the anonymized identifiers.

We took several measures to tackle non-response. First, we are aware that many telephone numbers are not valid anymore as the owners have changed numbers since enrollment (which might date up to five years back). For the valid numbers, in order to build trust, text messages announcing the call were sent at least one day before the first call. If the respondent was not reached at the first call, s/he was re-called 2 times according to a protocol (call 2: one hour after unsuccessful call 1 and call 3: at a different time on the next day). Interviewers would also take appointments. We aimed to interview the main cardholder, which was successful in around 77% of the interviews (see appendix Table A 4). If it was not possible to conduct the interview with him/ her, e.g. due high age (above 65 years), hearing or language difficulties or not being willing to be interviewed, another household member (ideally the household head or the main cardholder's spouse) was interviewed. If the listed main cardholder was not a member of the household (anymore), this household was excluded.

Post-intervention interviews were conducted following the same procedure. The survey was designed to not last longer than 10-15 minutes and only includes a subset of questions from the previous more detailed survey. The questions that remained cover all outcome measurements (as described in section 3.4) as well as an additional section that elicits mobile phone usage as well as whether and how our messages were received. In order to keep the time between receiving the intervention and the interview constant across respondents, the start of the intervention was staggered according to the amount of interviews that the interviewer team was expected be able to conduct in one day. As the time between receiving the last message and the interview should be one week, the first interviews began on September 2nd. The termination rule for data collection was that all sampled households have been either interviewed or contacted at least three times according to the calling schedule. As this was not reached after one month as intended, the intervention and data collection was continued for another four weeks.

The enumerators were 14 students and graduates of Khyber Medical University. Beyond thorough survey piloting and enumerator training, where possible, we programmed error-checks into the questionnaire, and closely monitored the consistency of the data during and after data collection.

In addition to the survey data, we use calls to the Sehat Insaf Card helpline as well as the Public Health Association's (PHA) telemedicine helpline, which are both administered by ICU healthcare. Call-logs as well as a short questionnaire on the demographics and reason for calling were incorporated into the data portal and matched to the study sample using anonymized identifiers based either national identification number or telephone numbers.

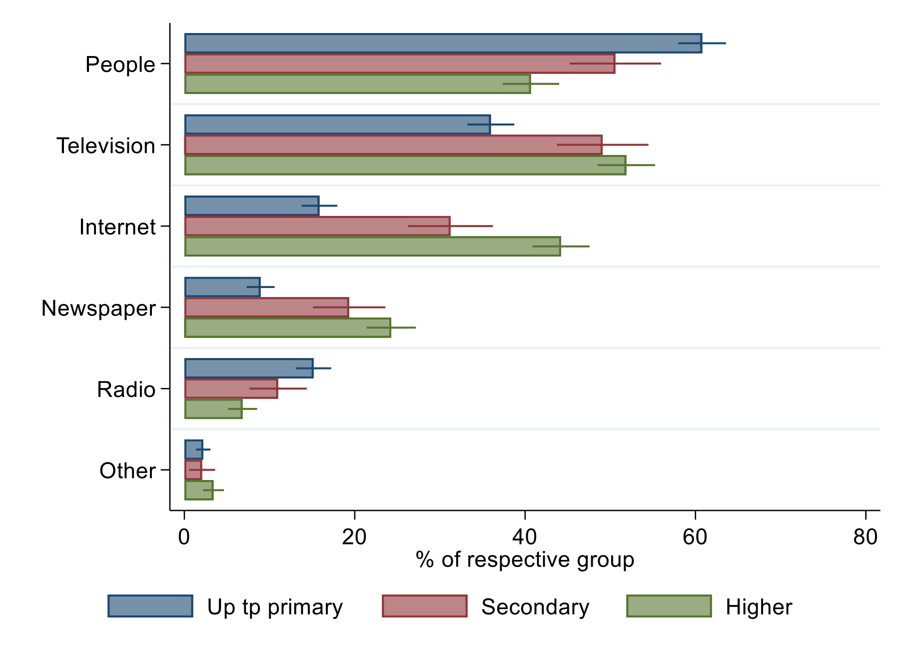
7.2.3. Insights from pre-intervention survey and intervention piloting

We relied on two kinds of pilot data: quantitative interview data and qualitative interviews after a pilot intervention.

The sample characteristics of the pre-intervention survey are very similar to the post-intervention survey. On average, respondents are around 50 years old, and more than half have a household member that is above 60 years old. Around half of the sample have not more than primary education, 14% have secondary and 32% tertiary. Most are either daily wage laborers or self-employed, 15% are unemployed, 11% are private employees and 9% civil servants. While only 51% of respondents say that they are literate, 75% have at least one family member who can read in Urdu, so that the majority of the sample would be able to read an SMS. Only 42% on the other hand report any access to the internet, so the majority of the sample could not be reached by a web-based information campaign.

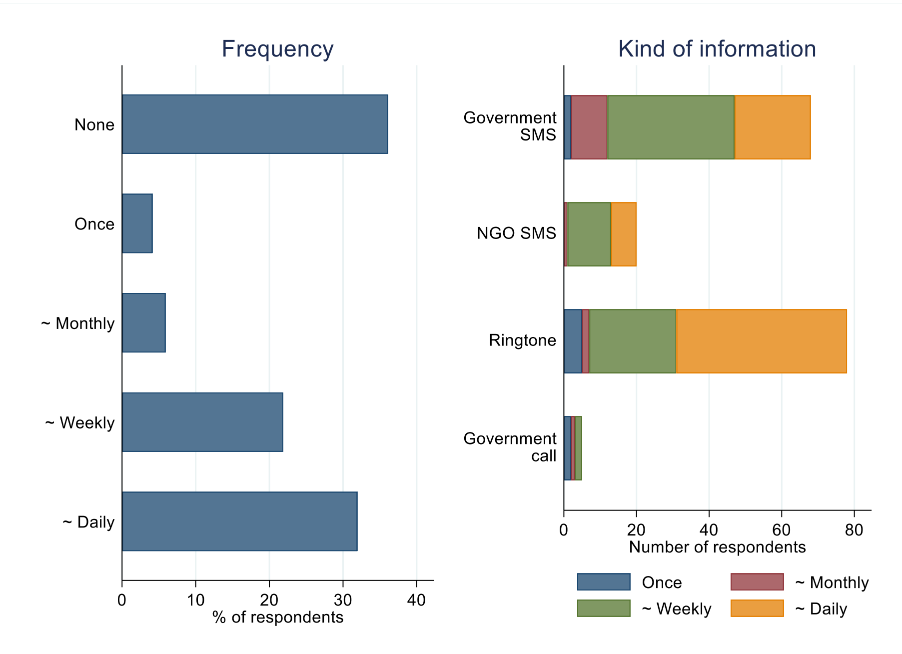
Literacy rate and internet access also reflect in the way information on COVID-19 is sought. As depicted in Figure A 7, the majority of respondents relied on others and television for information, while word of mouth is a much more common source of information among those with at most primary education and internet and newspaper only play a large role for those with above primary education. Though rarely mentioned freely in the open question on the information source, around 75% of the 250 interviewees in the last weeks of data collection confirmed that they had received some information on COVID-19 through their mobile phone. Out of those, around half reports to have received information on a daily basis, which is mainly driven by listening to the caller tune with every outgoing call, around the same number of respondents reports to have received government SMS, but with a lower frequency (see Figure A 8).

Figure A 7 COVID-19 information source during the previous week by education level of the respondent



Share of each education group reporting to have used the respective information source during the pre-intervention survey period; including 95% confidence interval around each group mean.

Figure A 8 Frequency and type of mobile-phone based COVID-19 information



For subset of 189 respondents during the last weeks of data collection that reported to have received COVID-19 information through their mobile phone.

Risk groups for a complicated COVID-19 disease course can be identified from the survey data based on age and preconditions following the definitions in chapter 3.4. We see that almost 50% of households have at least one household member that is over 60 years old and

38% reported to have at least one household member with the previously mentioned medical conditions that indicate increased risk based on a precondition. As depicted in Figure A 6, half of this age-based risk group can also be identified in the administrative data using the age of the main cardholder in the enrollment data and 28% of the households that reported a relevant precondition could be identified from the claims data. Within the risk group that can be identified in the survey data, around 2/3 are characterized by age risk. Precondition-based risk is dominated by household members with hypertension and diabetes followed by cardiovascular diseases more generally, while respiratory diseases and cancer take up a much smaller share. Similarly, a large share of households with a precondition also has a member over the age of 60 and different preconditions coincide within one household (Table A 19).

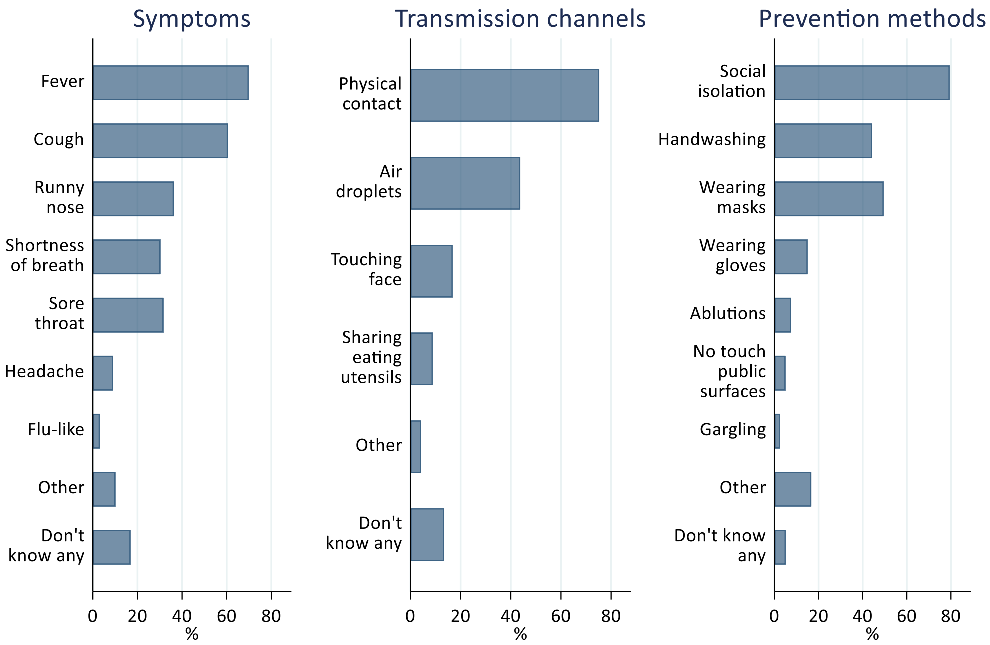
Table A 19 Risk factor distribution in the pre-intervention survey sample

	Any risk	Age >60	Any precon dition	Hypert ension	Diabet es	CVD	Respir atory	Cance r
% of households	69	60	26	14	14	10	3	1
Number of households	1,416	1,239	540	295	278	201	54	17
+ Age >60				203	189	134	40	12
+ Hypertension					86	67	12	7
+ Diabetes						64	13	7
+ CVD							11	5
+ Respiratory								3

Number of households in the pre-intervention survey sample with the respective risk factor and below the pairwise number of coinciding risk factors.

The survey sections on disease knowledge and action reveal a generally high awareness of COVID-19 and its severity, but also significant gaps. Only half of the respondents could name both fever and cough as symptoms of COVID-19, 75% knew that it can be transmitted through physical contact but only 44% that it can also be transmitted via air droplets (Figure A 9). Social distancing is widely known as prevention method, wearing masks by about half and increasing over time, but hygiene measures such as handwashing were named by less than half of the respondents and decreasingly over time. Risk group knowledge seems more limited, around 60% of respondents are aware of old age being a risk factor, but only 20% mention any precondition while over 30% falsely mention children.

Figure A 9 Listed symptoms, transmission or prevention knowledge item



Displays share of respondents who mentioned the respective item in open questions regarding symptoms, transmission channels and prevention methods respectively.

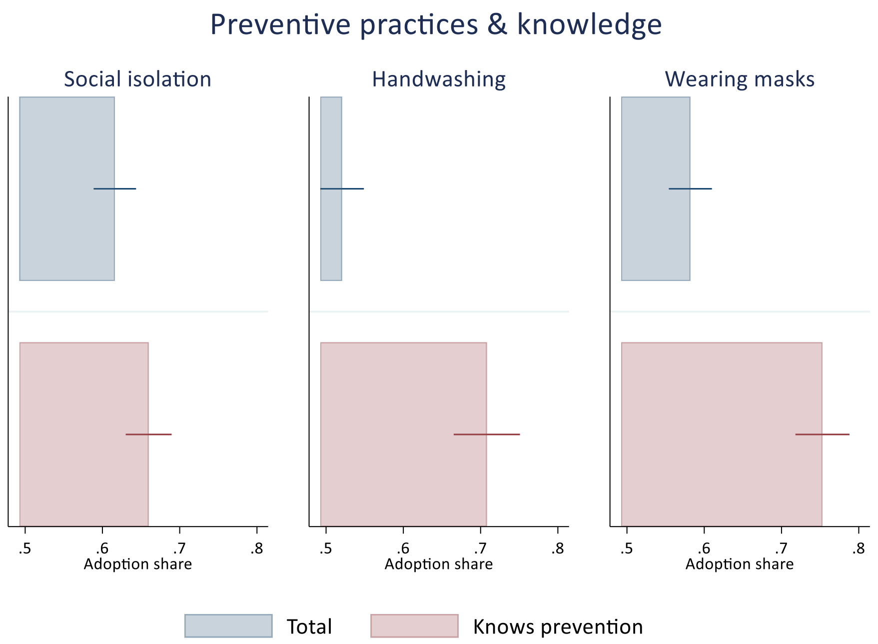
We further find evidence for the channels that are predicted by the theory of change. The overall state of knowledge and practice does not seem to differ strongly across risk and non-risk group (Table A 20). Only knowledge of all main symptoms (fever, cough and shortness of breath), old age being a risk factor and the use of masks are slightly higher in the risk group. One reason for the lack of a difference might be that almost half of risk group households are not aware of at least one of the own risk factors. The data further hints that the adoption of preventive practices is strongly associated with knowledge of the practice, both in the overall sample (Figure A 10) and the risk group in particular when knowledge of the own risk and the preventive practice are combined as it is the case in the intervention (*Figure A 11*). This gives us confidence that an intervention that raises awareness for both risk groups and effective preventive measures could contribute to closing gaps in adoption of preventive practices.

Table A 20 Mean knowledge and practice in full sample and comparison between risk and non-risk group

	Full sample	No risk	Any risk
Knowledge			
- main symptoms	0.1907	0.1650	0.2067**
	(0.3930)	(0.3714)	(0.4051)
- fever& cough	0.5272	0.5162	0.5341
	(0.4994)	(0.5000)	(0.4990)
- airborne	0.4380	0.4161	0.4513
	(0.4963)	(0.4932)	(0.4978)
- smear transmission	0.8004	0.7836	0.8106
	(0.3998)	(0.4120)	(0.3920)
- Social distancing	0.7941	0.8043	0.7878
	(0.4045)	(0.3970)	(0.4090)
- hygiene	0.4420	0.4220	0.4542
	(0.4967)	(0.4942)	(0.4981)
- mask	0.4949	0.4847	0.5012
	(0.5001)	(0.5001)	(0.5002)
- age risk	0.6264	0.6032	0.6401*
	(0.4839)	(0.4896)	(0.4802)
- precondition risk	0.2002	0.1824	0.2107
	(0.4002)	(0.3865)	(0.4080)
Practice			
- going out daily	0.6855	0.6640	0.6981
	(0.4644)	(0.4727)	(0.4593)
- times going out	4.1955	4.3354	4.1174
	(4.2549)	(5.2881)	(3.5501)
- for shopping	0.6234	0.6123	0.6295
	(0.4847)	(0.4876)	(0.4831)
- for work	0.4931	0.5089	0.4845
	(0.5001)	(0.5003)	(0.5000)
- for prayer	0.3829	0.3667	0.3917
	(0.4862)	(0.4823)	(0.4883)
- social distancing	0.6383	0.6747	0.6184
	(0.4808)	(0.4693)	(0.4862)
- hygiene	0.5279	0.5479	0.5169
	(0.4995)	(0.4986)	(0.5002)
- mask	0.6092	0.5685	0.6316*
	(0.4882)	(0.4961)	(0.4828)
Suspect action			
- consult health worker	0.4575	0.4798	0.4456
	(0.4983)	(0.5000)	(0.4972)
- get test	0.8129	0.8085	0.8152
	(0.3901)	(0.3938)	(0.3883)
Telemedicine			
Heard of	0.4153	0.3783	0.4340
	(0.4931)	(0.4859)	(0.4961)
Willing to use	0.8208	0.8068	0.8279
	(0.3837)	(0.3955)	(0.3778)
N	2,529	1,127	1,402

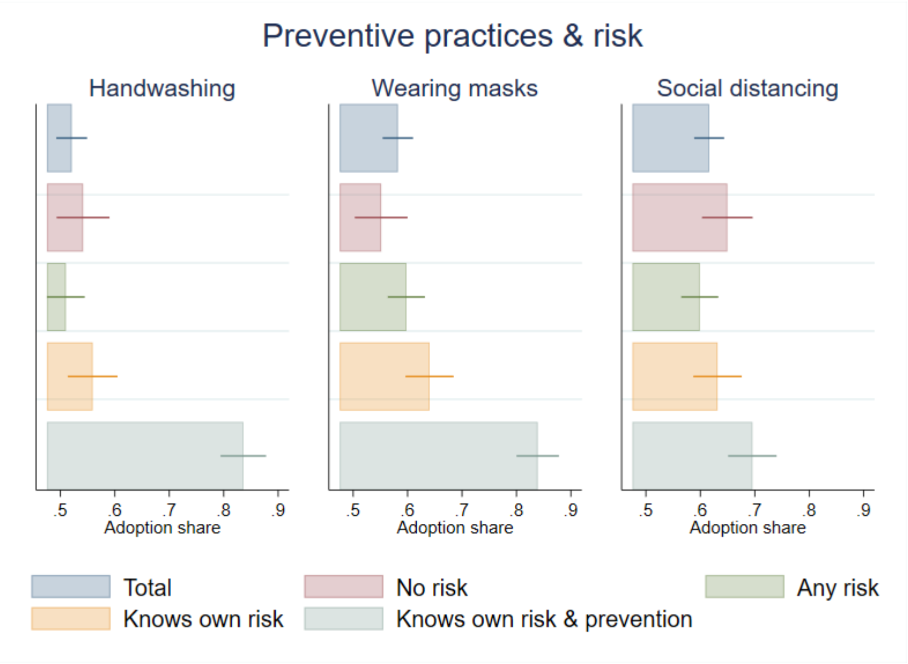
*Standard deviations below mean; stars indicate significant difference between no risk and risk group based on t-test, * 0.1 ** 0.05 *** 0.01*

Figure A 10 Mean preventive practice adoption in full sample compared to those who stated to know the practice



Bars depict mean estimates with 95% confidence intervals.

Figure A 11 Preventive practices in pre-intervention survey period



Bars depict the adoption share of the respective practice in the complete sample (blue), among those who have no at-risk household member (red), at least one at-risk household member (green), those who additionally know about the own household's risk (yellow) and those who additionally mentioned the respective practice in the prevention knowledge question (light green) with 95% confidence intervals.

Intervention piloting

We further pre-tested the intervention among 400 numbers. Shortly after receiving the messages, semi-structured in-depth interviews were conducted with 12 message recipients to understand whether and how they received the messages, and if so how they perceived them and whether they acted upon them. From the message delivery reports, we see that 15% of messages could not be delivered, so that these are likely invalid numbers. Out of the 85% for whom the messages were either delivered to phone or network, we can only find out at the time of the interview whether the number is still active and being used. Out of those who were reached and agreed to be interviewed, half report to have received and read our messages.

Respondents generally appreciated the messages even though they reported that most of the content was not new to them, but a good reminder and confirmation that this information is still valid. Many also reported that they normally have limited access to new information as they live remotely and rely on others sharing information with them. Along those lines, all but one respondent would like to keep receiving such messages with the same or even higher frequency, so that we do not expect our messages to be perceived as a burden. One respondent mentions that he doubts the existence of COVID-19 as there has not been a case in his village and therefore finds it irritating to keep receiving information on this topic.

Respondents received the messages in different ways: out of those who recalled the messages, four had read them themselves and two had either the brother or son read it to them. This shows that despite high illiteracy among respondents, they can rely on other household members to read the message and receive it nevertheless. One case also shows the limits of this strategy as one of the respondents said that his brother would have normally ready it to him, but could not as he was sick. If respondents did not recall the messages, the interviewers read the messages to them to still elicit their opinion on content and wording. After reading it, one further respondent remembered that his brother had told him about the message as it had mentioned his name. Three respondents did not even find the message on the phone anymore. For respondents who could not recall receiving the messages, interviewers read out the messages to them to still elicit their opinion on wording and content.

When asked about which content they recall, the preventive practices are mentioned most, so that this seems to be the content that is absorbed most. Two also mentioned the risk group information. Only three said that they received new information through the messages, one in general, one regarding telemedicine and one regarding diabetes. This is in line with most respondents saying that they feel in general well informed about COVID-19 as they know about basic preventive measures, but are unsure regarding more details. The focus on the practice information is also apparent in the actions that they have or plan to take after receiving the messages. Most mention following the preventive practices, but rather as a reinforcement of what they will continue to be doing instead of starting a new habit. In addition, eight respondents mention explicitly that they will tell others about the information, two of those specifically that they will share it with elderly members of the family.

Furthermore, the respondents confirmed that the messages were perceived to be trustworthy. Stated reasons are that they consider the sender (Sehat Insaf Card program) to be a trustworthy source for health information, that the helpline number was mentioned for further information and that they were addressed by their name. Most also say that the message concerns them as they are aware of the pandemic and perceive the message as a good service for them. As mentioned above, the language and literacy barrier can be lifted to some degree through other household members and some respondents also like about the messages that they are sent in Urdu rather than English like some other official campaigns. Yet, some respondents say that they would prefer the messages to be sent in Pashto rather than Urdu language or even prefer phone calls as they are more personal and accessible for illiterate people. These will remain the limits of an SMS-based intervention, which we have addressed to the degree possible by including the name, a reference to the family as well as the chance to call the helpline for more information. All in all, these results give us confidence that the messages will be received well if read and have the potential to reflect in preventive action.

7.3. Administrative information

Funding

This work was supported by the German Research Foundation's Research and Training Group 1723 "Globalization and Development" (DFG: RTG1723), Friedrich-Alexander Universität Erlangen-Nürnberg and Khyber Medical University.

Ethics

All necessary ethics approvals are in place.

Ethics approval and confirmation of adherence to European data protection laws issued by University of Göttingen's ethics commission

Ethics approval was issued by the Research Ethics Committee at Khyber Medical University with the registry number Dir/KMU-EB/IT/000784.

Data availability

The survey-based data that support the findings of this study are available upon reasonable request in Göttingen Research Online at https://doi.org/10.25625/VZJAFY , and will be made publicly available after publication of this study. The questionnaires are publicly available in the repository. Full replication data is available in the project repository at the Chair of Sebastian Vollmer.

Acknowledgements

We thank the Social Health Protection Initiative of Khyber Pakhtunkhwa for their support, and ICU healthcare, particularly Raheel Shahab for making the telephonic data collection and fielding of the intervention possible. We gratefully acknowledge the effort of our dedicated enumerators from Khyber Medical University under the coordination of Safat Ullah and Asad Ullah, as well as the support of student research assistants from the RTG "Globalization and Development", and Friedrich-Alexander Universität Erlangen-Nürnberg. We further thank colleagues, participants of several seminars and conferences for fruitful discussions and comments.