Marcus, Maja E.; Reuter, Anna; Rogge, Lisa; Vollmer, Sebastian

Working Paper

The effect of SMS reminders on health screening uptake: A randomized experiment in Indonesia

Discussion Papers, No. 284

Provided in Cooperation with:
Courant Research Centre 'Poverty, Equity and Growth in Developing and Transition Countries', University of Göttingen

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. 284

The Effect of SMS Reminders on Health Screening Uptake: A Randomized Experiment in Indonesia

Maja Marcus, Anna Reuter, Lisa Rogge, Sebastian Vollmer

August 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 SMS Reminders on Health Screening Uptake: A Randomized Experiment in Indonesia

Maja Marcus[1], Anna Reuter[1,2], Lisa Rogge[1,3], Sebastian Vollmer[1]

[1] Department of Economics & Centre for Modern Indian Studies, University of Göttingen; [2] Heidelberg Institute of Global Health, University of Heidelberg; [3] Leibniz University of Hannover

Abstract

While the burden of non-communicable diseases is rising in low- and middle-income countries, the uptake of screening for these diseases remains low. We conducted a community-based RCT in Indonesia to assess whether personalized and targeted text messages can increase the demand for existing public screening services for diabetes and hypertension in the at-risk population. Our intervention increased screening uptake by approximately 6.6 percentage points compared to the pure control group. Among those, who received and read the messages, the effect size is 17 percentage points. The intervention appears to work through a reminder rather than a knowledge effect. We conclude that text messages can be a cheap and easily scalable tool to reduce testing gaps in a middle-income country setting.

Keywords: Health, Noncommunicable Diseases, Information, Health Systems, Screening Uptake, mHealth, text message reminder.

JEL codes: C93; D83; I12; I15; I18; O15

Study pre-registration: Marcus M, Reuter A, Rogge L, Vollmer S. 2019. "A Mobile Phone-based Intervention to Improve Health Screening Uptake: A Randomized Experiment in Indonesia." AEA RCT Registry. November 2019. https://doi.org/10.1257/rct.5047-1.0.

1 Introduction

The ongoing epidemiological transition in low- and middle-income countries (LMICs) raises new challenges for their health systems. While the burden of infectious diseases remains high, non-communicable diseases (NCDs) are on the rise. Many of these diseases require a care very different from infectious diseases: They can be tackled effectively many years before individuals notice symptoms, and before severe complications develop. At the same time, individuals must be aware of this “invisibility” and take up preventive health behavior early on.

Diabetes and hypertension screening can be seen as a special case of preventive health behavior, for which it is not the aim to avoid an illness altogether but to detect a prevalent condition early enough to avoid or postpone complications. Screening is possible at very low costs, and at very early stages, behavioral changes can be sufficient to control these conditions. Yet, screening for diabetes and hypertension is underutilized in many LMICs (Geldsetzer et al., 2019; Manne-Goehler et al., 2019), even in settings with a free and easily accessible screening infrastructure, such as Indonesia.

In this study, we test whether a low-cost, low-touch text message intervention can increase the uptake of hypertension and diabetes screening in Indonesia. To understand the potential effect better, we explicitly test whether the intervention can transport new information, and whether risk aversion and patience are mediating the effect. Lastly, we examine household spillover effects to see whether the intervention can be effective beyond the direct message recipient.

We assessed these research questions via a community-based randomized controlled trial, in which half of the participants received the full intervention and half is the pure control group. The treatment group received two sets of three text messages, with each set sent before one of the monthly village screening dates between January and March 2020. The messages called upon the recipients to attend screening at the specified time and place and gave short information on their elevated risk, the necessity, and the benefit of screening. The intervention was targeted at individuals over the age of 40, who are at increased risk to develop diabetes or hypertension and should be screened once a year in accordance with WHO PEN screening guidelines (WHO, 2010). We randomly sampled 2,006 participants from two districts in Aceh province in a two-stage stratified design. Baseline data was collected in November and December 2019 and endline data was collected approximately one month after the last screening date via telephone surveys as the COVID-19 outbreak did not allow for in-person interviews.

We find that the intervention increased the uptake of screening services from 33% to 40%, which is a 6.6 percentage point or a 20% increase compared to the control group. For respondents who received at least one full set of messages and could remember any message content, this effect size increases to 17 percentage points. The text messages seem to work as a reminder for screening: While there is an overall increase in the uptake of screening, there is no impact on knowledge related to the text message or general disease knowledge. Respondents primarily remembered the content on the logistics and the advice to get screened. The only new information, which is remembered by a quarter of the respondents who recall any content is the fact that their age group implies a higher risk for hypertension and diabetes. In addition, the treatment effect is driven by attending screening at the health center rather than the specific village screening meeting that was mentioned in the messages. The treatment effect does not seem to differ across time and risk preferences. We cannot detect any spillovers to other household members.

In a standard model, investment in preventive health care such as screening would be the result of the monetary and non-monetary costs and benefits of the single options, as well as the time horizon over which they occur (Dupas & Miguel, 2017). In such a world, the individual's investment in preventive health care is optimal for the individual, and the societal optimum could be reached by changing the cost structures. However, in reality, an underinvestment in preventive health care is observed (Kremer et al., 2019). This underinvestment can be the result of various factors, such as inaccurate or motivated beliefs, trust, present bias, or limited attention.

Previous studies showed that preventive health behavior can be improved by both new information and reminders via text messages. In LMICs, text messages have been found effective to increase immunization rates among children (e.g. Jacobson Vann (2018)) or specific preventive behavior like dengue prevention (Dammert et al., 2014). As NCDs are rather new to the disease burden, community-wide screening programs, their benefits and relevance might not yet be salient to the population. Thus, it is unclear whether light-touch interventions such as text messages are sufficient to increase screening uptake, even if they were proven effective in high-income countries with a longer tradition of screening (e.g. Sallis et al. (2019)). To our knowledge, the only text message interventions addressing screening in LMICs focus on diseases that are very different from diabetes and hypertension, for example sexually transmitted diseases (Taylor et al., 2019). Other interventions to increase screening demand for diabetes and hypertension in LMICs are also rare; the only other study we know of uses a more intensive treatment, in-person scripts and pharmacy vouchers, in Armenia (de Walque et al., 2020; Gong et al., 2020). We expand this literature by demonstrating that text message reminders can also be an effective tool to increase diabetes and hypertension screening in the general population at risk in a middle-income country. Beyond the main treatment effect, we contribute to the scarce evidence base of spillover effects, particularly within the household, of preventive health interventions (Dupas & Miguel, 2017).

In the following chapter, we summarize the current prevalence of and screening for diabetes and hypertension in Indonesia. Then, we describe the experiment in detail by deriving the hypotheses from previous evidence and our own pre-studies, presenting the intervention design, estimation strategy, data collection and outcome definitions. The fourth chapter displays the experiment's results as well as implications for a potential scale-up. Finally, we conclude and give an outlook for further research.

2 Context

Similar to other LMICs, the burden of NCDs is rising in Indonesia. From 1990 to 2017, the share of NCDs in causes of death rose from 48% to 75% (IHME, 2018). In 2017, hypertension and diabetes were among the top three risk factors for morbidity (IHME, 2018). The most recent national health survey from the Ministry of Health revealed that diabetes prevalence has risen to 11% and hypertension to 34% (Riskesdas, 2018), both above the global average. To battle this trend, the national government has started implementing targeted prevention programs. In the last decade, nationwide programs were established to integrate a division responsible for NCD needs in every community health center (*Puskesmas*) (Mahendradhata et al., 2017).

One main effort is the village screening program *Posbindu (Pos binaan terpadu*). Once per month, trained nurses from the local *Puskesmas* offer information as well as screening and monitoring services for various NCDs to the general population at a central place within each village. This basic service is free of charge for the user and financed through a combination of the *Puskesmas* and village budget. At the village level, community health workers (*kader*) are responsible for organizing and advertising the meetings. In addition to *Posbindu*, it is possible to get basic screening at the local *Puskesmas* at all times. However, the national health survey shows that the general population has rarely used the NCD screening services so far. About one third of those aged above 45 report that they never had their blood pressure checked, and around 70% never had their blood sugar level checked (Riskesdas, 2018).

This pattern of high diabetes and hypertension prevalence and low screening uptake is also observed in our study region in Aceh province: the diabetes and hypertension prevalence is slightly above the national mean, and reported screening rates were below the national average in 2018 (Riskesdas, 2018). A focus group discussion with 12 *kaders* from our study area revealed that *Posbindu* tends to be visited by elderly women and those who were already diagnosed[1]. The *kaders* perceive it as more difficult to motivate the general population to attend the meetings even though sufficient time and equipment would be available. In addition, the province has close to universal health insurance coverage for over a decade, which makes it a suitable setting to study the demand-side barriers to screening uptake.

3 Method

3.1 The Intervention

Our intervention is a repeated set of SMS text messages on the necessity and logistics of diabetes and hypertension screening. It was designed to address disease misperceptions as well as behavioral barriers to screening uptake. The intervention was piloted in mid-January 2020 (see appendix D) and fielded from late January until March 2020.

Targeted mechanisms

As a high prevalence of NCDs is a rather new phenomenon in LMICs, individuals might not yet be aware of the role of screening as preventive health behavior, or might not have internalized regular check-ups. Text messages on screening dates might tackle several of these barriers: They might convey new information, thus update beliefs, make the screening decision more salient to the individual, thus serving as a reminder, or introduce a deadline to be screened.

To find out which factors keep at-risk individuals from taking up screening in the Acehnese context, we conducted a qualitative and a quantitative pre-study[2] (see Table A 2 for the detailed study timeline). For the qualitative arm, twelve in-depth semi-structured interviews with individuals from the target population were conducted in November 2019. These findings were quantified and extended in the quantitative baseline data collection from late November until December 2019 (see chapter 3.3 for data collection details).

These pre-studies showed that the majority of our respondents were informed about the main characteristics of hypertension and diabetes, as well as the possibility to screen free of charge. There are some perceived non-monetary costs such as fear of diagnosis and the notion that preventive health programs are designed for the elderly or women, but no strong stigmatization. On the other

[1] The focus group discussion was part of our pre-study to gather information on the supply-side perspective.
[2] The detailed design and findings will be made available in a separate paper.

hand, respondents are aware that early treatment initiation can help and that especially diabetes likely leads to high treatment costs. However, to most respondents it was not salient that their age implied a higher risk for both conditions, and most did not know that one could have them without feeling any symptoms. Studies from other parts of Indonesia confirm that even if individuals could identify risk factors, the own susceptibility was underestimated (Pujilestari et al., 2014), and even diagnosed respondents did not yet internalize that the need for screening does not depend on feeling ill (Rahmawati & Bajorek, 2018). Informing individuals about the need for screening independent of symptoms and their age-based risk might thus increase screening uptake.

Furthermore, forgetfulness and limited attention might prevent screening uptake. Reminders and fixed dates might simply make the decision for screening more salient and induce planning (Milkman et al., 2013), or increase the perceived urgency of screening. Similarly, evidence from other LMICs suggests that present bias can be a substantial barrier to screening uptake, as individuals postpone the health investment infinitely (Kremer et al., 2019). Deadlines can be efficient countermeasures as they signal that on the deadline, individuals cannot decide between now or later, but only between now or never (Kremer et al., 2019). Hence, individuals might not procrastinate the health investment any longer, but might be inclined to take up screening at the deadline. While the screening date is a non-binding deadline, the mere notion that missing the date implies a waiting period of one month might be effective to reduce naïve procrastination (O'Donoghue & Rabin, 2015).

Previous studies showed that impatient individuals are less likely to seek screening (Picone et al., 2004), resulting in a higher risk of underdiagnoses (Kim & Radoias, 2016). Information on the urgency of early action might reinforce this heterogeneity, by making the time sensitivity more salient, while deadline setting might help especially impatient individuals to take up screening. Similarly, more risk-averse individuals invest more in preventive health in some cases (Tsaneva, 2013), but not in all (Goldzahl, 2017; Picone et al., 2004), depending on how uncertain the outcomes of screening and treatment are perceived (Selden, 1993). Thus, the information conveyed in text messages might impact screening demand differently for relatively more and relatively less risk-averse individuals.

Finally, text messages could impact individuals beyond the targeted respondents due to information sharing, social learning, or mere convenience when respondents are accompanied to the screening facility. Spillovers of health interventions are rarely examined (Dupas & Miguel, 2017), but are of interest when analyzing the overall impact of an intervention. In the case of text messages, this might be a special concern, as they can be shared easily.

Thus, to assess the effectiveness of the intervention, we test the following hypotheses:

H1: The intervention increases screening uptake of the message recipient.

H2: The intervention increases screening and disease knowledge.

H3: There is a heterogeneous treatment effect along risk and time preferences.

H4: The intervention increases screening uptake of other household members.

Content & Personalization

The messages' content included the village-level *Posbindu*[3] screening date and location as well as selected information about hypertension and diabetes. We opted to emphasize the benefits of early screening uptake, in order to positively frame the messages, rectify respondents' misconceptions, and confirm their correct beliefs. Furthermore, as very few respondents were aware of age being a significant risk factor for diabetes and hypertension, we included this information to increase relevance and urgency for the recipients. Also, we included a note that the community health worker (*kader*) or the community health center (*Puskesmas*, abbreviated to PKM) can be contacted for further information. This aimed at increasing the trustworthiness and legitimacy of the messages, while at the same time providing respondents with contacts should any questions arise. To maximize their potential impact (e.g. Head et al. (2013)), the messages were personalized by providing village-level information, addressing the age of the recipient, as well as including the recipient's name in the greeting.

Based on these considerations, we formulated the following messages (see Table A 1 in the appendix for the translation of each message):

Message 1: Greetings [Mr/Ms name], do you know that [diabetes|hypertension| does not always show symptoms but can be treated better if detected earlier. Check for FREE at POSBINDU [date].

Message 2: Greetings [Mr/Ms name], do you know that people over 40 years old have a high risk of diabetes & hypertension? Ask kader / PKM & check for FREE at POSBINDU [date].

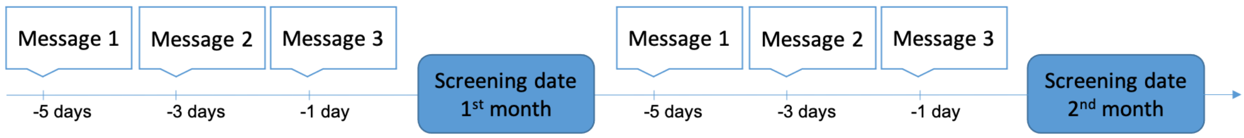
Message 3: Greetings [Mr/Ms name], remember to benefit from a FREE diabetes and hypertension CHECK in POSBINDU tomorrow morning at [place within the village]. Contact nearest kader or PKM.

[3] 17 out of 146 villages did not have a *Posbindu* screening during our study period. In these cases, participants were invited to the *Posbindu* in a neighboring village as participation is not restricted to village residents.

Implementation

Each individual in the treatment group received six SMS messages to the telephone number that s/he chose to be his/her contact number at baseline. As depicted in Figure 1, three messages were sent before the first village screening date and three were sent before the second date one month later. In the first cycle, the first message addressed diabetes, while in the second cycle, it addressed hypertension. In both screening cycles, messages were sent five days, three days and one day before the screening date. For 12 respondents in the treatment group, the first screening date took place end of January 2020, whereas for everyone else in the treatment group it took place in February.[4] The screening dates were enquired by our local research assistants from the respective *Puskesmas* up to two weeks before the start of the intervention to ensure their accuracy. As the *Puskesmas* only coordinates the screening services for all the villages in their catchment area, and the organization at the village level is done by the village health worker, we do not expect this enquiry to have any supply side effects.

Figure 1. Intervention timeline



The messages were sent out by the research team using the bulk SMS gateway provider *bulkgate*. We received delivery reports from the portal stating which messages failed to be delivered.

Treatment assignment was done in a random draw after baseline data collection in Stata 14 using the procedure proposed in DIME (2019). Half of the phone numbers were randomly allocated to the treatment group, which received the full intervention, while the control group received no intervention. Interviewers were fully blinded to treatment assignment and respondents were not aware of the existence of a control and treatment group throughout the study.

3.2 Estimation Strategy

We assess the impact of our intervention using intention-to-treat and local-average-treatment-effect estimates. Our regression specifications include the following outcome, treatment, and control

[4] To not interfere with newly implemented recommendations of social distancing, SMS were no longer sent after March 24, 2020, such that 10 people did not receive the full second cycle of the text messages. Most of the intervention period was not affected by the COVID-19 pandemic as *Posbindu* typically takes place in the beginning of a month and the second treatment cycle was therefore finished for most participants in early March when case numbers were still very low in Indonesia (and none in Aceh) and there were no restrictions in place.

variables, all of which were specified in the pre-analysis plan and implemented accordingly (Marcus et al., 2020).

Outcome Variables

Our primary outcome is screening uptake, which is measured in two ways. First, we use self-reported data at endline on whether respondents went to any diabetes or hypertension screening within the intervention period.[5] Secondly, we measure whether respondents went to at least one of the two *Posbindu* dates specified in our text message intervention.

Secondary outcomes are SMS-related knowledge, broader diabetes and hypertension knowledge, and household spillovers. SMS-related knowledge aims to capture the direct effect of the information that is transmitted in the messages. This is measured by a count index from 0 to 7, which increases by one for each correctly answered question that relates to the message content. All dimensions are measured by separate survey items that are part of the larger block of knowledge and screening questions (refer to appendix Table A 4 for the list of questions). We measure broader diabetes and hypertension knowledge with an index derived from a model of the determinants of health seeking behavior (Becker, 1974; Janz & Becker, 1984). The index includes items that can be influenced by information into a clear direction. An increase in the index therefore reflects both an increase in knowledge and should, as the model hypothesizes, increase the propensity to take up screening services. We measure the single dimensions using the survey items displayed in appendix Table A 5. For the main results, we use a count index that increases by one with each correctly answered knowledge question. To test the sensitivity of this result, we employ principal component analysis to reduce the dimensions to one variable, weighted by their explanatory power. This index gives a holistic picture of health knowledge with a focus on diabetes and hypertension.

We measure household spillovers through a binary variable indicating whether any other household members went for diabetes or hypertension screening within the intervention period.

Treatment Status

Treatment is defined in two ways. First, we categorize respondents into treatment and control group according to their randomized status. Secondly, we define a "treatment exposure" variable, which indicates whether the respondent received all three messages in one month and can recall the content of at least one message. The former is measured using delivery reports from the bulk SMS provider.

[5] We further pre-specified the aim to measure screening uptake across all villages in the sample districts using *Posbindu* attendance rates from administrative data, but full access could not yet be granted.

The latter is a self-reported measure collected at endline. It is based on listing at least one of the elements of our text messages when asked about the content of the NCD/ screening related message in an unaided recall question, if the respondent claims to have received such a message.

Variables for heterogeneous treatment effects

We measure risk and time preferences with one self-reported baseline survey question each, taken from and validated by the Global Preferences Survey (Falk et al., 2016, 2018). Patience is elicited by asking respondents to indicate how generally willing they are to give up something today in order to benefit from it in the future (on a scale from 0 to 10). Willingness-to-take risks is elicited by asking respondents to indicate on a scale from 0 to 10 how generally willing they are to take risks.

Control Variables

We measure age, sex, education, and phone ownership[6] using self-reported survey questions. Furthermore, we construct a wealth index based on self-reported asset ownership using the standard DHS approach. All control variables were elicited at baseline.

Regression Specifications

We estimate treatment effects on primary and secondary outcomes in the following framework:

a) Intention-to-treat (ITT)

$$Y_i = \alpha + \beta Treat_i + \delta Control_i + \varepsilon_i \quad (1)$$

where *Y* is our outcome variable (screening uptake in the main specifications and household spillover effects, SMS-related knowledge, and broader hypertension and diabetes knowledge in secondary analyses), *Treat* is an indicator variable for treatment status, and *Control* denotes the variables age, sex, education, wealth, and phone ownership.

b) Local Average Treatment Effect (LATE)

Additionally, we estimate the local average treatment effect using an instrumental variable approach (Imbens & Angrist, 1994). Specifically, we use assigned treatment status to instrument the treatment exposure variable.

[6] Due to a technical problem, phone ownership was not elicited for 7 individuals. We created a separate indicator for missing phone ownership information to keep them in the estimation sample. Neither phone ownership nor the indicator are significantly different from zero in the regressions.

$$Exposed_i = \eta + \theta Treat_i \ + \pi \ Control_i + \upsilon_i \quad (2)$$

$$Y_i = \alpha + \beta \widehat{Exposed}_i \ + \delta \ Control_i + \varepsilon_i \quad (3)$$

We explore potential heterogeneities in treatment uptake along time and risk preferences using the following specification:

$$Y_i = \alpha + \beta Treat_i + \gamma Trait_i + \theta Trait_i * Treat_i + \delta \ Control_i \ + \varepsilon_i \quad (4)$$

Where $Trait$ is the respective continuous indicator of baseline risk or time preference.

For all main hypothesis, p-values will be adjusted for multiple hypothesis testing following the Benjamini-and-Hochberg method (Benjamini & Hochberg, 1995) as a robustness check.

3.3 Data and Sample Characteristics

The baseline sample was drawn in a two-stage stratified random sampling procedure. First, we randomly drew 147 villages from a complete list of villages in the districts Aceh Besar and Banda Aceh. This draw was stratified by district to have an equal number of villages from the mostly rural Aceh Besar and the mostly urban provincial capital Banda Aceh (refer to appendix Figure A 1 for a map of the sampled villages). Within the villages, we selected households using a random walk following the procedure described in appendix B2. Around half of the identified houses were found to be occupied, out of which 85% agreed to undergo the short eligibility check. The eligibility criteria ensured that the respondent would be recommended to be screened on a yearly basis (being over the age of 40[7]), and is neither diagnosed with diabetes or hypertension nor adheres to the recommended screening schedules. Out of those who did the eligibility check, one third of households was eligible[8]. If several household members met the inclusion criteria, one was randomly chosen as respondent. This yielded a sample of 2,006 individuals[9].

The endline survey was conducted from end of March until beginning of May 2020 and was shifted to phone interviews due to the outbreak of the COVID-19 pandemic (call pattern described in appendix Figure A 2). The analysis sample comprises of 1,386 individuals, 704 of the control and 682 of the

[7] We set the upper age limit of 70 to ensure that the respondent is able to complete the interview. Refer to appendix B1 for a detailed list and reasoning for each in- and exclusion criterion.
[8] Out of those ineligible, 36% did not have a member between the ages of 40 and 70, 28% had a member with a prior diabetes or hypertension diagnosis, 15% went for regular screening, in 8% of households eligible members were not at home and only 6% of households had to be excluded because they did not have any mobile phone (Table A 3).
[9] An additional 94 baseline respondents were excluded before randomization as they had not supplied us with a valid telephone number until the end of data collection. This also led to the drop-out of one village in the final sample.

treatment group. This implies a re-contact rate of slightly more than 70%[10], which is high for a telephone survey, but lower than we expected from the planned in-person endline data collection. The endline sample is hence slightly smaller than was deemed necessary in the power calculation (see appendix B3).

We depict endline sample characteristics across treatment and control group in Table 1. The average age of the respondents is 50 years, slightly more than 60% of the sample population is female, and 73% have at least lower secondary education. Literacy in Bahasa Indonesia is over 90%. Compared to the same age group who owns a mobile phone in the representative national socio-economic survey (SUSENAS 2017), our respondents are to a higher proportion female and slightly less educated, but generally similar across basic sociodemographic characteristics (see appendix Table A 7).

Treatment and control group were balanced across all key variables of interest at baseline, except for phone ownership (see appendix Table A 6). At endline, respondent age is slightly lower in the treatment group and the share of phone owners remains slightly higher. As displayed in appendix Table A 8 to Table A 10, there was no differential attrition between treatment and control group, but respondents who were lost to follow-up seem to be to a higher proportion female, less educated and wealthy and to a lesser proportion phone owners. These differences likely occur due to the need to shift the administration of the survey to the phone: Additional analyses reveal that phone ownership is more likely across younger, male and better educated individuals from households in the fifth wealth quintile. If controlling for all base characteristics simultaneously, having no educational degree and not being the phone owner are the only significant drivers of attrition (appendix Table A 11).

[10] 1,409 respondents could be re-interviewed, but due to missing information on the month of screening for 23 respondents, and missing information on age, gender and wealth quintile for one respondent each, the final analyses sample consists of 1,386 respondents.

Table 1 Endline sample characteristics across treatment and control group

	Control group			Treatment group			
	Mean	Standard deviation	N	Mean	Standard deviation	N	p-value
Age	50.26	8.22	704	49.52	7.85	682	0.088
Female	0.64	0.48	704	0.61	0.49	682	0.285
Highest level of schooling							0.850
None	0.04	0.19	704	0.03	0.18	682	
Primary	0.23	0.42	704	0.24	0.42	682	
Junior Secondary	0.23	0.42	704	0.21	0.41	682	
Senior Secondary	0.35	0.48	704	0.36	0.48	682	
Tertiary	0.15	0.36	704	0.17	0.37	682	
Literacy	0.91	0.29	568	0.93	0.26	555	0.160
Wealth quintile							0.389
1	0.22	0.41	704	0.19	0.39	682	
2	0.19	0.39	704	0.18	0.38	682	
3	0.19	0.39	704	0.22	0.41	682	
4	0.20	0.40	704	0.19	0.39	682	
5	0.20	0.40	704	0.22	0.42	682	
Own phone	0.64	0.48	700	0.68	0.47	679	0.101
Joint F-test							0.277

Means, standard deviation and number of observations of main respondent characteristics by treatment group; p-values based on t-tests of difference in mean between treatment and control group, except in the case of education, wealth quintile, and the total, where we used F- tests on joint significance of the different levels respectively variables.

According to the message delivery reports, 97% of treatment group individuals received at least one full cycle of intervention messages before one of the Posbindu dates. For 84% of our sample, we have also self-reported measures of exposure[11]: Out of those who received at least one full cycle, 30% could correctly recall at least one item of the message content, indicating that the messages were not only delivered, but also received, read, and understood. Consequently, around 28% of the treatment group constitute the exposed group in the LATE estimation.

4 Results

4.1 Screening uptake

We find that our intervention had a positive effect on screening uptake of the message recipient (Figure 2). In the intention-to-treat analysis, treatment increased screening uptake from 33% in the control to 40% in the treatment group. This is an increase by around 6.6 percentage points (p.p.) or 20% at a statistical significance level of less than 1%. This effect is robust in all pre-specified model

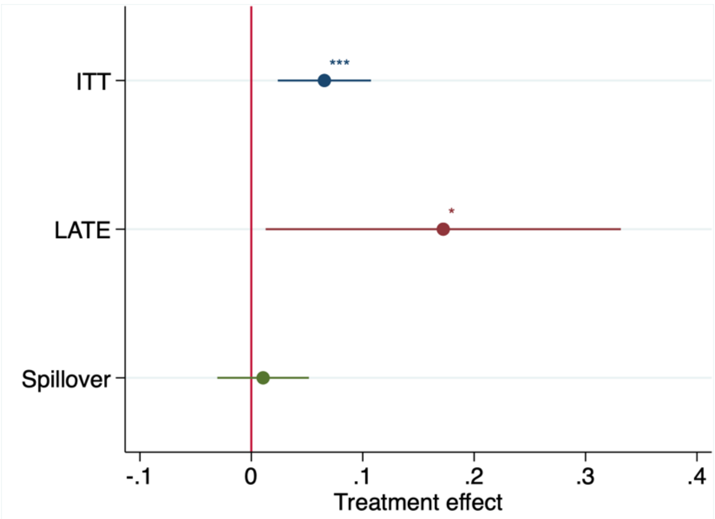
[11] As the questions about message content were asked only in the very end of the interview, the estimation sample for the LATE excludes 204 respondents who terminated the interview before this question. Respondents in this subsample are to a higher proportion male, to a lesser proportion phone owner, but otherwise similar.

specifications (Table A 12), adjustments for multiple hypothesis testing (Table A 13) or alternative estimation strategies (Table A 15).

When treatment exposure (having received the full cycle of text messages and being able to recall message content) is instrumented by treatment status, the effect is more than twice as high (17 p.p.), which indicates the potential for a higher treatment effect if barriers to message reception are reduced. In section 4.4, we explore the main barriers from sending up to acting upon the messages in detail. It needs to be mentioned that the precision of the LATE estimate is lower than for the ITT due to the above-mentioned reduction in the sample and hence a loss in statistical power.

The effect on screening uptake of the message recipient did not lead to within-household spillover effects. We do not find evidence for other household members taking up screening more often, neither in the aggregate as displayed in Figure 2, nor when restricting the sample to household members in the same age group as our respondents (between the age of 40 and 70). Receiving the messages through another household member's phone or a family phone could have increased other household member's attention to the messages, but even if accounting for phone ownership, we do not find evidence for substantial spillover effects (Table A 20).

Figure 2 Treatment effect on screening uptake of the message recipient and household members.

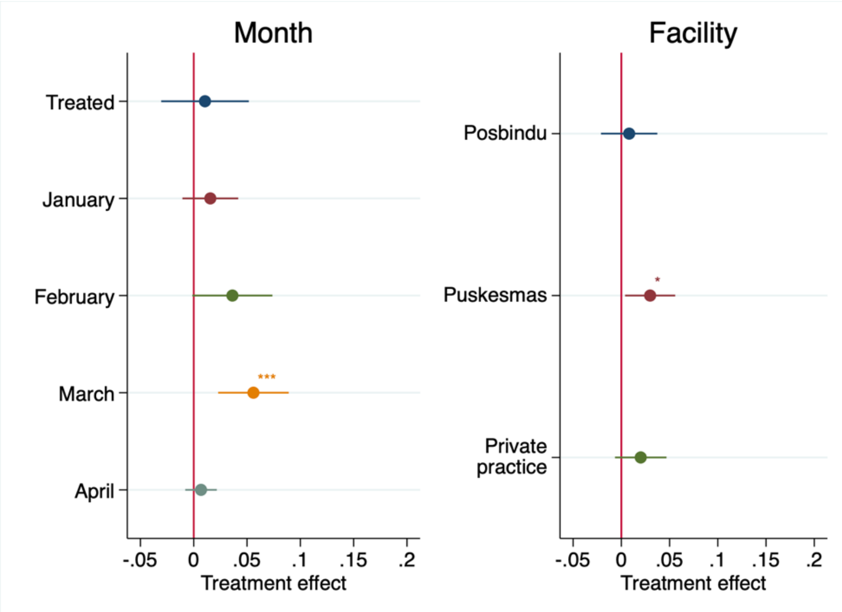


*Point estimates of the treatment coefficient from equation 1 (ITT), the instrumented treatment coefficient from equation 3 (LATE) for the message recipient and other household members (ITT), controlling for age, gender, wealth and phone ownership; see Table A 12 for tabular display with and without covariates; displayed with 90% confidence intervals; * p < 0.1, ** p < 0.05, *** p < 0.01*

To understand the treatment effect of the message recipient better, we further examine the timing and location of screening (Figure 3). For all respondents, we see low screening uptake in November and December, and increasing visits to testing facilities from January on. Even though treatment is positively correlated with screening uptake in all months, it is only statistically significantly different from zero in March and is comparable to the size of the aggregate treatment effect. This suggests a concentration of the treatment effect after having received the second set of text messages. When disaggregating the treatment effect according to screening provider, we see that the effect is not driven by treatment group respondents going to the specific *Posbindu* meeting that was mentioned in the messages, but rather by going for screening at the *Puskesmas*. Even though the focus of the messages was on the *Posbindu* meeting, the *Puskesmas* was always mentioned as a point of contact, and might have posed a suitable alternative for some respondents.

Apart from merely going for screening, we see that this uptake translated in significantly higher blood pressure testing rates and checks of the medical history in the treatment group. Blood glucose testing, physical measurements, and other blood checks are also positively correlated with treatment, but not statistically significantly different from zero (Table A 23).

Figure 3 Treatment effect on message recipient screening uptake by month and facility



Point estimates of treatment coefficient from equation 1 with different binary screening uptake indicators as outcomes (coded as 1 if the individual indicated to have gone to screening in the respective month/ facility and 0 otherwise); controlling for age, gender, wealth and phone ownership; see

*Table A 21 and Table A 22 for tabular display; displayed with 90% confidence intervals; * p < 0.1, ** p < 0.05, *** p < 0.01.*

4.2 Channels

We find that the intervention did not increase knowledge, as shown in Table 2. We can neither detect a treatment effect for the specific knowledge items mentioned in the text messages, nor for general diabetes and hypertension knowledge. These patterns hold when defining the indices via PCA rather than as a count index (Table A 16), and for each element of the respective index (Table A 17, Table A 18, Table A 19). In addition, the point estimates are small with rather precise confidence bounds, so that these results can be interpreted as a null effect. It is hence likely that the intervention increased screening uptake of the message recipient purely via a channel that does not imply an updating of beliefs through new information.

Table 2 Treatment effect on knowledge outcomes

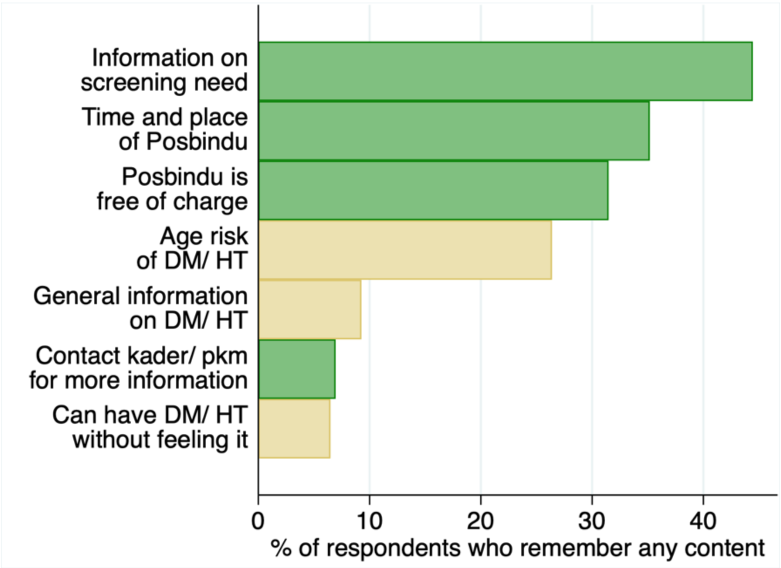
	(1) SMS knowledge	(2) SMS knowledge	(3) General disease knowledge	(4) General disease knowledge
Treated	-0.0009	-0.0029	-0.0365	-0.0570
	(0.0609)	(0.0610)	(0.0616)	(0.0597)
Covariates	No	Yes	No	Yes
Observations	1088	1088	1042	1042

ITT estimates on SMS-related and general disease knowledge indices following equation 1. Both indices are standardized to a sample mean of 0 and a standard deviation of 1. Covariates are age, gender, wealth and phone ownership. Standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 4 displays which information the respondents who report to have received any text message on *Posbindu* are able to recall. We see that these respondents tend to remember the actionable elements of the messages (green), rather than the disease information components (yellow). More precisely, the principal directive that the respondent should be tested for diabetes and hypertension is remembered most frequently – namely by 45% of all respondents, who self-reported being exposed to the treatment. This is followed by logistical components, as 35% and 31% of these respondents remember being told when and where *Posbindu* takes place as well as that it offers free NCD check-ups. We interpret this as evidence for making existing information more salient to the message recipients as even in the control group, almost all of the 44% of respondents who knew the Posbindu program were aware that it is free of charge and where it takes place.

Similarly, the reported reasons for no screening indicate that our intervention works through increased salience rather than shifts in beliefs: Nearly all respondents who did not attend any screening since the baseline visit reported they did not attend any screening because they were not ill (93%), and only few mentioned time constraints (15%). This pattern is similar to the reasons at baseline and fits the null effect on disease-related knowledge. Hence, more intensive interventions might be needed to alter the beliefs which prevent a large share of the population from regular screening.

Figure 4 Ability to Recall Text Message Components



4.3 Heterogeneous treatment effects

We cannot detect any heterogeneous effects across time and risk preferences (Table 3). In most cases, the standard errors are also too large to retain the original treatment effect. One reason for not detecting any heterogeneous treatment effects might be a weak correlation between screening and willingness to take risks and patience in our study setting. At baseline, we observed a significant correlation between patience and hypertension screening within the last year, but no correlation for willingness to take risk. Another reason might be that the endline sample is too small to detect any heterogeneity.

Table 3 Analysis of Heterogeneous Effects

	(1) Screened	(2) Screened	(3) Screened	(4) Screened
Treated	0.055	0.082	0.090	0.118**
	(0.051)	(0.051)	(0.057)	(0.057)
Willingness to take risk	0.001	0.007		
	(0.007)	(0.007)		
Treated x Willingness to take risk	0.001	-0.004		
	(0.010)	(0.010)		
Patience			0.005	0.008
			(0.006)	(0.006)
Treated x Patience			-0.006	-0.009
			(0.009)	(0.009)
Covariates	No	Yes	No	Yes
Obs.	1386	1386	1386	1386
Control group mean	0.3310	0.3310	0.3310	0.3310

*Results of regressing the binary screening indicator on the binary treatment indicator, the respective time or risk preference as well as their interaction following equation 4; controlling for message recipient age, gender, wealth, and phone ownership; Standard errors clustered at the phone number in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01*

4.4 Implications for scale-up

In the following explorative analyses, we further investigate the scale-up potential and limits to the effectiveness of the intervention. We first focus on what hinders message recipients from reading the messages and hence being exposed to the treatment to shed more light on the potential to reduce the discrepancy between ITT and LATE. Then, we explore differences in screening experience between the three main facility types to assess the role of accessing a specific screening service. Finally, we provide a cost estimate of this intervention.

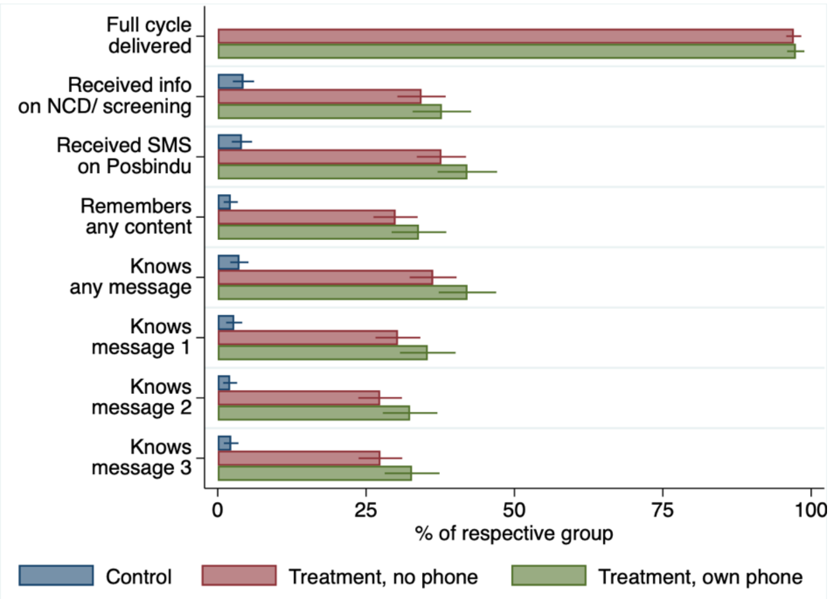
<u>*Treatment exposure*</u>

For an allocated message recipient to be exposed to the treatment, s/he needs to receive, become aware of, read, understand, and trust the messages. As stated above, message delivery by the provider does not pose a barrier. Rather, being aware or remembering to have received any information on screening appears to be the major barrier (Figure 5). Phone ownership appears to ease this barrier substantially: While 26% of the treated individuals without a phone remember to have received any information, the share increases to 37% among the treated phone owners. A main issue might be the transfer of the information from the owner to the respondent: 51% of the phone owners who were assigned by the respondents as contact person admitted they transmit messages only sometimes, rarely, or never (response rate: 40%). Once this barrier of becoming aware of the information is overcome, most respondents are able to remember some message content or remember to have received the messages after reading them out. Hence, with an increase in phone ownership over time, the exposure to the intervention can be expected to rise.

We do not find that illiteracy is a binding constraint to reading the messages as only 5% of the sample population reports to be illiterate and 80% face never or only rarely problems when reading Bahasa Indonesia. Alternatively, our messages might be ignored if there is already an overload of information via SMS. We find that around half of the sample receives any text message on a daily basis and on average around four messages per day. Even though this does not seem overly high, phone owners report to receive more messages. We also see that 90% of the respondents who receive SMS in general receive advertisements and 60% would like to receive less advertisement. However, our messages are rather perceived as an official announcement and not an advertisement, thus it is unlikely that our messages are perceived as a burden. This is strengthened by the statement that 68% of respondents, who recall receiving the text messages, report they found the information very relevant to them, and 30% report they found it somewhat relevant. Thus, associating the text messages with the health services might mitigate any information overload.

Taken together, any scale-up needs to consider that even though targeted more broadly, population groups who are more likely to be telephone owners (younger, male and more educated) will be more likely to be exposed to the intervention. See Table A 24 for a detailed list of socio-demographic and other baseline characteristics by different exposure measures.

Figure 5 Exposure to Treatment



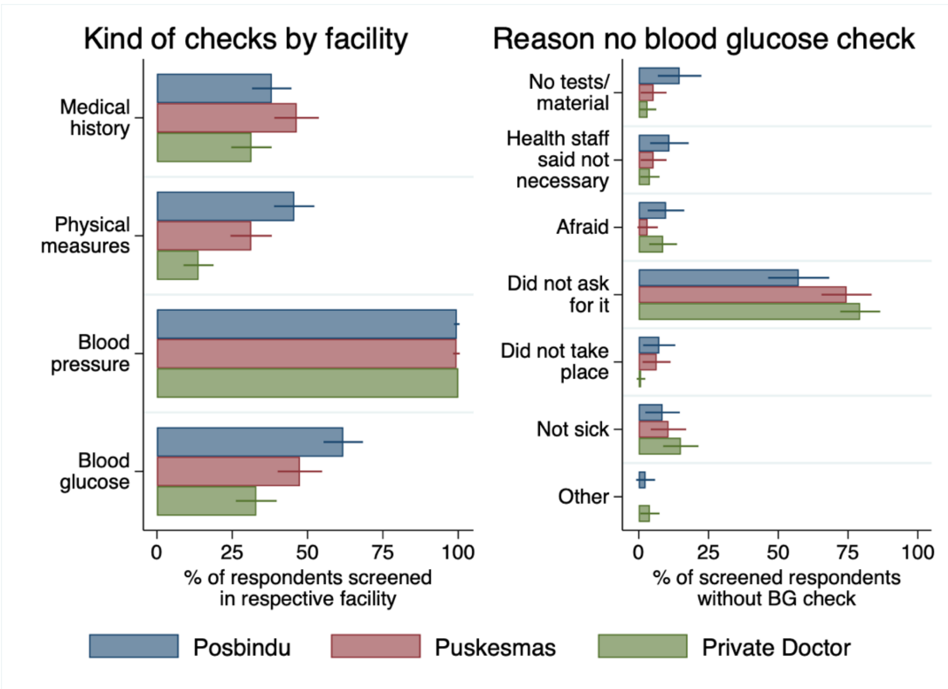
"Full cycle delivered" is based on the provider delivery reports, the remaining indicators are based on the respondent's self report at endline;"Knows any content" indicates whether the respondent could name any message content when asked in an open-ended question (compare Figure 4); "Knows any message" until "Knows message 3" is based on whether the respondent remembered the respective message when the enumerator read it out.

Screening services across facilities

Increased screening uptake can translate into improved control of the NCD burden the better the screening service. Our treatment effect is driven by respondents screened at the Puskesmas, but their recall of which services and recommendations for future screening were provided to them suggest that currently Posbindu offers the more comprehensive package. As depicted in Figure 6, nearly all respondents who reported to have undergone a screening report a blood pressure reading. However, which further checks were performed varied across facilities. While 62% of the Posbindu visitors had a blood glucose measurement, this only applies to 47% of the Puskesmas visitors and 33% of the visitors of private practices. In these two facility types, more than two thirds of the visitors who did not get a blood glucose check missed it, because they did not ask for this specifically. This might be caused by different reasons for visiting the respective facility type, but we cannot disentangle this further with our data.

Posbindu visitors were also more likely to report that they were asked to return for blood pressure screening another time, especially compared to visitors of private practices. As our treatment effect is mainly driven by increased use of the Puskesmas services, any potential scale-up might thus consider either increasing awareness towards blood glucose screening to ensure it is actively asked for at the Puskesmas, or stressing the benefits of Posbindu to nudge participants into the more specialized Posbindu.

Figure 6. Medical checks performed by facilities



Cost estimation

To improve the comparability of our text message reminders with other demand-enhancing interventions, we estimate the costs of our intervention per targeted person and per additionally screened person (Table 4). In the first column, we consider costs directly related to the intervention, i.e., costs of sending out the text messages and of inquiring the village-specific Posbindu details, assuming that any implementer would be able to target recipients using a register, such as a health insurance database. We base this estimate on the complete treatment group, rather than only the endline sample for a conservative estimate that assumes no treatment effect on the individuals lost to follow-up. In the second column, we additionally provide estimates on the screening costs occurring to the health system in the form of medical staff and material. We assume that a person presenting at a facility would take up 15 minutes of time with a medical practitioner, and price this using wage data from the National Statistical Office (Badan Pusat Statistik, 2021). In addition, we calculate the costs for blood glucose tests with a point-of-care machine, assuming that 47% of the individuals accessing the service are screened for diabetes (as observed in our sample). As every health worker has an own blood pressure monitor, no additional costs are borne for a blood pressure reading. For the scale-up, we assume that Posbindu dates can be transmitted directly to the implementer at a fix cost, such that these costs are not included in the scale-up calculation. On this basis, we estimate that a scale-up would cost IDR 5,277 or USD 0.38 per targeted person, and IDR 129,293 or USD 9.21 per additionally screened person.

Table 4. Cost estimates

	Intervention costs	Total costs	Scale-up (per Person)
SMS	4,651,101	4,651,101	4,500
Request for Posbindu dates	1,000,000	1,000,000	
Medical staff		640,313	638
Blood glucose test		140,121	140
Per targeted person	5,629	6,406	5,277
Per additionally screened person	137,899	156,943	129,293
Per targeted person (USD)	0.40	0.46	0.38
Per additionally screened person (USD)	9.83	11.18	9.21

All prices denoted in IDR, unless noted differently. Costs are calculated based on the targeted 1,004 respondents of the treatment group after the baseline. SMS costs were EUR 300 and are converted with an exchange rate of 15503.67 IDR/EUR. Costs for medical staff were taken from the National Statistical Office (BPS) as monthly net wages for employees in the health sector with university degree and doubled to receive an upper bound of gross wages to the health system (Badan Pusat Statistik, 2021). It was assumed that medical staff would spend about 15 minutes on each examination. It was assumed that point-of-care machines were used for the blood glucose check, as they are used at the Posbindu, such that one test would cost IDR 7,275, including lancet, stick, gloves, and disinfect. Costs for medical staff were calculated for the share of respondents who went to a screening facility due to the intervention (6%) times the share of treatment group respondents who were reached for the endline interview and for whom screening data was non-missing (68%). Costs for blood glucose tests were calculated for the share of respondents who went to a facility due to the intervention (6%) and conducted a blood glucose check (47% of the visitors) times the share of treatment group respondents who were reached for the endline interview and for whom screening data was non-missing (68%). USD were calculated using an exchange rate of 14032.02 IDR/USD. All costs were assessed between November 2019 and February 2020. If the targeted respondents who were not reached for the endline interview or for whom screening data is missing had the same treatment effect as the observed respondents, costs would reduce to USD 6.69 for the intervention costs, USD 8.04 for the total costs, and USD 6.70 for the scale-up costs per additionally screened person.

5 Discussion and conclusion

Like many other LMICs, Indonesia suffers from a high and increasing burden of diabetes and hypertension. Despite providing opportunities for easily accessible and free screening, uptake remains limited. Diabetes and hypertension screening are specific cases of preventive health behavior that can avoid or postpone complications rather than the disease itself, and are a relatively new component in the Indonesian health system. Thus, it is unclear whether light-touch policy measures proven effective in high-income countries, or for different preventive health behavior work in this context. We conducted a community-based RCT to test whether the uptake of screening programs can be increased with a low-touch text messaging intervention targeted at at-risk individuals.

We find that sending two sets of three text messages before two village-based screening meetings increased screening rates by approximately 6.6 percentage points from 33% in the control group. For participants who received at least one full treatment cycle and remembered any message content, this translates into an increase by approximately 17 percentage points. We do not find a significant difference in the SMS-conveyed or general disease knowledge between treatment and control group. Also, we cannot detect any spillover effects within households, or heterogeneous effects along levels of patience or willingness-to-take-risks.

The intervention appears to work as a reminder rather than conveying new information. Even though our pre-studies revealed gaps in disease knowledge, neither the information that was mentioned in the message nor a larger set of facts and beliefs about diabetes and hypertension changed as a result of the intervention. We find several hints that the intervention might have increased the salience of the decision to take up screening and hence rather works through addressing behavioral barriers related to procrastination or limited attention. First, the elements that respondents remember most from the messages are the general need for screening and its logistics, which were both widely known at baseline already. Secondly, message recipients react more strongly after receiving the second set of text messages and opt to get screened at the Puskesmas rather than the explicitly mentioned Posbindu meeting. Nevertheless, the awareness of a concrete date for screening might have been perceived as a deadline and pushed the recipient to no longer postpone asking for a preventive check-up at the Puskesmas at their convenience.

Possibly, the personalization of the text messages was effective in increasing the relevance for the recipients but did not give them the notion to share this information, such that no spillovers occurred within households. Alternatively, spillovers might exist but be too small to be detectable in our sample.

Similarly, we cannot detect heterogeneous treatment effects based on risk or time preferences. One reason might be the lack of a meaningful update of beliefs on disease risk and treatment efficacy. On the other hand, the countervailing forces of the lotteries of becoming ill and being effectively treated might cancel out any heterogeneous effects. For patience, however, we would have expected that the reminder channel alone would impact respondents with different degrees of patience differently.

The size of our treatment effect is comparable to other SMS interventions on preventive behavior in LMICs: With a risk ratio of 1.174, our findings lie between the results from the systematic reviews on immunization rates by Mekkonen et al. (2019) (RR: 1.11) and Jacobson Vann et al. (2018) (RR: 1.29). With an odds ratio of 1.284, the effect size is slightly lower than the average effect size of studies on STD detection as reported by Taylor et al. (2019) (OR: 1.73). Thus, even though the uptake of immunization or STD screening might underlie very different barriers compared to hypertension or diabetes screening, the impact of text messages can be similar. In addition to finding increased screening attendance after adding SMS reminders to routine invitations in the UK, (Sallis et al., 2019) found that adding the prompt to screen in a specific month increased the effectiveness, suggesting that mentioning a concrete deadline might counteract procrastination in this high-income setting. Similar to our results on knowledge transmission, recent evidence on broadcasting SMS to increase COIVD-19 preventive behavior found changes in behavior despite no updates in knowledge (A. Banerjee et al., 2020).

An advantage of text message interventions is their comparatively low cost. We estimate that our intervention costs USD 11.18 per additionally screened person, incorporating the costs of the screening service. A scale-up might decrease these costs even further, especially if screening dates can be centrally collected. Thus, such interventions can be used to reach out to wide parts of the population, such as the population over the age of 40. For people at higher risk due to preconditions, more intensive interventions might be a good addition to push screening rates even more, albeit at higher costs: Using personally delivered invitation letters and pharmacy voucher, (de Walque et al., 2020) measure an increase in screening rates by even 15 to 30 percentage points at about 60 USD per screened person. Hence, combining large-scale low-touch interventions as ours with intensive interventions in more selected higher risk groups might be a route to reach the population while keeping the costs balanced.

We conclude that our intervention is cost-effective and has the potential to be scaled up in the Indonesian setting, keeping in mind the limitations that are inherent to SMS interventions. First, being targeted and exposed to the intervention highly depends on owning and regularly using a mobile phone. This implies people who are more likely to own a phone, such as younger, male and more educated individuals are more likely to be reached, and not necessarily the most vulnerable. As mobile phone ownership, network coverage as well as familiarity in usage increases, so does the potential to reach a broader set of the population. As of now, we do not see evidence that our messages induced an overflow of information, but during implementation this needs to be monitored closely and implementers need to bear in mind to target carefully and keep messages to the necessary minimum. Secondly, who is reached by the intervention strongly depends on how the target population is sampled. At scale-up, collecting numbers by visiting households is likely not feasible and would increase the costs substantially. At the same time, previous literature established that personalization matters, such that mere broadcasting might not be advisable. Instead, drawing numbers from an existing register would be ideal. With the expansion of public health insurance in many middle-income countries, health insurances might be suitable implementers. In Indonesia, for example, the recently established, centrally administered health insurance JKN covers the majority of the Indonesian population and could likely target its members based on age and potentially even previous diagnosis.

This study comes with some limitations regarding the recruitment of participants and the telephonic endline data collection. Apart from being unfeasible for scale-up, we cannot rule out that our in-person baseline survey already worked as a reminder to take up screening 2-3 months prior to the intervention. Both treatment and control group saw higher propensities to be screened from January onwards, so that the high control group uptake might in part be driven by our baseline visit. However, we can still detect a systematic difference between treatment and control group, especially as time to the baseline interview increased. Secondly, measuring the main outcome as self-report is subject to the concern of misreporting. To minimize this concern, we added detailed follow-up questions on what happened at the screening visit and the consistency of the answers gives us confidence in the main result. Similarly, part of the reason that we do not find an update of beliefs could be that many knowledge questions were posed in a strict way, like asking for the risk factors in an unaided recall question. It might be that more nuanced updates of beliefs happened, but these are unlikely to explain the main treatment effect.

Switching the endline data collection to the telephone was the only possibility after the outbreak of the COVID-19 pandemic, but poses additional limitations. First, we could only re-interview 70% of the sample, with significant attrition across several socioeconomic characteristics. Though we do not expect that the attrition was selective due to factors other than the mode of contact, the true size of the treatment effect might be slightly different when taking the full initial sample into account. To the extent that phone ownership is correlated with both, a higher rate of recall receiving the message and a lower probability to be lost to follow-up, it is likely that our treatment effect would be slightly smaller in this case. Secondly, respondents may be less trusting over a telephone call in comparison to face to face interviews conducted in the privacy of their own home. As our study team visited the respondents during baseline, we think this problem might be less severe compared to phone surveys when the call is the first point of contact. To minimize this concern further, we assigned the enumerator who visited the respondent at baseline whenever possible and re-introduced our team and the survey in the beginning of the interview.

Our study opens several areas of complementing research. First, a scale-up study without baseline contact would be needed to validate the effectiveness of our study. Fielding the intervention in a larger sample would also offer the opportunity to test for the discussed mechanisms and heterogeneities more clearly. A second important extension would be to include longer-term outcomes such as regular or repeated screening. Beyond the intervention itself, our results showed that substantial misperceptions on screening recommendations prevail despite including this information in the messages, calling for designing and testing more intensive interventions to address this gap.

With the expansion of mobile phone coverage around the globe, policy makers gain access to a new toolbox of low-cost and low-touch interventions at scale. We show that text messages can induce preventive health behavior and reduce the screening gap for fairly new, yet severe contributors to the health burden of middle-income countries. As universal health coverage expands and is digitized, such text messages can become cost-effective and easily customizable measures to remind a target population of preventive health behavior and stimulate new health care habits.

6 List of References

Badan Pusat Statistik. (2021, August 11). *Average of Net Income per Month of Casual Worker by Province and Main Industry*. https://www.bps.go.id/subject/19/upah-buruh.html#subjekViewTab3

Banerjee, A., Alsan, M., Breza, E., Chandrasekhar, A., Chowdhury, A., Duflo, E., Goldsmith-Pinkham, P., & Olken, B. (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* (No. w27496; p. w27496). National Bureau of Economic Research. https://doi.org/10.3386/w27496

Becker, M. H. (1974). The Health Belief Model and Sick Role Behavior. *Health Education Monographs*, *2*(4), 409–419. https://doi.org/10.1177/109019817400200407

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Chavarría, E., Diba, F., Marcus, M. E., Marthoenis, Reuter, A., Rogge, L., & Vollmer, S. (2021). Knowing Versus Doing: Protective Health Behaviour Against COVID-19 in Aceh, Indonesia. *The Journal of Development Studies*, *57*(8), 1245–1266. https://doi.org/10.1080/00220388.2021.1898594

Dammert, A. C., Galdo, J. C., & Galdo, V. (2014). Preventing dengue through mobile phones: Evidence from a field experiment in Peru. *Journal of Health Economics*, *35*, 147–161. https://doi.org/10.1016/j.jhealeco.2014.02.002

de Walque, D., Chukwuma, A., & Ayivi-Guedehoussou, N. (2020). Invitations, Incentives, and Conditions: A Randomized Evaluation of Demand-Side Interventions for Health Screenings in Armenia. *World Bank Group Policy Research Working Paper*, *9346*.

DIME. (2019). *Randomization in Stata - DIME Wiki*. https://dimewiki.worldbank.org/index.php?title=Randomization_in_Stata&oldid=5389

Djimeu, E. W., & Houndolo, D.-G. (2016). *Power calculation for causal inference in social science: sample size and minimum detectable effect determination*. International Initiative for Impact Evaluation (3ie).

Dupas, P., & Miguel, E. (2017). Chapter 1 - Impacts and Determinants of Health Levels in Low-Income Countries. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of Economic Field Experiments* (Vol. 2, pp. 3–93). North-Holland. https://doi.org/10.1016/bs.hefe.2016.09.003

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global Evidence on Economic Preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692. https://doi.org/10.1093/qje/qjy013

Falk, A., Becker, A., Dohmen, T. J., Huffman, D., & Sunde, U. (2016). The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences. *IZA Discussion Papers*, *9674*, 66. https://doi.org/10.2139/ssrn.2725874

Geldsetzer, P., Manne-Goehler, J., Marcus, M.-E., Ebert, C., Zhumadilov, Z., Wesseh, C. S., Tsabedze, L., Supiyev, A., Sturua, L., Bahendeka, S. K., Sibai, A. M., Quesnel-Crooks, S., Norov, B., Mwangi, K. J., Mwalim, O., Wong-McClure, R., Mayige, M. T., Martins, J. S., Lunet, N., … Jaacks, L. M. (2019). The state of hypertension care in 44 low-income and middle-income countries: a cross-sectional study of nationally representative individual-level data from 1·1 million adults. *The Lancet*. https://doi.org/10.1016/S0140-6736(19)30955-9

Goldzahl, L. (2017). Contributions of risk preference, time orientation and perceptions to breast cancer screening regularity. *Social Science & Medicine*, *185*, 147–157. https://doi.org/10.1016/j.socscimed.2017.04.037

Gong, E., Chukwuma, A., Ghazaryan, E., & de Walque, D. (2020). Invitations and incentives: a qualitative study of behavioral nudges for primary care screenings in Armenia. *BMC Health Services Research*, *20*(1), 1110. https://doi.org/10.1186/s12913-020-05967-z

Head, K. J., Noar, S. M., Iannarino, N. T., & Grant Harrington, N. (2013). Efficacy of text messaging-based interventions for health promotion: A meta-analysis. *Social Science & Medicine*, *97*, 41–48. https://doi.org/10.1016/j.socscimed.2013.08.003

IHME. (2018). *GBD Compare Data Visualization*. IHME, University of Washington. Institute for Health Metrics and Evaluation (IHME)

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475. https://doi.org/10.2307/2951620

Jacobson Vann, J. C., Jacobson, R. M., Coyne-Beasley, T., Asafu-Adjei, J. K., & Szilagyi, P. G. (2018). Patient reminder and recall interventions to improve immunization rates. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD003941.pub3

Janz, N. K., & Becker, M. H. (1984). The Health Belief Model: A Decade Later. *Health Education Quarterly*, *11*(1), 1–47. https://doi.org/10.1177/109019818401100101

Kim, Y., & Radoias, V. (2016). Education, individual time preferences, and asymptomatic disease detection. *Social Science & Medicine*, *150*, 15–22. https://doi.org/10.1016/j.socscimed.2015.11.051

Kremer, M., Rao, G., & Schilbach, F. (2019). Chapter 5 - Behavioral development economics. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of Behavioral Economics: Applications and Foundations 1* (Vol. 2, pp. 345–458). North-Holland. https://doi.org/10.1016/bs.hesbe.2018.12.002

Mahendradhata, Y., Trisnantoro, L., Listyadewi, S., Soewondo, P., Marthias, T., Harimurti, P., & Prawira, J. (2017). *The Republic of Indonesia Health System Review* (Vol. 7). WHO.

Manne-Goehler, J., Geldsetzer, P., Agoudavi, K., Andall-Brereton, G., Aryal, K. K., Bicaba, B. W., Bovet, P., Brian, G., Dorobantu, M., Gathecha, G., Singh Gurung, M., Guwatudde, D., Msaidie, M., Houehanou, C., Houinato, D., Jorgensen, J. M. A., Kagaruki, G. B., Karki, K. B., Labadarios, D., … Jaacks, L. M. (2019). Health system performance for people with diabetes in 28 low- and middle-income countries: A cross-sectional study of nationally representative surveys. *PLOS Medicine*, *16*(3), e1002751. https://doi.org/10.1371/journal.pmed.1002751

Marcus, M. E., Reuter, A., Rogge, L., & Vollmer, S. (2020). *A Mobile Phone-based Intervention to Improve Health Screening Uptake: A Randomized Experiment in Indonesia*. AEA RCT Registry. https://doi.org/10.1257/rct.5047-2.0

Mekonnen, Z. A., Gelaye, K. A., Were, M. C., Gashu, K. D., & Tilahun, B. C. (2019). Effect of mobile text message reminders on routine childhood vaccination: a systematic review and meta-analysis. *Systematic Reviews*, *8*(1), 154. https://doi.org/10.1186/s13643-019-1054-0

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2013). Planning prompts as a means of increasing preventive screening rates. *Preventive Medicine*, *56*(1), 92–93. https://doi.org/10.1016/j.ypmed.2012.10.021

O'Donoghue, T., & Rabin, M. (2015). Present Bias: Lessons Learned and To Be Learned. *American Economic Review*, *105*(5), 273–279. https://doi.org/10.1257/aer.p20151085

Picone, G., Sloan, F., & Taylor, Jr., D. (2004). Effects of Risk and Time Preference and Expected Longevity on Demand for Medical Tests. *Journal of Risk and Uncertainty*, *28*(1), 39–53. https://doi.org/10.1023/B:RISK.0000009435.11390.23

Pujilestari, C. U., Ng, N., Hakimi, M., & Eriksson, M. (2014). "It is not possible for me to have diabetes"–Community Perceptions on Diabetes and Its Risk Factors in Rural Purworejo District, Central Java, Indonesia. *Global Journal of Health Science*, *6*(5), p204. https://doi.org/10.5539/gjhs.v6n5p204

Rahmawati, R., & Bajorek, B. (2018). Understanding untreated hypertension from patients' point of view: A qualitative study in rural Yogyakarta province, Indonesia. *Chronic Illness*, *14*(3), 228–240. https://doi.org/10.1177/1742395317718034

Riskesdas. (2018). *Laboran Nasional Riskesdas 2018*. Badan Penelitian dan Pengembangan Kesehatan; Ministry of Health.

Sallis, A., Sherlock, J., Bonus, A., Saei, A., Gold, N., Vlaev, I., & Chadborn, T. (2019). Pre-notification and reminder SMS text messages with behaviourally informed invitation letters to improve uptake of NHS Health Checks: a factorial randomised controlled trial. *BMC Public Health*, *19*(1), 1162. https://doi.org/10.1186/s12889-019-7476-8

Selden, T. M. (1993). Uncertainty and health care spending by the poor: The health capital model revisited. *Journal of Health Economics*, *12*(1), 109–115. https://doi.org/10.1016/0167-6296(93)90043-E

Taylor, D., Lunny, C., Lolić, P., Warje, O., Geldman, J., Wong, T., Gilbert, M., Lester, R., & Ogilvie, G. (2019). Effectiveness of text messaging interventions on prevention, detection, treatment, and knowledge outcomes for sexually transmitted infections (STIs)/HIV: a systematic review and meta-analysis. *Systematic Reviews*, *8*(1), 12. https://doi.org/10.1186/s13643-018-0921-4

Tsaneva, M. (2013). The Effect of Risk Preferences on Household Use of Water Treatment. *Journal of Development Studies*, *49*(10), 1427–1435. https://doi.org/10.1080/00220388.2013.790960

WHO. (2010). *WHO PEN Protocol 1 - Package of essential noncommunicable disease interventions for primary health care in low-resource settings*. https://www.who.int/ncds/management/Protocol1_HeartAttack_strokes_kidneyDisease.pdf?ua=1

7 Appendices

A. Wording of messages

Table A 1 Wording of messages

Message (English)	Message (Indonesian)	Sending date
Greetings [Mr/Ms] [name], do you know that diabetes does not always show symptoms but can be treated better if detected earlier. Check for FREE at POSBINDU [date]	Salam [Pak/Ibu] [name], tahukah Anda diabetes tdk selalu menunjukan gejala namun dapat diobati lbh baik jika diketahui lbh awal. Periksa GRATIS di POSBINDU [date]	5 days before the first village screening date
Greetings [Mr/Ms] [name], do you know that people over 40 years old have a high risk of diabetes & hypertension? Ask kader / PKM & check for FREE at POSBINDU [date]	Salam [Pak/Ibu] [name], tahukah Anda umur diatas 40 tahun memiliki risiko tinggi diabetes & darah tinggi? Tanyakan Kader/PKM & Periksa GRATIS di POSBINDU tgl [date]	3 days before the first village screening date
Greetings [Mr / Mrs] [name], remember to benefit from a FREE diabetes and hypertension CHECK in POSBINDU tomorrow morning at [place within the village]. Contact nearest kader or PKM.	Salam [Pak/Ibu] [name], Jangan Lupa untuk PERIKSA Darah Tinggi dan Diabetes GRATIS di POSBINDU Besok pagi di [place within village]. Hubungi Kader dan PKM terdekat	1 day before the first village screening date
Greetings [Mr/Ms] [name], remember that hypertension does not always show symptoms but can be treated if detected earlier. Check for FREE at POSBINDU [date]	Salam [Pak/Ibu] [name], ingatlah darah tinggi tdk selalu menunjukan gejala namun dapat diobati lbh baik jika diketahui lbh awal. Periksa GRATIS di POSBINDU [date]	5 days before the second village screening date
Greetings [Mr/Ms] [name], remember that people over 40 years old have a high risk of diabetes & hypertension. Ask Cadre / PKM & check for FREE at POSBINDU date [date]	Salam [Pak/Ibu] [name], ingatlah umur diatas 40 tahun memiliki risiko tinggi diabetes & darah tinggi. Tanyakan Kader/PKM & Periksa GRATIS di POSBINDU tgl [date]	3 days before the second village screening date
Greetings [Mr / Mrs] [name], remember to benefit from a FREE diabetes and hypertension CHECK in POSBINDU morning at [place within the village]. Contact nearest kader or PKM.	Salam [Pak/Ibu] [name], Jangan Lupa untuk PERIKSA Darah Tinggi dan Diabetes GRATIS di POSBINDU Besok pagi di [place within village]. Hubungi Kader dan PKM terdekat	1 day before the second village screening date

B. Data collection details

Table A 2 Data collection timeline

	2019			2020			
Month	**October**	**November**	**December**	**January**	**February**	**March**	**April**
Qualitative pre-studies							
Baseline data collection (enrolment)							
Treatment allocation				X			
Pilot Intervention				X			
Intervention							
Endline data collection							

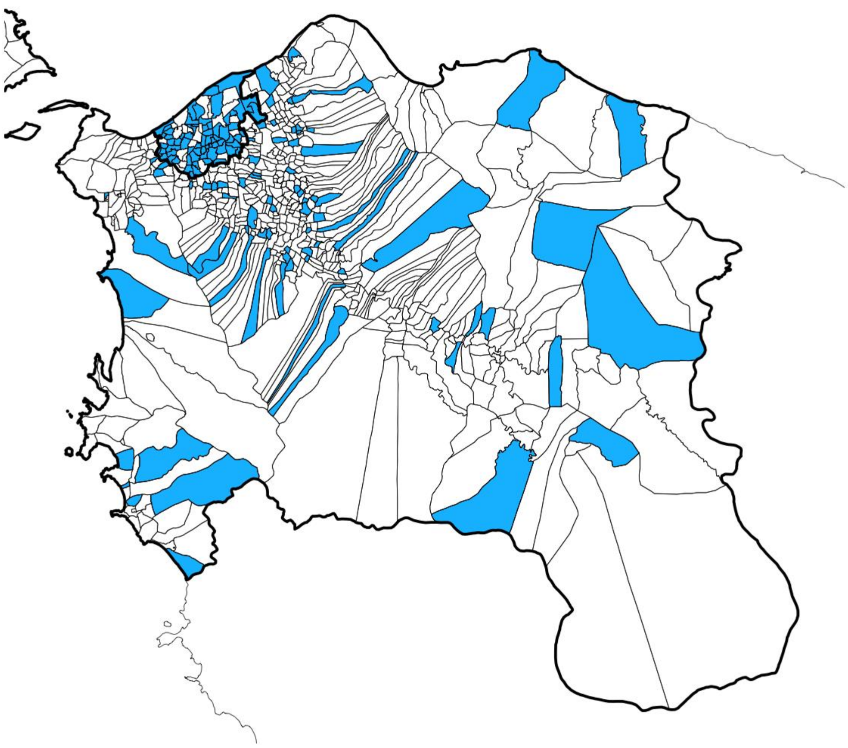


Figure A 1 Sample villages. Boundaries of the city Banda Aceh and the district Aceh Besar are in bold. Taken from the supplementary material in Chavarría et al. (2021).

B1 *Inclusion Criteria*

We targeted the population at high risk for NCDs, who do not yet adhere to the recommended screening schedule. Based on this, we formulated six inclusion and exclusion criteria:

1. The respondent must be between 40 and 70 years old. The WHO PEN Protocol for essential NCD interventions for primary health care in low-resource settings specifies that individuals over 40 years old should undergo routine screening for hypertension and diabetes (WHO, 2010).
2. The respondent cannot already be diagnosed with diabetes or hypertension, as this would render screening unnecessary.
3. The respondent did not undergo diabetes screening within the last year. Individuals that have done so seem to be adhering to recommended screening schedules, and would therefore not fall within our target population. Hypertension screening is not included in this restriction, as blood pressure checks are usually carried out whenever individuals visit a community health center and are hence much more common in this context.
4. The respondent must not be in regular care for another disease. If they are in regular contact with health system services, a lack of NCD screening may not stem from a lack of demand but rather from further downstream health system failures, which we do not aim to address in our intervention.
5. The respondent must be reachable via phone and text messages on either their own or another household member's phone.
6. The respondent must be at home at the time of the interview. Logistically, it was not feasible to re-visit households. Furthermore, seeking out respondents outside of their home would have violated the comparability of interview conditions across our sample. For instance, respondents might feel most comfortable answering sensitive questions regarding their health in their own home. This criterion might bear the risk to exclude the working population, which we sought to reduce by extending the enumeration time to the evening and the weekends. Overall, this might not be as severe in our age group as in younger age groups, as some are retired already or work from home.

B2 *Random walk scheme*

Taken from the supplementary material in Chavarría et al. (2021).

The enumerators conducted the random walk according to the following instructions to ensure that the walk yields a representative sample of the target population:

1. Get permission and number of village subdivisions from the village head.
2. Ask for a description of the village boundaries, including remote houses.
3. Get the total number of houses in the village and divide this number by 100. This number indicates the skip-pattern of houses. It takes into account the aim of having around 20 respondents per village that should be evenly distributed throughout the

village, how many interviews one enumerator can do in one day, and the likelihood of finding a household member that meets the inclusion criteria.

4. Then, randomly select which village subdivision to visit first and at which house (a random number between 1 and the skip number) to begin with. The count begins from the point of entry to the respective subdivision.
5. If a person is at home, check and record the eligibility and conduct the interview if the criteria are fulfilled and the respondent is willing to.
6. After each contact, continue with the next house according to the skip pattern.
7. In case of an empty house, contact the direct neighbor until an occupied house was found and record the number of empty houses.
8. When walking, turn left on every turn and only count houses to your left. Whenever you reach the end of the village subdivision or the road, turn around and continue.
9. One village was considered finished if 20 interviews were conducted or all houses that should be contacted according to the skip pattern were contacted.

Table A 3 Overview of baseline contacts

	Total	**Of all contacts**			**Of all consenting**		**Of all eligible**		
	Contacts	Empty houses	Refusal/ busy/ other	Consent	Eligible	Ineligible	Refusal	Incomplete	Complete
N	15,128	7,682	946	6,500	2,115	4,385	11	98	2,006

	Of all ineligible			
	No member 40-70	No member 40-70 present	No phone access	No member without diagnosis/ screening/care
N	1,589	414	270	2,112

Disaggregation of the number of contacts and respondents at baseline. Contacts refer to all dwelling units drawn by the random walk within the villages. Empty houses are dwellings where no one was present at the first contact, including dwellings which might not be inhabited. Refusal/busy/other denotes to reasons for non-participation stated at the first contact. Consent signifies that at least one household member agreed to respond to the screening questions to assess eligibility. Eligible refers to all contacts where at least one eligible member was present. Ineligible are all contacts where no member was eligible or no eligible member was present. Refusal denotes those (eligible) contacts for which no eligible member was willing to participate in the study. Incomplete denotes the interviews which were missing information on the telephone number. Complete refers to all conducted interviews with information on the telephone number. The columns 'no member 40-70' till 'no phone access' refer to the household eligibility criteria, the last column to the individual-level criteria (if multiple members were eligible, one was randomly selected). Among individuals, ineligibility could occur due to previous hypertension or diabetes diagnosis (59.36%), being in continued care (8.42%), being tested for diabetes in the last year (31.98%), or not answering one of the eligibility questions (0.24%). Taken from the supplementary material in Chavarría et al. (2021).

B3 <u>Power Calculations</u>

The following procedure of power calculation was set in the pre-analysis plan and under the assumption of an in-person endline data collection, which we had to deviate from due to the start of the COVID-19 pandemic.

The sample size was determined based on sufficient statistical power to determine a meaningful change in the primary outcome, screening uptake. Prior to baseline data collection, we could approximate the base levels of diabetes and hypertension separately from the most recent round of the Indonesian health survey Riskesdas (Riskesdas, 2018). This data supplies self-reported figures on whether the individual respondent attends screening regularly, irregularly or never, where regularly is defined as according to the doctor's advice for patients and once a year for the non-diagnosed. As our outcome is measured during approximately two months, the most appropriate base value is the *regular* category. The national average of the age group between 45 and 74 years is 5.2% for diabetes and 16.7% for hypertension screening[12]. As there are no previous studies on the effect of text message reminders on diabetes and hypertension screening, the minimum detectable effect size was approximated from studies that measure the effect of text message reminders on the initial take-up of other health services. A review on vaccination uptake found an average effect size of 4.5 percentage points (Jacobson Vann et al., 2018). With a power of 80% and 5% significance, a sample size of 1,800 individuals would be required to detect such an effect for both diabetes and hypertension screening. We would be able to detect a 4.4 percentage point increase for blood pressure measurement and a 2.6 percentage point increase in blood glucose measurement.[13] This implies that we would be able to detect a significant effect on any screening if at least 24 more respondents of the treatment group attend diabetes screening during the intervention period compared to the control group at the same time. With this sample size, we will also be likely to detect a small change in the secondary knowledge outcomes. For the SMS knowledge, the mean points of the treatment group need to be 0.1 points higher than for the control group, which means that on average every tenth respondent needs to know one item more. For the broader health knowledge index, we will be able to detect a 0.56 point difference, which means that on average about every other individual in the treatment group needs to know at least one item more than the control group. As these changes are smaller than a meaningful effect that we would expect to be a channel for the primary outcome, we expect to be able to detect every meaningful effect of the intervention on health knowledge.

We account for potential sample reductions by over-sampling by about 15%. The main reason for a high over-sampling rate is that we rely on functioning phone numbers for the intervention. The over-sampling also accounts for respondents that need to be excluded from the treatment group because the messages could not be delivered to their mobile phone. One reason might be that the respondent changed his/her telephone number, which is common in this context. We tried to avoid this by asking for a contact number that is likely to be active until April 2020, and by planning a short duration between baseline interview and intervention. Another reason might be a typo when entering the phone number. Non-compliance might be a problem if the respondent does not own a mobile phone and the

---

[12] From our baseline data, we know that slightly more individuals (23%) had a blood pressure check during the previous year. This would increase the minimal detectable effect size by 0.5 percentage points.

[13] We used the 3ie Sample size and minimum detectable effect calculator as described in Djimeu and Houndolo (2016). For screening uptake, we used the formula for binary outcomes and for the knowledge index the formula for continuous outcomes.

stated contact person does not transfer the message. We minimize this by specifically asking for a contact person from whom a message can be received and by including the name of the recipient in each message. Finally, we expect attrition at endline as it is likely that some respondents either cannot be found or are unavailable or unwilling to participate in a second interview. However, we expect overall attrition to be low: at baseline, each respondent has agreed to a second interview, we have taken detailed information on the place of residence (name, address, and geolocation), and we can contact him/her through the mobile phone number.

Calling procedure at endline

Taken from the supplementary material in Chavarría et al. (2021).

The telephone interviews were scheduled according to the call pattern that is displayed below. Initially, each respondent received five calls, which were staggered with time delays of one hour to three days any at varying times of the day. After the second unanswered call, a standardized text message was sent announcing another call on the following day. Whenever feasible, the same enumerator who had visited the respondent during the baseline survey was deployed to call them during the phone interview, in order to maximize the response rate as well as the respondents' trust towards the enumerator. In the end of the data collection period, each number that was not answered during five calls received one additional call from another interviewer (with a different telephone number).

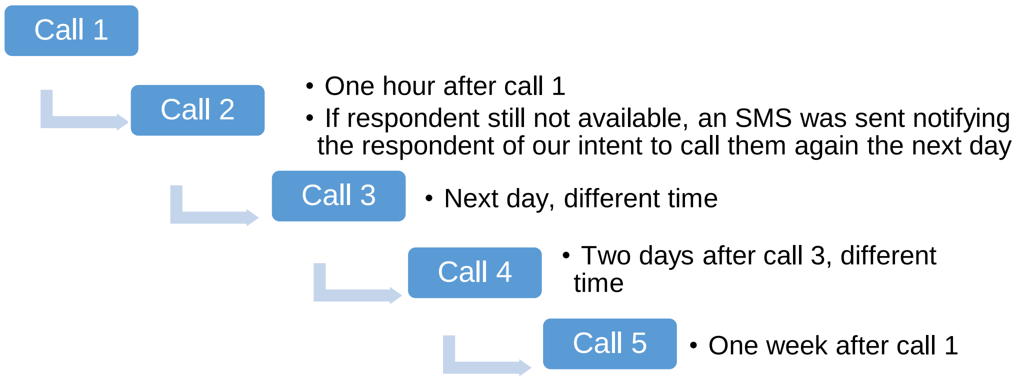


Figure A 2 Call Pattern at endline

C. Variable definitions

C1 Knowledge Indices

Table A 4 Composition SMS knowledge index

Question	***Coding***
"One can feel whether one experiences diabetes/ hypertension "	0 if (strongly) agree, 1 if (strongly) disagree
"It makes a difference to start diabetes/ hypertension treatment early"	0 if (strongly) disagree, 1 if (strongly) agree
Which risk factors of diabetes/ hypertension do you know?	1 if mentioned age, 0 otherwise
Have you ever heard of *Posbindu*?	0 if (strongly) disagree, 1 if (strongly) agree

Note: Each question with diabetes / hypertension is included for both diseases separately. "Don't know" coded as 0.

Table A 5 Composition knowledge index

Question	***Coding***
"Which risk factors of diabetes / hypertension do you know?"	1 count for each correctly identified factor
Do you know someone with diabetes/ hypertension?	Binary variable for the answers: Family member, friend, neighbour, other, none.
Which complications of disease diabetes/ hypertension do you know?	1 count for each correctly identified factor
"Who do you think should be screened?"	0 if "everyone who feels sick", 1 if "everyone" or "people at risk"
Which ways of controlling diabetes/ hypertension do you know?	1 count for each correctly identified factor
"It makes a difference to start treatment early"	0 if (strongly) disagree, 1 if (strongly) agree
"There is nothing one can do to prevent diabetes/ hypertension, it is destiny."	0 if (strongly) agree, 1 if (strongly) disagree
"One can feel whether you experience diabetes/ hypertension "	0 if (strongly) agree, 1 if (strongly) disagree
"Checking your level regularly helps to detect diabetes/ hypertension early"	0 if (strongly) disagree, 1 if (strongly) agree
"Diabetes/ hypertension is treatable"	0 if (strongly) disagree, 1 if (strongly) agree

Note: Each question with diabetes / hypertension is included for both diseases separately. "Don't know" coded as 0.

D. Intervention piloting

We piloted the messages in January 2020 to find out whether the contents were understandable, deemed trustworthy, and to assess whether the time of sending (morning/evening) and order of information (age as risk factor/having it without feeling it) mattered. However, the messages were not sent according to the time schedule of the intervention, i.e., not 5, 3 and 1 day before a *Posbindu* date. The messages 1 and 2 were sent to the respondents on two consecutive days, and respondents were interviewed via phone a few days after. In 10 out of 14 cases, the phone was answered on the designated survey day (no second contact attempts on another day were made). The messages were received in 9 out of 10 cases, although in two cases they were received by the children of the main respondent and were not yet transferred to him/her. In both cases, the *Posbindu* dates were a few weeks ahead, so the children might not have felt the urgency to deliver the message directly. We assumed that this would be different when the dates are close by.

Qualitative semi-structured interviews were conducted with the remaining eight respondents. All respondents confirmed that they trusted the message. Reasons stated were the connection to the interview conducted two months before, the mentioning of a public program (*Posbindu*) and the *kaders*, the mentioning of the respondent's name, and confirmation of the content by the *kader*. Most respondents remembered that the messages were reminding them to go to *Posbindu*, and some specifically mentioned the *Posbindu* date. Three respondents could recall that the messages contained information regarding diseases, and two additional respondents recalled information regarding risk factors. The respondents liked in particular that the messages served as reminders, and two respondents explicitly stated that they liked how the messages were written. Time of message sending and order of the messages did not appear to make a difference in how the messages were perceived.

While experimenter demand biases are always a concern in these types of interviews, we believe them to be minimal here. First of all, respondents may feel less inclined to cater to experimenter demand during phone interviews, as they are less personal than in-home visits. This was confirmed by our enumerators, who qualitatively assessed that respondents were likely to report their true opinions. Second of all, respondents always gave specific reasons and arguments for their opinions, making them more credible.

E. Sample characteristics and attrition

Table A 6 Baseline balance across treatment and control group

	Control group			Treatment group			
	Mean	Standard deviation	N	Mean	Standard deviation	N	p-value
Age	50.35	8.24	1,002	49.91	8.08	1,003	0.226
Female	0.64	0.48	1,001	0.64	0.48	1,003	0.936
Highest level of schooling							0.876
None	0.05	0.22	49	0.05	0.22	49	
Primary	0.25	0.43	253	0.24	0.42	236	
Junior Secondary	0.21	0.41	215	0.22	0.41	219	
Senior Secondary	0.35	0.48	346	0.35	0.48	348	
Tertiary	0.14	0.35	139	0.15	0.36	152	
Wealth quintile							0.611
1	0.22	0.42	225	0.21	0.41	213	
2	0.20	0.40	203	0.18	0.39	182	
3	0.19	0.39	192	0.20	0.40	200	
4	0.19	0.39	188	0.20	0.40	198	
5	0.19	0.39	193	0.21	0.41	211	
Own phone	0.58	0.49	995	0.62	0.49	1,000	0.044
Posbindu in own village	0.90	0.30	1,002	0.90	0.30	1,004	0.666
Ever had blood pressure or blood glucose checked	0.58	0.49	999	0.59	0.49	1,002	0.610
Disease knowledge index	18.30	5.53	923	17.97	5.42	936	0.196
Patience	5.73	2.83	1,002	5.70	2.86	1,004	0.823
Willingness to take risks	4.57	2.66	1,002	4.45	2.62	1,004	0.298
Joint F-test				0.880			

Means, standard deviation and number of observations of main respondent characteristics by treatment group; p-values based on t-tests of difference in mean between treatment and control group, except in the case of education and wealth quintile, where we used Pearson chi-squared tests due to the categorical nature of the variables.

Table A 7 Comparison of sample characteristics to SUSENAS

	SUSENAS Banda Aceh, Aceh Besar	Baseline	Endline
Age	50.5935	50.1203	49.9404
	(0.3088)	(0.1826)	(0.2306)
Above 50	0.4878	0.4656	0.4592
	(0.0207)	(0.0111)	(0.0142)
Female	0.5239	0.6379***	0.6224**
	(0.0207)	(0.0107)	(0.0161)
Education			
- Up to primary	0.2424	0.2926**	0.2720***
	(0.0188)	(0.0100)	(0.0162)
- Lower secondary	0.2347	0.2164	0.2188
	(0.0179)	(0.0092)	(0.0120)
- Upper secondary and above	0.5229	0.4910	0.5092**
	(0.0207)	(0.0109)	(0.0194)
Wealth above median		0.4923	0.5082**
		(0.0112)	(0.0201)
Banda Aceh	0.4074	0.4372	0.4511*
	(0.0182)	(0.0061)	(0.0220)
N	863	2,006	1,412

*SUSENAS samples are obtained from SUSENAS 2017 and restricted to respondents aged 40 – 70 with a mobile phone in the household. Standard errors accounting for survey design (sampling weights in SUSENAS, district stratification in both samples, PSU when comparing base- and endline sample) below mean; stars indicate significant difference from mean listed in previous column based on adjusted Wald test, * 0.1 ** 0.05 *** 0.01. Columns on SUSENAS and Baseline as in (Chavarría et al., 2021).*

Attrition

We test for differential attrition using three approaches. First, we test whether attrition differs across treatment and control group:

$$Attrit_i = \alpha + \beta T_i + \omega_{ij} \qquad \text{(A1)}$$

Second, we analyze attrition based on the set of baseline characteristics used for testing balance across treatment and control group – namely age, sex, education, wealth quintile, knowledge index, time preferences, risk preferences, phone ownership and *Posbindu* in own village:

$$y_i = \alpha + \beta Attrit_i + \omega_{ij} \qquad \text{(A2)}$$

Third, we examine whether these baseline characteristics of attrited treated individuals are significantly different from the attrited control individuals, restricting the sample to attriting respondents only:

$$(y_i | Attrit = 1) = \alpha + \beta T_i + \omega_{ij} \qquad \text{(A3)}$$

Table A 8 Attrition I: between treatment and control group

	(1) Attrition
Treated	0.0273
	(0.0207)
Observations	2006

*Regression of a binary attrition indicator (not re-interviewed at endline) on a binary treatment indicator (equation A1). Standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table A 9 Attrition II: endline sample compared to those lost to follow-up

	(1) Age	(2) Female	(3) Education	(4) Wealth quintile	(5) Baseline disease knowledge	(6) Willingness to take risks	(7) Patience	(8) Own phone	(9) Own Posbindu
Attrition	0.630	0.055**	-0.218***	-0.182**	-2.465***	-0.057	-0.111	-0.200***	0.008
	(0.406)	(0.023)	(0.056)	(0.071)	(0.304)	(0.129)	(0.138)	(0.024)	(0.015)
Observations	2005	2004	2006	2005	1580	2006	2006	1995	2006

*Separate regressions of each characteristic on the binary attrition indicator (not re-interviewed at endline) (equation A2). Standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table A 10 Attrition III: between treatment and control in those lost to follow-up

	(1) Age	(2) Female	(3) Education	(4) Wealth quintile	(5) Baseline disease knowledge	(6) Willingness to take risks	(7) Patience	(8) Own phone	(9) Own *Posbindu*
Treated	0.149	0.060	0.047	0.042	-0.849*	-0.236	-0.246	0.065	0.029
	(0.688)	(0.038)	(0.096)	(0.119)	(0.487)	(0.218)	(0.230)	(0.041)	(0.024)
Observations	594	593	594	594	532	594	594	590	594

*Separate regressions of each characteristic on the binary treatment indicator in the sample that was not re-interviewed at endline (equation A3). Standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table A 11. Role of phone ownership for attrition

	(1) Own phone	(2) Attrition
Age	-0.008***	-0.000
	(0.001)	(0.001)
Female	-0.113***	0.032
	(0.021)	(0.021)
Primary	0.088*	-0.142***
	(0.050)	(0.054)
Junior Secondary	0.156***	-0.155***
	(0.053)	(0.056)
Senior Secondary	0.360***	-0.121**
	(0.051)	(0.055)
Higher	0.517***	-0.146**
	(0.053)	(0.060)
Wealth quintile 2	0.011	0.001
	(0.032)	(0.033)
Wealth quintile 3	0.043	-0.048
	(0.033)	(0.031)
Wealth quintile 4	0.042	-0.012
	(0.033)	(0.033)
Wealth quintile 5	0.079**	-0.028
	(0.034)	(0.033)
Own phone		-0.161***
		(0.023)
Observations	1991	1991

*Regression of the binary phone ownership indicator (column 1) and the binary attrition indicator (column 2) on the respective characteristics in the whole intervention sample. Reference categories: No formal education, wealth quintile 1; Coefficient estimates for education in column (2) are statistically not distinguishable from each other. Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.*

F. Main tables and robustness checks

Table A 12 Treatment effects on screening uptake, with and without covariates

	(1) ITT	(2) ITT	(3) LATE	(4) LATE	(5) Any other member	(6) Any other member
Treated	0.0576**	0.0656***	0.144	0.172*	0.0152	0.0106
	(0.0257)	(0.0254)	(0.0970)	(0.0969)	(0.0250)	(0.0250)
Covariates	No	Yes	No	Yes	No	Yes
Observations	1386	1386	1175	1175	1070	1070
Control group mean	0.331	0.331	0.357	0.357	0.205	0.205

Results of regressing the binary screening uptake indicator following equation 1 for the message recipient (columns 1 and 2) and any other household member (columns 5, 6) and the local average treatment effect following equation 3 (columns 3, 4); if covariates are included, they are message recipient age, gender, wealth and phone ownership; standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 13. Adjustments for multiple hypothesis testing in main specification for primary and secondary outcomes.

	(1) Screening uptake (ITT)	(2) Screening uptake (LATE)	(3) Spillovers	(4) SMS Knowledge	(5) General Knowledge
Treated	0.066	0.172	0.011	-0.002	-0.336
	(0.010)***	(0.076)*	(0.672)	(0.962)	(0.340)
	[0.090]*	[0.227]	[0.808]	[0.962]	[0.510]
Covariates	Yes	Yes	Yes	Yes	Yes
Observations	1386	1175	1070	1088	1042

Results of regressing the binary screening uptake indicator following equation 1 for the message recipient (col 1) and any other household member (col 3), the respective knowledge index (col 4, 5), and the local average treatment effect following equation 3 (col 2); controlling for message recipient age, gender, wealth, and phone ownership; unadjusted p-values in parentheses, adjusted q-values following the Benjamini-Hochberg method for the 9 main hypotheses in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 14. Adjustments for multiple hypothesis testing in main specification of heterogeneity analysis.

	Screening uptake	Screening uptake
Willingness to take risk	0.082	
	(0.105)	
	[0.236]	
Patience		0.118
		(0.037)**
		[0.165]
Treated x Willingness to take risk	-0.004	
	(0.719)	
	[0.808]	
Treated x Patience		-0.009
		(0.301)
		[0.541]
Covariates	Yes	Yes
Observations	1386	1386

*Treatment coefficients from estimating equation 4 controlling for message recipient age, gender, wealth, and phone ownership; unadjusted p-values in parentheses, adjusted q-values following the Benjamini-Hochberg method for the 9 main hypotheses in square brackets. * p < 0.1, ** p < 0.05, *** p < 0.01*

Table A 15. Binary outcomes with probit and logit specifications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Screening uptake		Heterogeneity: Risk		Heterogeneity: Time		Spillover	
	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit
Treated	0.182***	0.301***	0.229	0.375	0.332**	0.546**	0.033	0.063
	(0.070)	(0.116)	(0.141)	(0.231)	(0.158)	(0.260)	(0.088)	(0.153)
Preference			0.019	0.031	0.022	0.036		
			(0.019)	(0.032)	(0.018)	(0.029)		
Treated x Preference			-0.010	-0.016	-0.026	-0.043		
			(0.027)	(0.044)	(0.025)	(0.040)		
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1386	1386	1386	1386	1386	1386	1065	1065

*Results of regressing the binary screening uptake indicator following equation 1 for the message recipient (col 1, 2) and any other household member (col 7, 8), as well as heterogeneous treatment effects along a continuous risk and time preference scale following equation 4; controlling for message recipient age, gender, wealth and phone ownership; each model is separately estimated using probit and logit; standard errors clustered at the phone-number level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01*

Table A 16. Knowledge outcomes measured through PCA

	SMS knowledge (PCA)	SMS knowledge (PCA)	Disease knowledge (PCA)	Disease knowledge (PCA)
Treated	0.0215	0.00198	-0.0328	-0.0551
	(0.0596)	(0.0581)	(0.0612)	(0.0594)
Covariates	No	Yes	No	Yes
Obs.	1088	1088	1042	1042
Control group mean	-0.00301	-0.00301	0.0215	0.0215

Regressions for an alternative definition of both knowledge indices via Principal Component Analysis; if covariates are included, they are message recipient age, gender, wealth, and phone ownership; standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 17 Treatment effect on each element of the SMS knowledge index

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Feel it		Early treatment		Age risk		Knows
	Hypertension	Diabetes	Hypertension	Diabetes	Hypertension	Diabetes	Posbindu
Treated	0.0051	-0.0133	0.0040	-0.0033	-0.0171	0.0178	0.0047
	(0.0089)	(0.0156)	(0.0109)	(0.0129)	(0.0173)	(0.0163)	(0.0171)
Covar.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1088	1088	1088	1088	1088	1088	1088
C. mean	0.0185	0.0775	0.9613	0.9502	0.1015	0.0664	0.9151

Regressions of the components of the SMS knowledge index as defined in Table A 4 on the binary treatment indicator controlling for message recipient age, gender, wealth, and phone ownership; standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 18 Treatment effect on each element of the disease knowledge index (Hypertension)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Number of					Share with correct answer			
	Risk Factors	Compli-cations	Control	Target group	Start early	Destiny	Feel it	Regular checks	Treat-able	Know someone
Treated	-0.0627	0.0311	-0.0959	-0.0044	0.0026	0.0010	0.0072	-0.0134	-0.0022	0.0014
	(0.0680)	(0.0439)	(0.0705)	(0.0306)	(0.0106)	(0.0283)	(0.0140)	(0.0101)	(0.0189)	(0.0251)
Covar.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1042	1042	1042	1042	1042	1042	1042	1042	1042	1042
C. mean	2.1612	1.1478	2.1440	0.5566	0.9655	0.2917	0.9424	0.9789	0.8983	0.7908

Regressions of the components of the disease knowledge index as defined in Table A 5 on the binary treatment indicator controlling for message recipient age, gender, wealth, and phone ownership; the outcomes in columns 1-3 are the number of correct items and binary measures in columns 4-10; standard errors clustered at the phone-number level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A 19 Treatment effect on each element of the general knowledge index (Diabetes)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Number of				Share with correct answer				
	Risk Factors	Compli-cations	Control	Target group	Start early	Destiny	Feel it	Regular checks	Treat–able	Know someone
Treated	-0.0623	-0.1026	-0.0722	0.0138	-0.0047	0.0072	0.0258	0.0061	0.0172	0.0321
	(0.0607)	(0.0706)	(0.0628)	(0.0307)	(0.0125)	(0.0278)	(0.0226)	(0.0105)	(0.0268)	(0.0297)
Covar.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1042	1042	1042	1042	1042	1042	1042	1042	1042	1042
C. mean	1.8330	1.6046	1.7697	0.5182	0.9559	0.2726	0.8292	0.9655	0.7486	0.6180

Regressions of the components of the disease knowledge index as defined in Table A 5 on the binary treatment indicator controlling for message recipient age, gender, wealth, and phone ownership; the outcomes in columns 1-3 are the number of correct items and binary measures in columns 4-10; standard errors clustered at the phone-number level in parentheses. * *$p < 0.1$,* ** *$p < 0.05$,* *** *$p < 0.01$.*

Table A 20 Different versions of spillover analysis

	Any member (main specification)	Member 40-70	Other phone owner
Treated	0.0106	0.0134	0.0167
	(0.0250)	(0.0308)	(0.0305)
Other's phone			0.0399
			(0.0392)
			-0.0180
Treated x other's phone			(0.0530)
			0.0399
Covariates	Yes	Yes	Yes
Obs.	1070	727	1070
Mean	0.205	0.212	0.205

Results of regressing the binary indicator of household member screening uptake (col 1), screening uptake among other household members aged 40-70 years (col 2) on the binary treatment indicator following equation 1, and the heterogeneous treatment effect of the binary phone ownership indicator, which takes value 1 if the intervention was either received on a family phone or the private phone of another household member, and zero if it belongs to the message recipient; controlling for age, gender, wealth and phone ownership; standard errors clustered at the phone-number level in parentheses. * *$p < 0.1$,* ** *$p < 0.05$,* *** *$p < 0.01$.*

Table A 21 Treatment effect on screening uptake by month

	(1) January	(2) February	(3) March	(4) April
Treated	0.0156	0.0363	0.0560***	0.0068
	(0.0159)	(0.0228)	(0.0201)	(0.0090)
Covariates	Yes	Yes	Yes	Yes
Obs.	1386	1386	1386	1386
Control group mean	0.0895	0.2216	0.1435	0.0256

*Results of regressing different binary screening uptake indicators on the binary treatment indicator (equation 1), controlling for age, gender, wealth and phone ownership; the outcome indicator takes the value 1 only if the individual indicated to have gone to screening in the respective month and zero otherwise; standard errors clustered at the phone-number level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A 22 Treatment effect on screening uptake by location

	(1) Went on correct date to Posbindu	(2) Posbindu	(3) Puskesmas	(4) Private doctor/midwife
Treated	0.0067	0.0081	0.0298*	0.0201
	(0.0177)	(0.0178)	(0.0158)	(0.0162)
Covariates	Yes	Yes	Yes	Yes
Obs.	1386	1386	1386	1386
Control group mean	0.1335	0.1335	0.0810	0.0895

*Results of regressing different binary screening uptake indicators on the binary treatment indicator (equation 1), controlling for age, gender, wealth and phone ownership; the outcome indicator takes the value 1 only if the individual indicated to have gone to screening in the respective facility and zero otherwise; the screening outcome in col 1 additionally conditions on the correct month; standard errors clustered at the phone-number level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A 23 Treatment effect on disaggregated screening outcome: kind of check done

	(1) Medical history	(2) Physical measurement	(3) Blood pressure	(4) Blood glucose	(5) Other blood check
Treated	0.0420**	0.0151	0.0652**	0.0302	0.0091
	(0.0176)	(0.0165)	(0.0254)	(0.0200)	(0.0134)
Covariates	Yes	Yes	Yes	Yes	Yes
Obs.	1386	1386	1386	1386	1386
Mean	0.1023	0.1009	0.3295	0.1548	0.0639

*Results of regressing different binary screening indicators on the binary treatment indicator (equation 1), controlling for age, gender, wealth and phone ownership; the outcome indicator takes the value 1 only if the individual indicated that at the screening visit the respective check was conducted and zero if the respondent either did not go for screening or did not get the respective check done despite going for screening; standard errors clustered at the phone-number level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A 24 Characteristics of sub-groups of treatment group who remember receiving messages on NCDs and specific elements of these messages

				Remembers content on:			
	Total treatment	Received message	LATE definition	Screening need	Posbindu logistics	Posbindu free	Age risk
Demographics							
Age	49.52	48.31***	48.54	47.79	48.36	48.42	49.60*
	(7.85)	(7.55)	(7.43)	(7.31)	(6.76)	(7.54)	(8.01)
Female	0.61	0.56*	0.57	0.55	0.60	0.55	0.56
	(0.49)	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)
Education							
- None							
	0.03	0.02	0.02	0.01	0.00	0.00	0.02
- Primary	(0.18)	(0.12)	(0.13)	(0.11)	(0.00)	(0.00)	(0.13)
	0.24	0.18**	0.17	0.14	0.19	0.19	0.18
- Lower	(0.42)	(0.39)	(0.38)	(0.35)	(0.40)	(0.39)	(0.39)
Secondary	0.21	0.18	0.17	0.21	0.21	0.17	0.11
- Higher	(0.41)	(0.39)	(0.38)	(0.41)	(0.41)	(0.38)	(0.31)
Secondary	0.36	0.43***	0.43	0.41	0.42	0.45	0.38
- Tertiary	(0.48)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)
	0.17	0.20	0.21	0.23	0.18	0.19	0.31**
Banda Aceh	(0.37)	(0.40)	(0.41)	(0.42)	(0.39)	(0.39)	(0.47)
	0.52	0.50	0.49	0.49	0.44	0.31***	0.51
SMS-related characteristics							
Phone owner	0.68	0.80***	0.80	0.79	0.81	0.77	0.80
	(0.47)	(0.40)	(0.40)	(0.41)	(0.40)	(0.43)	(0.40)
Messages							
- daily	0.48	0.57***	0.58	0.67**	0.58	0.60	0.61
	(0.50)	(0.50)	(0.50)	(0.47)	(0.50)	(0.49)	(0.49)
- < daily	0.36	0.39	0.38	0.29**	0.36	0.38	0.39
	(0.48)	(0.49)	(0.49)	(0.46)	(0.48)	(0.49)	(0.49)
- never	0.16	0.04***	0.04	0.04	0.06	0.02	0.00*
	(0.37)	(0.19)	(0.20)	(0.20)	(0.24)	(0.13)	(0.00)
Messenger use	0.47	0.48	0.49	0.61***	0.55	0.56	0.52
	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	(0.51)
Prefers less SMS							
- in general	0.15	0.22***	0.23	0.23	0.29*	0.14**	0.24
	(0.36)	(0.42)	(0.42)	(0.42)	(0.46)	(0.35)	(0.43)
- advertisement	0.60	0.57	0.57	0.61	0.54	0.66*	0.53
	(0.49)	(0.50)	(0.50)	(0.49)	(0.50)	(0.48)	(0.50)
- no	0.25	0.21*	0.20	0.16	0.17	0.21	0.22
	(0.44)	(0.41)	(0.40)	(0.37)	(0.38)	(0.41)	(0.42)
Baseline characteristics							
Disease	18.42	19.58***	19.76	20.07	19.10	19.87	20.00
knowledge	(5.30)	(4.88)	(4.99)	(5.18)	(4.42)	(4.99)	(4.44)
H- feel it	0.12	0.10	0.11	0.10	0.07	0.06	0.09
	(0.33)	(0.30)	(0.31)	(0.31)	(0.26)	(0.24)	(0.29)
D- feel it	0.19	0.19	0.20	0.23	0.13	0.16	0.18
	(0.39)	(0.39)	(0.40)	(0.42)	(0.34)	(0.37)	(0.39)
H- start early	0.95	0.96	0.95	0.93*	1.00**	0.95	0.98
	(0.22)	(0.20)	(0.21)	(0.25)	(0.00)	(0.21)	(0.13)
D- start early	0.94	0.94	0.94	0.92	0.99**	0.94	0.96
	(0.24)	(0.23)	(0.24)	(0.28)	(0.12)	(0.25)	(0.19)

H- age risk	0.06	0.05	0.05	0.07	0.03	0.05	0.02
	(0.23)	(0.22)	(0.21)	(0.25)	(0.17)	(0.21)	(0.13)
D- age risk	0.04	0.05	0.06	0.09**	0.06	0.06	0.04
	(0.20)	(0.22)	(0.24)	(0.29)	(0.23)	(0.25)	(0.19)
Knows Posbindu	0.50	0.56*	0.56	0.53	0.53	0.62	0.64
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.49)
Ever screened	0.59	0.61	0.57**	0.57	0.65	0.64	0.64
	(0.49)	(0.49)	(0.50)	(0.50)	(0.48)	(0.48)	(0.49)
Last year screened	0.29	0.28	0.25*	0.06***	0.15***	0.22	0.37
	(0.45)	(0.45)	(0.43)	(0.24)	(0.36)	(0.42)	(0.49)
N	682	199	170	87	72	65	55

*Simple means of the respective characteristic across groups: complete treatment group, individuals who stated to have received a message on Posbindu, those who received at least one full message cycle according to the delivery reports and remember any message content (LATE definition) and the four most commonly recalled content elements: the recommendation to take up screening, when and where Posbindu takes place, that Posbindu is free and higher age implies a higher NCD risk. Standard deviations in parentheses below mean; stars indicate the p-value of the two-sample t-test for difference of the respective group and characteristic compared to the rest of the treatment group; * p < 0.1, ** p < 0.05, *** p < 0.01.*

G. Administrative information

Funding

This work was supported by the German Research Foundation’s Research and Training Group 1723 “Globalization and Development” (DFG: RTG1723).

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 issued by the ethics board of the nursing faculty of Syiah Kuala University, Banda Aceh, Indonesia

Declaration of Interest

None.

Data availability

The data that support the findings of this study are available upon reasonable request in Göttingen Research Online at https://data.goettingen-research-online.de/, DOI 10.25625/SE4IDP, and will be made publicly available after publication of this study. The questionnaires are publicly available in the repository.

Acknowledgements

We thank Farah Diba and Marthoenis from Syiah Kuala University, who assisted in logistical support and insights on the Indonesian health care system. We further thank the members of the Chair of Sebastian Vollmer, the 3rd cohort of the RTG “Globalization and Development”, participants of various seminars and conferences for fruitful discussions and comments on the project. We thank in particular our enumerators Cut Voenna Nestya, Dian Islamy, Iqlima Sari, Tria Fitri Deswinda, Indah Sukma Dewi, Evi Afri Yani, Nelfi Ahul Lishani, Siti Amalia Husna, Viona Delinda, Sisca Pramesti Wulandari, Octami Ruliani, Ledi Iksarina, Sri Intan Khairunnisa, Julianti, Fitri Ani, Yuni Shelma, Qurrata Aini, Fityah Rasyiqah, Fiya Irma Safiya, Sulistian Nugraha, Nova Friska, Alya Nurul Mahfudhah, Meka Suci Fitria, Intan Zahrina, Ulpana, Ulya Hakim, Alvi Arhamunnisa, Rheka Fitria, Suhardi, Ema Putri, Khairi Rizkana and Zulfira.