

NBER WORKING PAPER SERIES

BIOMETRIC TRACKING, HEALTHCARE PROVISION, AND DATA QUALITY:
EXPERIMENTAL EVIDENCE FROM TUBERCULOSIS CONTROL

Thomas Bossuroy
Clara Delavallade
Vincent Pons

Working Paper 26388
<http://www.nber.org/papers/w26388>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2019

We are deeply grateful to Sandeep Ahuja, Shelly Batra, and Prateek Ahuja at Operation ASHA for supporting this study. For comments and suggestions, we thank Marcella Aslan, Jishnu Das, Puneet Dewan, Esther Duflo, Rifat Hasan, Sophie Huddart, Reshmaan Hussam, Gabriel Koehler-Derrick, Manoj Mohanan, Patrick Mullen, Madhukar Pai, Dr KS Sachdeva, Bill Thies, as well as seminar participants at Harvard, Warwick, the World Bank, and Duke University, and conference participants at GREThA, NEUDC, PAA, the Hewlett Foundation Workshop on Quality of Healthcare, and USAID's TB Industry Day. For excellent field management and research assistance, we are grateful to Siddhartha Baral, Sruthi Chandrasekaran, Rachna Nag Chowduri, Sadish Dhakal, Radhika Jain, Chand Mazumdar, Mansi Shah, and Ameet Singh. Miral Kalyani and Umry Shukla analyzed the qualitative data. Ravi Agarwal carried out scalability analysis. This project was funded by grants from USAID, AusAid, and the Health Results Innovations Trust Fund administered by the World Bank. Ethical approvals were obtained from the MIT Committee on the Use of Humans as Experimental Subjects (COUHES), the Institutional Review Boards of the Institute for Financial Management and Research (IFMR) in India, and the University of Cape Town, South Africa. This study is registered in the AEA RCT Registry, and the unique identifying number is AEARCTR-0000038. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the World Bank, Harvard Business School, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Thomas Bossuroy, Clara Delavallade, and Vincent Pons. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Biometric Tracking, Healthcare Provision, and Data Quality: Experimental Evidence from Tuberculosis Control

Thomas Bossuoy, Clara Delavallade, and Vincent Pons

NBER Working Paper No. 26388

October 2019

JEL No. I12,I18,O15,O33

ABSTRACT

Developing countries increasingly use biometric identification technology in hopes of improving the reliability of administrative information and delivering social services more efficiently. This paper exploits the random placement of biometric tracking devices in tuberculosis treatment centers in urban slums across four Indian states to measure their effects both on disease control and on the quality of health records. The devices record health worker attendance and patient adherence to treatment, and they automatically generate prompts to follow up with patients who miss doses. Combining data from patient and health worker surveys, independent field visits, and government registers, we first find that patients enrolled at biometric-equipped centers are 25 percent less likely to interrupt treatment—an improvement driven by increased attendance and efforts by health workers and greater treatment adherence by patients. Second, biometric tracking decreases data forgery: it reduces overreporting of patient numbers in both NGO data and government registers and underreporting of treatment interruptions. Third, the impact of biometric tracking is sustained over time and it decreases neither health worker satisfaction nor patient satisfaction. Overall, our results suggest biometric tracking technology is both an effective and sustainable way to improve the state's capacity to deliver healthcare in challenging areas.

Thomas Bossuoy
The World Bank
1818 H Street, NW
Washington, DC 20433
tbossuoy@worldbank.org

Clara Delavallade
The World Bank
1818 H Street, NW
Washington, DC 20433
cdelavallade@worldbank.org

Vincent Pons
Harvard Business School
Morgan Hall 289
Soldiers Field
Boston, MA 02163
and NBER
vpons@hbs.edu

1 Introduction

Principle-agent theory posits that employees may reduce effort in the presence of information asymmetry (Alchian and Demsetz, 1972; Laffont and Martimort, 2002), with damaging consequences for the capacity of the state to deliver social services (Banerjee, Deaton, and Duflo, 2004). Besides adverse effects on actual performance, the quality of performance data may itself suffer when agents exploit information asymmetry to misreport the outcomes of their work (Jacob and Levitt, 2003). Misreporting in the public sector hampers the government’s ability to assess the needs of its citizens, let alone address them: historical accounts of the creation of modern bureaucracies show that possessing accurate data on populations and territories is, in itself, a key component of state capacity (Tooze, 2001; Desrosières, 2002; Brambor, Goenaga and Lindvall, 2019).

These issues are especially salient and consequential in tuberculosis-control programs. With 1.6 million deaths per year, tuberculosis (TB) is the second-deadliest communicable disease globally. India has the highest number of cases: about 1,100 Indians die of TB every day (WHO, 2018). Yet most cases of TB can be cured with a range of affordable, widely available antibiotics. For this to happen, patients need to be diagnosed and they then need to comply with a six-month treatment regimen.

To accomplish effective TB control around the world, the World Health Organization (WHO) promotes a strategy called Directly Observed Therapy, Short Course (DOTS). In India, this strategy is implemented by the Revised National TB Control Program (RNTCP), whose state- and district-level offices leverage numerous partnerships with local NGOs. The RNTCP tasks community-based health workers with referring potential cases to detection labs and directly observing that patients take their medication every few days at the DOTS center. This decentralized organization makes monitoring health workers both critical and difficult. The program’s limited monitoring capacity stymies its efficiency, with direct consequences on morbidity and survival (WHO, 2018). Failure to detect cases early and interruption of treatment are highly likely to result in patient death and the spread of infection.

This paper examines the benefits of using biometric tracking to improve service delivery and reduce misreporting in TB control programs in India. We partnered with Operation ASHA (OpASHA), the largest NGO delivering primary care to TB patients in India, to randomize the roll-out of biometric tracking devices across 129 areas located in urban slums across four states in Northern India, with each

area covering a population of over 20,000 individuals. The innovative technology was developed in partnership with Microsoft Research India and specially designed to optimize usability for low-skill operators in environments with limited or unreliable internet connectivity. It was devised to perform three main functions: ensuring that patients themselves receive their medication from health workers; generating alerts when patients fail to come to the center to take their pills, thereby facilitating rapid follow-up by health workers; and creating a real-time tool for program managers to monitor attendance and performance of health workers. The devices also drastically reduce opportunities for health workers to alter health records by overreporting new cases or underreporting instances of treatment interruption.

While biometric tracking may improve the quality of health care provision and the truthfulness of health records, this technology may also have unintended consequences. Patients weakened by the disease may no longer be able to take home several doses at a time for self-administration or send a relative to the center to pick up their pills, which may lead to more strain and more frequent treatment interruptions. In addition, health workers focusing on ensuring treatment adherence may reallocate efforts away from other important activities, such as identifying early cases.

Our impact evaluation uses information covering 3393 patients and 85 health workers and combines multiple rich sources: several rounds of survey data collected both on patients and health workers; program data collected from the NGO; administrative data collected from government hospitals; independent data collected through random spot checks and observation days at DOTS centers; and semi-structured qualitative interviews with 45 current and former OpASHA health workers and 47 patients. The combination and comparison of program-generated and independently-collected data allows us to reliably estimate impacts both on real TB control outcomes and on the extent of misreporting in public health data.

Our empirical analysis yields two main findings. First, biometric tracking increases patient adherence to the TB treatment regimen: patients' interruption of their treatment, or "default," decreases by 25 percent, and in-person pill intake, a core requirement of the DOTS protocol, improves by 26 percent. These results appear to be driven by a 17-percent boost in provider attendance at the center and an increased frequency of follow-up home visits to patients who miss a dose. Second, biometric tracking measurably decreases misreporting in NGO data and official government registers. The magnitude of the

reduced discrepancies between reported and independently-collected outcome measurements indicates that, absent the devices, at least one-fourth of all default cases go unreported in government registers and one-fifth of all new patients declared are overreported.

We rule out the possibility that patient selection accounts for these results. One possibility suggested by multitasking models in the agency literature (Holmström and Milgrom, 1991; Van Den Berg and Van Der Klaauw, 2006) is that health workers using biometric devices reallocated their efforts towards the indicators tracked by the devices, namely center attendance and patient treatment adherence, at the expense of active case finding. In addition, they may have strategically avoided enrolling patients with a presumed higher likelihood to miss doses and interrupt their treatment, and limited patient notification altogether to keep their caseloads reasonable for treatment adherence activities. However, we do not find any significant impact on the number of patients detected or on their observable characteristics. This decreases the likelihood that our impact estimates may be confounded by differential patient selection induced by the intervention. To provide further support for this conclusion, we check that our results on patient-level outcomes are robust to restricting the analysis to a subset of patients whose selection could not have been affected by the intervention: patients who were enrolled prior to the experiment (but whose treatment continued into the time of experiment).

In addition to these findings, we provide positive evidence on the sustainability and acceptability of the technology. The decrease in overreported patient detections reduced the earnings of health workers with biometric tracking devices by 7 percent. However, biometric tracking also reduced their workload, and overall, their job satisfaction did not worsen. Patients did not report lower levels of satisfaction with their treatment either, despite having to visit the DOTS centers more frequently themselves. Consistent with maintained health worker and patient satisfaction, the technology remained in place and functional over time in the vast majority of the centers, and its impact did not taper off significantly over the duration of the experiment.

The main contribution of this paper relates to the literature on public administration and state capacity in developing countries. As the primary provider of social infrastructure, the state is called to play a critical role in eradicating poverty and building human capital, which requires strong capacity (Besley and Persson, 2010; Page and Pande, 2018). Weak capacity can result in vast inequalities in access to

healthcare across countries and along the wealth distribution (WHO and World Bank, 2017).

The first dimension of state capacity we address is the efficiency of the personnel of public agencies and their partner NGOs as well as their ability to serve the poor and underprivileged (Rasul and Rogger, 2016; Finan, Olken, and Pande, 2017). We make three contributions to the corresponding literature.

First, a large number of experiments tackle absenteeism of public service providers, a key factor undermining state capacity to deliver education and healthcare (Chaudhury and Hammer 2004; Banerjee and Duflo, 2006). While technology-based attendance monitoring has shown positive impacts on the presence of teachers (Duflo, Hanna, and Ryan, 2012) and health workers (Banerjee, Duflo, and Glennerster, 2008; Dhaliwal and Hanna, 2017; Callen et al., 2018), the evidence is mixed when it comes to attempts to strengthen service delivery through community-based monitoring (Björkman and Svensson, 2009; Banerjee et al., 2010; Banerjee et al., 2018; Raffler, Posner, and Parkerson, 2019). Results found by Duflo, Hanna, and Ryan (2012) and Banerjee, Duflo, and Glennerster (2008) suggest that monitoring works better in the presence of financial incentives. Our paper extends this literature by exploring whether positive results can be achieved by combining monitoring with productivity gains for the frontline worker.

Second, we contribute to the evidence on ways to boost medication adherence, which poses a particularly difficult challenge for state capacity as it requires shaping daily individual decisions (Sabate, 2003). As a result, even the most effective interventions targeting medication adherence have modest effects (McDonald, Garg, and Haynes 2002). Demand-side interventions, targeted to the patients, include the provision of tailored information on pill packaging (Cohen and Saran, 2018), text message reminders (Pop-Eleches et al., 2011; Raifman et al., 2014), and financial incentives (Giuffrida and Torgerson, 1997; DeFulio and Silverman, 2012). Our study contributes to the less-studied field of supply-side interventions targeting healthcare providers acting on behalf of the state, which are generally part of larger intervention packages and whose impacts are not as well documented (Nieuwlaat et al., 2014).

Third, we contribute to a larger literature on leveraging technology to increase the state capacity to deliver services. Two aspects of technology are used in the intervention evaluated here: connectivity that enables real-time oversight of service delivery in remote areas, and biometric identification that enables the creation of reliable registers of service recipients and their adherence to treatment. In the health sector, mobile technology has been used to send text-message reminders to boost adherence to anti-malarial

treatment (Raifman et al., 2014) and antiretroviral treatment (Lester et al., 2010; Pop-Eleches et al., 2011). Mobile communication can also be targeted to health workers, to help them manage appointments and drug inventory and to improve their adherence to treatment protocols (Zurovac et al., 2011; Otto et al., 2015). Biometric identification has seen a rapid growth in the past decade and is now used as a mechanism for delivering public programs in more than 80 developing countries, with applications ranging from financial services to social transfers, civil service reforms, and health policies (Zuniga, Win, and Susilo, 2010; Gelb and Clark, 2013). Evidence on the impact of biometric identification for healthcare provision is encouraging but still scarce. Bhatnagar et al. (2012) and Snidal et al. (2015) study the use of biometric devices for TB control through qualitative and observational studies. We contribute a randomized design to solve selection issues and rigorously quantify the impact of the technology. Beyond the health sector, studies have shown biometric technology to increase the efficiency of payments in cash-for-work and pension programs (Muralidharan, Niehaus, and Sukhtankar, 2016) and fuel subsidies (Barnwal, 2016), and to improve credit repayment rates (Giné, Goldberg, and Yang, 2012).

A second, less-studied, dimension of state capacity pertains to the ability to measure individual citizens' needs and public-sector performance accurately, in order to base policy decisions on a sound assessment of priorities (Lee and Zhang, 2017; Koehler-Derrick, 2019). Historians of developed countries have shown how the construction of reliable public statistical systems was a driving force behind the establishment of modern states (Scott, 1998; Tooze, 2001; Desrosières, 2002; Sretzer, 2007). Collecting accurate and comprehensive data on service delivery and individual-level outcomes is particularly critical to support effective public health systems and high-quality patient care (Oomman, 2013; OECD, 2013). But health care data quality limitations persist in developed countries (e.g., Peabody et al., 2004), and they are more acute and widespread in the developing world (Mikkelsen et al. 2015; Mohanan, Hay, and Mor, 2016; Downey et al., 2018). Little improvement has been made in the past decades, so that the quality even of statistics as fundamental as birth and death rates remains poor overall in the world (Mahapatra et al., 2007). In the context of health systems in India, Morton et al. (2016) show that data from a hospital insurance program are practically unusable, eliminating critical information for state efforts to improve healthcare for the poor, and Phillips et al. (2019) find significant overreporting in government health facility data relative to externally collected data, in particular on incentivized indicators. Despite the im-

portance of data quality for healthcare delivery, there is a paucity of rigorous evidence on the causes of health data incompleteness and misreporting and on interventions that can successfully address them. We provide groundbreaking evidence on the effects of biometric tracking technology on data accuracy. We find that by significantly strengthening the quality and reliability of NGO records and government health registers, this technology has the potential to substantially improve state capacity in the consequential area of TB control and beyond.

By measuring impacts on the accuracy of administrative data, we finally advance the literature studying the effects of monitoring on corruption (e.g., Slemrod, Blumenthal, and Christian, 2001; Nagin et al., 2002; Di Tella and Schargrodsky, 2003; Ferraz and Finan, 2008; Kleven et al., 2011; Ichino and Schündeln, 2012; Callen and Long, 2015). Becker and Stigler (1974) propose a seminal model linking fraudulent behavior both to the nature of monetary incentives attached to corruption and to the likelihood and consequences of being caught. Among the related literature, our paper is closest to Olken (2007) and Duflo et al. (2013), who study fraudulent misreporting in the context of road construction in Indonesia and industrial plant pollution in Gujarat, respectively. Similar to the measures used in these studies, our measure of misreporting is based on the comparison between official data and independently-collected data for the same outcomes of interest.¹ However, the intervention we study differs substantially: it strengthens top-down monitoring through technology instead of by government or third-party auditors. This difference is critical given that in-person auditors may themselves engage in misreporting, unless properly incentivized (Duflo et al., 2013).

From a policy perspective, our results contribute to ongoing efforts in India to strengthen the delivery of social programs through biometrically enhanced mechanisms. They point to the benefits that unique identification such as India's Aadhaar system may have on TB control programs, both to facilitate the delivery of primary care and to dramatically enhance the reliability of national health registers. This latter point is of even greater importance as direct benefits transfers to TB patients such as financial and nutritional support will be implemented to incentivize adherence (RNTCP, 2017) and could also be used to encourage patient referrals (Goldberg, Macis, and Chintagunta, 2019). Policy implications of our study

¹Recent papers use a comparable measurement strategy to study misreporting in macroeconomic indicators: Henderson (2011) and Martinez (2019) identify the manipulation of GDP growth statistics by comparing them to nighttime luminosity data, and Cavallo and Rigobon (2016) track the difference between official inflation numbers and actual increases in online prices.

for TB control programs in India are discussed in Section 6.

The remainder of the paper is organized as follows. Section 2 provides background on TB in India and on the intervention evaluated in this paper. Section 3 describes the experimental design, measurement strategy, and internal validity checks. Section 4 presents the results on TB control outcomes and provider efforts. Section 5 discusses the empirical strategy and findings related to misreporting. Section 6 presents evidence on the acceptability, sustainability, and scalability of the intervention, and Section 7 concludes the paper.

2 Setting: policy challenge and intervention

2.1 State capacity and TB control in India

TB is the largest infectious killer of adults, claiming nearly 1.3 million lives annually, of which about 410,000 are in India (WHO, 2018). As a comparison, the number of TB deaths every ten days in India is roughly equal to the number of deaths during the entire 2014–2015 Ebola outbreak in West Africa. While severe coughing is the hallmark of the disease, those infected also experience weight loss, listlessness, and crippling pain. Over 80 percent of the patients we surveyed reported suffering from respiratory problems, body pain, and other general health problems. As a result, affected people are often unable to work or care for their families, with direct consequences for household livelihoods. In our surveys, about half of the patients who were gainfully employed at the time they contracted TB reported losing more than 80 percent of their workdays and income as a direct consequence of the disease.

Roughly 97 percent of all new TB cases are affected by drug-susceptible strains of the disease, and effective first-line antibiotics are available at no cost in India. These baseline facts should make reducing TB mortality attainable, yet challenges endure in two areas that reveal limits of Indian state capacity: early diagnosis and enrollment of infected patients; and ensuring completion of the full course of treatment.

The first challenge, known as early case finding, is critical to improve the condition of infected patients and prevent the spread of this highly contagious, airborne disease. Contagion is particularly rapid in urban slums characterized by high population density and poor sanitation. In our sample, the average patient lives in a two-room house with five other family members. Fewer than a quarter of patients have

in-home access to tap water. As a result, disease transmission occurs primarily within families and communities. Over 40 percent of the patients we surveyed reported that at least one other member of their family had a history of the disease.

Early case finding is hampered by low levels of awareness and information on TB, poor access to care, and fear of social stigma about contracting the disease. Just half of the patients know about the most prominent symptoms of TB before starting the treatment, and only two-thirds know that it is a communicable disease (Huddart et al., 2018). In qualitative interviews, many patients showed lack of knowledge about TB symptoms. They did not suspect that they had TB at their first consultation with a health provider, and they continued to cycle through different private and government doctors without suspecting TB. In fact, low-quality primary health care in India (Das and Hammer, 2007; Mohanan, Hay, and Mor, 2016) and the prevalence of poorly qualified practitioners as primary points of care also hinder the timely referral to a TB detection center (Satyanarayana et al., 2015; Das et al., 2015). Among the surveyed patients who consulted a medical facility in the three months before starting their treatment in OpASHA's centers, 60 percent went to see a private doctor and 35 percent received an injection—a blatantly inadequate response to TB symptoms. Finally, almost half of the patients in our sample reported fearing social discrimination if it became known in their community that they had TB.

The second challenge, treatment adherence, is important to prevent relapses and the development of drug-resistant strains of the disease, which are much more difficult and up to six times more expensive to treat (WHO, 2018). However, improving adherence is far from straightforward. To be cured, an infected patient needs to maintain a complicated regimen of up to seven drugs at a time, three times a week for a two-month period—and then once a week for a further four months. These drugs often have strong side-effects. Two-thirds of the patients in our study reported experiencing physical or psychological side effects, including half complaining of severe effects such as nausea, vomiting, fever, and stomach pains. After about two months of treatment, TB symptoms begin to wane, but the side effects of the drugs remain. This leads patients to discontinue treatment prematurely. Treatment interruptions in our sample occurred most frequently in the third month (median at 70 days after starting the treatment). Of the patients who were going through their second round of TB treatment, only about half had completed their earlier round.

2.2 The DOTS model

In response to the joint challenges of increasing detection and ensuring adherence, India and other countries have adopted the Directly Observed Therapy, Short Course delivery method. In the DOTS model, patients are required to ingest each TB treatment dose under direct observation of trained providers (DOTS providers) based in local care centers (DOTS centers).

DOTS centers are widely dispersed, to facilitate patient access and the detection of new patients close to home. In remote or overcrowded places where public hospitals are overburdened or inadequate, the strategy calls on governments to collaborate with the private sector to achieve adequate service delivery. Private care providers include a vast array of actors ranging from unqualified practitioners to mission hospitals or NGOs, and the modalities of the partnerships vary according to the nature of the provider. Such arrangements are meant to strengthen the state outreach capacity and improve TB-control outcomes, especially for vulnerable, hard-to-reach populations.

In practice, to enroll in the DOTS system, symptomatic patients first need to visit a Designated Microscopy Center, a medical unit generally located within a public hospital. Lab technicians detect TB by examining sputum samples under a microscope to determine the presence of bacteria—a technique called sputum smear microscopy. If a patient is found to be TB-positive, she is recorded in a government register under a unique TB number and assigned to a DOTS center, depending on her location. The TB Control Program sends a box of drugs bearing the patient's name and containing the full treatment regimen as well as a personal treatment card to that center.

The patient can then start her treatment. The DOTS provider in charge of her center tracks treatment adherence by recording every visit made by the patient to the center on her treatment card, and declares treatment interruption, or default, if the patient misses doses for four consecutive weeks. Additional tests are done during the course of the treatment on sputum samples collected at the Designated Microscopy Center to assess progress and determine outcomes. At the end of the six-month period, the DOTS provider records the treatment outcome on the treatment card and reports this information to government registers by handing the card over to the local Designated Microscopy Center.

Although the implementation of DOTS has significantly improved efforts to contain TB over the last two decades (WHO, 2014; also see McMillen, 2015 for a historical discussion), the system still

suffers from several shortcomings. Despite the spread of the DOTS centers, visiting a center three times a week remains a significant burden, especially for those physically debilitated by the disease. In our study sample, patients live on average three miles away from their treatment center, corresponding to an average 11 minutes of reported travel time. In addition, the high level of decentralization reduces the TB Control Program's administrative capacity to monitor DOTS providers and ensure robust service delivery. The record system in DOTS centers is entirely paper-based, and treatment cards are only collected by government staff from the TB Control Program once every one or two months. DOTS providers monitor pill intake and TB treatment outcomes by cross-checking different roster sheets, a time-consuming and error-prone process which also leaves room for opportunistic misreporting.

2.3 Operation ASHA

OpASHA, our partner organization, is a large NGO based in New Delhi and operating TB DOTS centers under public-private partnerships in Northern Indian states. OpASHA establishes most of its centers in host facilities such as private medical practices (72 percent of the centers in the sample), dispensaries (9 percent), temples, schools, and pharmacies. Taking advantage of existing facilities saves operating costs while ensuring that the centers stay open at least as much time as the facility does. The DOTS provider is a private doctor, pharmacist, or shopkeeper, depending on the hosting facility.

In addition to the DOTS provider, OpASHA hires health workers and assigns them to treatment centers. Because patients must show up every other day to their center, the organization has each health worker operate two different centers, open on alternate days, in accordance with the pill intake schedule of the patient population. In areas with lower population density, OpASHA does not establish fixed treatment centers but hires mobile health workers to go around the communities and deliver information and treatment pills at home. Operations are supervised by one manager per city.

The health workers are the central piece of OpASHA's model. When physically present in the DOTS center, they take up the responsibility of the DOTS provider, including supervising patients' pill intake and filling out their treatment cards. In addition, they are responsible for the following two objectives. First, they are tasked with promoting early case finding. They canvass the area spreading awareness of TB symptoms and the availability of treatment, as well as asking health questions to households.

When they come across individuals showing TB symptoms, they provide information on the disease and the treatment process and encourage them to get tested at the Designated Microscopy Center. When respondents are too weak to go to the Center in person, health workers collect sputum samples, which they deliver on their behalf. A significant share of symptomatic individuals visit the Designated Microscopy Center of their own accord because they have prior knowledge of the medical care system or are referred by existing or former patients. In these cases, the interaction with the health worker only starts when they begin taking their treatment at the DOTS center. Second, health workers are responsible for tracking patients who missed a dose and following-up with them by phone or in person, at their house, to bring them back to the regular course of treatment. In qualitative interviews, health workers describe strategies to boost medication adherence of defaulting patients, such as educating their families and neighbors to encourage the patient, providing emotional support, increasing the frequency of home visits, or giving patients several doses at a time to spare them trips to the center.

Health workers receive a salary averaging Rs. 5600 per month (about USD 90 at the time of the experiment), with variations across cities. OpASHA measures their performance by seven indicators, including the number of new TB cases identified and the number of patients completing their treatment, as well as other indicators related to the maintenance of the center. In all cities but three, health workers receive financial incentives based on the number of new cases identified and amounting to 19 percent of their total salary on average. Those disparities do not threaten the validity of results since treatment assignment was stratified at the city level.² Qualitative interviews show that health workers perceive a strong pressure to achieve performance targets measured by the number of detections and the number of defaults. They commonly complain about the level of stress and their fear that poor performance would lead to salary cuts, sanctions, or contract termination.

2.4 The intervention

Our experiment evaluates biometric tracking devices installed in DOTS centers and designed to help health workers monitor the treatment adherence of TB patients. OpASHA developed this technology, called eCompliance, in partnership with Microsoft Research India. The devices consist of a tablet equipped with a SIM card and a 2G internet data plan connected to a fingerprint scanner. A user-friendly

²Three health workers—two in the control group and one in the treatment group—also received default-based incentives.

application allows health workers and DOTS providers to navigate the system intuitively. In biometric-equipped centers, all patients register their fingerprints at the beginning of their treatment and scan them again every time they come to the center to take their pills. The health workers can use a second portable terminal to record their home visits to patients. The interface allows health workers to access the list of patients that have missed a daily dose and warrant some specific follow-up action.

The system also uses the health workers' fingerprints for authentication and records the date and time when they log in and out of the system. In addition to a tool for health workers to monitor patients, it is therefore also a tool for NGO management to monitor health workers' effort and performance. The data are stored in the terminal located in the health center and sent daily to a server located in New Delhi for both backup and monitoring purposes. Communication with the server is done through 2G and regular SMSs, ensuring the maximum coverage possible even in remote areas with limited or unreliable internet connectivity. The entire system costs less than USD 250 per center.

3 Research design

3.1 Experimental design and sample

The study sample includes nine cities in four North Indian states, shown in Figure 1: Bhopal, Gwalior, Indore, Sagar (in the state of Madhya Pradesh); Korba, Raipur, Durg-Bhilai (Chhattisgarh); Bhubaneswar (Odisha); and East Delhi (Delhi). It includes 129 scattered and non-overlapping areas, grouped in 65 clusters. All clusters include two areas except one cluster which includes only one area. The vast majority of the clusters (52) have a fixed DOTS center established in each area. Another seven clusters do not have any center; each of these clusters has one mobile health worker devoted to that cluster's two areas. Finally, there are six hybrid clusters which each have one area with a fixed center and one area that receives mobile treatment.

[Figure 1 about here]

At the beginning of the experiment, OpASHA allocated a unique health worker to each cluster. The study team then conducted the randomization at the cluster level, after stratifying by city and health worker type (fixed, mobile, or hybrid). 34 clusters were allocated to the treatment group, and the re-

maining 31 to the control group. Two to four weeks after randomization, OpASHA installed biometric tracking devices in centers (or provided them to mobile health workers) in the treatment group and gave all the health workers who were selected to use the technology a dedicated short training. The experiment lasted between 12 and 14 months, depending on the state.

The experiment includes a total of 85 health workers. 65 (one per cluster) were hired at the outset of the experiment and 20 joined as replacements. 16 health workers (24.6 percent) present at the beginning of the experiment as well as six replacements left the organization for various reasons including poor performance (48 percent), alternative work opportunities (24 percent), and illness (4 percent). The likelihood to leave was not significantly affected by the intervention (Appendix Table A1, Panel A), and differences in observable characteristics of attritors (Panel B, columns 1 to 6) or replacements (columns 7 to 12) across treatment and control groups are small in magnitude. Out of 73 differences shown in Table A1, Panel B, only two are significant at the 5 percent level and three at the 10 percent level.

The equality in retention rates across treatment and control groups is consistent with the fact that biometric tracking did not significantly affect health workers' satisfaction with their jobs (we return to this in Section 6). The logistics of replacing health workers also limited the scope for differential selection. The simplicity of the technology made it unnecessary for OpASHA to select particular profiles to operate centers equipped with the devices. To allow for a rapid replacement of health workers terminating their contract, employee selection was based on a waitlist of existing ranked applicants. As a result, employee turnover was handled rapidly, with departing health workers being replaced on the day of their departure in most cases, and within eight days on average. Anecdotal evidence from the NGO management suggests near to 100 percent acceptance rates by candidates who were offered the position.

3.2 Data sources

We use many data sources in the analysis: several rounds of closed-ended surveys administered both to patients and health workers, independent day-long monitoring of DOTS centers (henceforth, "observation days"), random spot-checks of the centers, qualitative interviews with patients and health workers, and program and administrative data from OpASHA and government registers. Apart from

program and administrative data, all data were collected specifically for the purpose of this study.³

Survey data

Patient surveys were administered to all patients enrolled in OpASHA centers up to the last month before the end of the experiment. Patients starting and finishing their treatment within the time span of the experiment received surveys at two points: an entry survey about a month after the beginning of the treatment (specifically between two and six weeks from the start date), and an exit survey within a month after the end of the treatment, regardless of whether the end was due to completion or interruption. Patients who were already well into their treatment when the experiment started did not receive an entry survey but were instead administered an augmented exit survey including socio-demographic information. Patients who were still undergoing treatment by the end of the experiment were administered an early exit survey.

The sampling frame was provided by the individual treatment cards maintained at each center. In total, 4910 patients were recorded as undergoing treatment in one of the areas included in the experiment. The survey team contacted all patients to set up appointments, unless the health worker asked them not to, citing reasons such as the patient hiding her condition from her family. Patients had the option to conduct the survey outside of their homes if they wished to have privacy from their own family or community, or to refuse the survey. Most patients consented to being interviewed at home. Surveyors made three attempts to locate the patient before declaring them “Not found.” Survey completion rates are 70.4 percent for entry surveys and 64.2 percent for exit surveys. We discuss factors responsible for non-completion in Section 5.3. 25 percent of patients were selected to receive both an entry and an exit survey and were successfully surveyed both times, resulting in a database of 4600 observations from 3393 unique patients.

Health workers were scheduled to be surveyed three times: just before the installation of the biometric devices (baseline), after about six months of intervention (midline), and at the end of the intervention, six to eight months after midline (endline). More than two-thirds of them successfully completed all three, and 89 percent completed at least a baseline survey and one follow-up (either midline or endline). Health worker survey data include 219 surveys of 85 health workers. We systematically approached

³Table A2 in Appendix A checks that the presence of biometric devices did not create systematic differences in the availability of the various data sources across the treatment and control groups. Out of 14 differences shown in the table, two are statistically significant: the number of random spot checks and the number of observation days patient data are slightly larger in the treatment group.

those who quit their job during the intervention and administered an additional module on the reasons for employment termination.

Observation days and random spot checks

Enumerators conducted random day-long visits to the fixed DOTS centers or to mobile health workers. They started the observation day by gathering the list of all patients scheduled to come on that day, then proceeded to collect information on attendance (both of the health worker and of each individual patient visiting the center), activities carried out at the center and during home visits, and the usage of biometric devices (in the equipped centers). Each center received an average of two random visits per month over the duration of the experiment, resulting in a dataset of 3,084 observation days. We merged these data with the roster of patients in the experiment using three identifiers sequentially (TB number, lab number, and patient name) and verifying each uncertain merge by hand. The coverage of observation days is large. Overall, 85.7 percent of patients in the full sample appear in at least one observation day, and the average number of observations per patient is 3.7. This provides us with a thorough independent measure of medication adherence at the patient level, in both biometric and control centers.

Although the dates of the observation days were random, health workers could of course adjust their behavior, for instance the time at which they left the center, once they saw the enumerator. To further strengthen the measurement of health worker attendance, senior survey staff conducted random, unannounced spot checks of the DOTS centers and collected additional information on whether the center was open and the health worker present. On average, each center received seven spot checks over the duration of the experiment, for a total of 916 observations. These data provide a second independent measurement of provider attendance.

Program and administrative data

We collected center-wise data sheets, salary slips, and data from official TB registers. The center-wise data sheets are monthly reports prepared for each center or mobile area by health workers and used by OpASHA to track their performance and calculate salaries. They include data on the number of new patients enrolled and the number of defaulting patients. Salary slips include health worker-level performance data and the breakdown of monthly salaries. TB registers are the official source of patient-level information for the TB Control Program. They are kept by government TB officers at the central

hospital and typically cover several Designated Microscopy Centers. They list the name and address of all enrolled patients, the dates and results of their initial and follow-up sputum tests, and the outcome of their treatment. We obtained access to TB registers in all study sites except Indore. We digitized and matched these data with our other datasets using several identifiers in sequence and verifying unclear cases by hand. Overall, TB register data were successfully matched for 94.1 percent of all patients in cities where access was granted, allowing us to use the experiment to assess the extent of misreporting in official government registers.

Qualitative interviews

The study team conducted semi-structured interviews with 45 OpASHA health workers and 47 patients. Respondents were randomly sampled from the pool of all the health workers and patients who were part of the experiment at the time. To elicit candid responses, we complemented the sample with former health workers no longer employed by OpASHA.

Interviews were conducted primarily in Hindi based on an interview template including “grand-tour,” structured, and hypothetical interaction questions (Morse, 2014; Bhattacharya, 2017). Open-ended questions to health workers related to their intrinsic motivation, effort allocation, and data reporting. Questions to patients related to their TB knowledge, pathways to diagnosis, and experience with their current treatment. We transcribed and translated interviews and used a qualitative data analysis software to code them along major themes of interest, some pre-defined at the inception of the study and others emerging from the transcripts.

3.3 Summary statistics and balance checks

Table I presents summary statistics for health workers in the full sample used in the analysis (columns 1 to 3). Health workers were 31 years old on average. Almost three-quarters were men, 40 percent belonged to the general caste, and 81 percent were Hindus. They were well-educated on average, with more than a third having completed tertiary education. Although 75 percent of the health workers had previous work experience (for an average of nine years), only 16 percent worked in the social/NGO sector before joining OpASHA. A majority of health workers lived in one of the areas they covered, mostly in decent conditions: almost all of them had electricity and 56 percent had access to tap water in their home.

More than half owned their house. They had some exposure to technology: 57 percent knew how to use a computer and 40 percent had an email account. Although some health workers expressed interest and motivation in social work in qualitative interviews, most mentioned doing this work because it is close to home and salaries are acceptable.

[Table I about here]

Table I further presents balance checks for the subsample of health workers present at the outset of the experiment (columns 4 to 8). We show means and standard deviations separately for the control and treatment groups, the difference between the means of the two groups, and the p-value of a test of the null hypothesis that they cannot be distinguished from each other. Out of 41 differences, two are significant at the 5 percent level and seven at the 10 percent level. In addition to a few differences in their socio-demographic characteristics, health workers in the treatment group were less likely to have any previous work experience and they reported a lower number of patients enrolled in their centers prior to randomization.

Table II shows summary statistics for all patients (columns 1 to 3). Patients were 34 years old on average, and 57 percent were men. Most belonged to the most-deprived castes in India: Scheduled Castes and Scheduled Tribes (33 percent) and Other Backward Classes (36 percent). Only 18 percent belonged to the better-off general caste category, as opposed to 40 percent of health workers. Most of them lived in crowded, unsanitary conditions with high exposure to TB (see Section 2). Balance tests shown in columns 4 to 8 are based on patients enrolled prior to the outset of the experiment because of the possibility that the intervention affected the pool of patients enrolled in treatment and control centers after its start (we explore this possibility in Section 4.3 and show it did not materialize). Out of 31 differences shown in Table II, two are significant at the 5 percent level and three at the 10 percent level, which is in line with what would be expected.

[Table II about here]

In the analysis, we show the robustness of our results to controlling for health workers' and patients' socio-demographic characteristics as well as baseline outcomes.

4 Impact on TB control outcomes

4.1 Empirical strategy

We report impact estimates in two main areas. First, we measure the impact of biometric tracking on treatment adherence, which was the main TB control outcome targeted, as well as other patient-level outcomes, with the following OLS specification:

$$Y_{it} = \alpha_1 + \beta_1 T_i + X_i' \gamma_1 + \sum_s \delta_{i1}^s + \varepsilon_{i1}, \quad (1)$$

where Y_{it} is the outcome for patient i at time t , T_i is a dummy equal to one if the patient was assigned to a center equipped with a biometric tracking device (or to a mobile health worker using a biometric device) and zero otherwise, δ_{i1}^s are strata fixed effects, and X_i is a vector of patient controls. Patient controls include gender, age, household size, time taken to go to the center, dummies indicating whether the patient belongs to the general caste category, whether she is Hindu, whether she can read and write, whether she owns her house, whether she has always lived in the area, whether she has lived there for more than six years, whether she was enrolled after the beginning of the experiment, and whether the observation was collected during the entry survey. For each control variable, we finally include a dummy equal to one when the control is missing and replace the missings by 0's to avoid dropping observations. We use various measures of adherence, rule out the existence of a confounding factor linked to differential patient selection, and look into the impact on health outcomes. In Appendix B, Tables B1 through B6, we check the robustness of all patient-level results to restricting the sample to patients enrolled before the beginning of the experiment (but whose treatment continued into the time of experiment) and whose selection can thus not be affected by the intervention.

Second, we turn to the impact of the intervention on health worker effort, measured by attendance and intensity of activities, using the following specification:

$$Y_{it} = \alpha_2 + \beta_2 T_i + Z_i' \gamma_2 + \sum_s \delta_{i2}^s + \varepsilon_{i2}, \quad (2)$$

where Y_{it} is an outcome for health worker i at time t , T_i is a dummy equal to one if the health worker

was assigned to the treatment group and zero otherwise, δ_{i2}^s are strata fixed effects, and Z_i is a vector of health worker controls. Health worker controls include gender, age, household size, dummies indicating whether the health worker belongs to the general caste category, whether she is Hindu, whether her level of education is class 12 or below, whether she completed tertiary education, whether she has any previous experience, and whether she owns her house, as well as baseline outcomes when available. Again, for each control variable, we include a dummy equal to one when the control is missing and replace the missings by 0's.

In this section, all outcomes are constructed using data collected independently by the research team. In particular, we focus on verified patients, defined as patients who were successfully surveyed by independent enumerators in face-to-face interviews, due to evidence indicating that some of the patients found on treatment cards (our sampling frame for patient surveys) were fabricated, an issue we return to in Section 5. Strata fixed effects are included in all regressions, and patient and health worker controls in some of them. In all regressions, we adjust the standard errors for clustering at the level of the cluster of areas, since randomization was conducted at this level. Although our main tables use regular cluster robust standard errors, Tables C1 through C9 and D1 through D9 in Appendices C and D show all the results with p-values obtained using the wild cluster bootstrap procedure (Cameron, Gelbach, and Miller, 2008) and the pairs cluster bootstrap procedure (Esarey and Menger, 2019). Using these alternative clustering methods does not affect our results in any substantial way.

4.2 Impacts on treatment adherence

Table III reports results on treatment interruption and adherence, the main disease control outcomes targeted. TB treatment outcomes are measured using exit patient surveys. Following the WHO guidelines, we define default as a dummy equal to 1 for patients who stopped taking pills before completing four months of treatment or missed doses for a period of two months or more, and 0 otherwise. Columns 5 and 6 of Panel A show that the likelihood to default drops by 1.8 percentage points, from a mean of 7.3 percent in the control group. This 25 percent drop is statistically significant at the 5 percent level and robust to the inclusion of patient controls. Its magnitude makes it highly significant from economic and public health standpoints too, given the risk that defaulting patients develop and spread a drug-resistant

strain of the disease and the cost and difficulty of treating it.⁴ There is no significant impact on other treatment outcomes, but there is suggestive evidence that the decrease in defaults resulted in an increased likelihood that patients were still receiving treatment by the time of the exit survey: indeed, the effects on both outcomes are of similar magnitude and opposite signs.

[Table III about here]

In Table III, Panel B, we report program impact on various measures of treatment adherence. Daily adherence is measured routinely by health workers through the individual treatment cards provided by the TB Control Program. Yet in practice, the quality and accuracy of these cards vary tremendously and there is very limited capacity to ensure truthful reporting. Therefore, the results presented here rely instead on the large number of independent observation days organized in all DOTS centers. On these days, an independent enumerator would record all individual visits made to the center by patients or surrogates, and all visits made by the health worker to the patients' homes.

We compute a first adherence measure as the fraction of patients who did receive their dose on an observation day relative to the total number of patients who were supposed to. Columns 1 and 2 show that on an average day, the share of patients who came to the center in person increased by 13.9 percentage points in the treatment group, from a mean of 54.2 percent in the control group. This 26 percent jump is significant at the 1 percent level.⁵ The impact does not simply correspond to a substitution effect, in which patients who would normally send a trusted relative to the center to pick up the pills now have to make it to the center themselves to scan their fingerprints. As columns 3 and 4 show, the impact on adherence is lower but still large and significant after including the possibility that patients send a relative or that the health worker deliver the pills at the patients' homes.

We then compute another measure of treatment adherence as the likelihood that observation-day data demonstrate that the patient did not default from treatment according to the WHO definition. This requires having enough patient-level observation points to ascertain that they did not interrupt their treatment after the fourth month and did not miss doses for more than two months. In columns 5 and 6 of

⁴The effect on defaults is also visible and of almost exact identical size among patients enrolled before the beginning of the experiment, but not significant: our statistical power is much lower in this subsample, which includes only 28 percent of the observations in the full sample (Appendix Table B1, Panel A).

⁵The effect is of similar magnitude and also significant at 1 percent when restricting the sample to patients enrolled before the beginning of the experiment (Appendix Table B1, Panel B).

Table III, Panel B we first take a strict definition of pill intake in which the patient needs to come to the center in person (similar to the first definition of adherence used in columns 1 and 2). With an average of 3.7 data points per patient, we can independently rule out the occurrence of default for 4.6 percent of patients in the control group—admittedly a very low fraction of non-defaulters since the default rate in the control group is measured at 7.3 percent. Results show an increase of 3.9 percentage points (85 percent), significant at the 1 percent level, in this fully independent measure of confirmed non-defaults. In columns 7 and 8 we expand the scope of pill intake methods to include delivery to a relative or at home (similar to the definition used in columns 3 and 4), and again find a positive impact of biometric tracking on treatment adherence, significant at 10 percent.

Finally, Table III, Panel C leverages patient surveys to confirm impacts on compliance with some key features of the DOTS protocol defined by the WHO. The share of patients who declare sending another person to pick up their pills for them, a breach of the DOTS protocol, is massively reduced by the introduction of biometric tracking. From a mean of 36.7 percent in the control group, this share drops to 14.9 percent in the treatment group, a 21.8 percentage points difference. As discussed above, this reduction does not offset the positive impact of the intervention on patient adherence. Similarly, there is a large reduction in the share of patients who declare picking up pills for one week or more (i.e., three or more separate doses of medication) and would then not be observed ingesting them. Both impacts are significant at 1 percent.

Overall, these results show that biometric tracking is an effective instrument for improving adherence to the DOTS protocol and increasing daily pill intake, translating into a lower likelihood of default.

4.3 Impacts on detections and patient selection

We now investigate whether the sizeable impacts of the intervention on treatment adherence came at the cost of harming the other key objective of TB control programs: early case finding. Health workers in the treatment group may have reallocated effort toward default-prevention activities to the detriment of detection activities (for detrimental multitasking effects, see Holmström and Milgrom [1991]). There could also be strategic decisions by health workers to avoid enrolling patients presumed to be non-compliant and to limit patient notification altogether in order to keep their caseload reasonable for

treatment adherence activities, which they know will be monitored more stringently.

Table IV reports the impact on the number of detections of verified patients. Panel A shows that there is no significant impact on the number of verified patients detected by health workers per month and treatment area. This rejects the multitasking hypothesis and suggests that there was no strategic adjustment of the number of detections by the health workers to facilitate future default-prevention work.

[Table IV about here]

We also examine whether health workers equipped with biometric tracking devices focused their detection efforts on patients who may be less prone to defaulting in the future. Panel B of Table IV considers the same observable characteristics used in Table II and displays balance tests for a different sample: patients enrolled after the beginning of the experiment. New patients were very similar overall in the treatment and control groups. Out of 31 differences, one is significant at the 5 percent level, and four at the 10 percent level. This result shows that health workers do not strategically adjust their patient detection efforts, and implies that our impact estimates on treatment adherence in the full sample of patients, shown above, are unlikely to be confounded by differential patient selection induced by the intervention.

Table A3 in Appendix A provides additional support for this conclusion. In Panel A, we consider verified patients enrolled after the beginning of the experiment in the control group and regress their likelihood to default on all observable characteristics. Three characteristics are significantly correlated with default: the fraction of defaults increases with patients' distance to the center, and patients of a religion other than Hindu or Muslim are relatively less likely to default, as are patients who can read. Reassuringly, Table IV shows that all three variables are well-balanced across the control and treatment groups. In Table A3, Panel B, we predict the likelihood to default of all verified patients enrolled after the beginning of the experiment based on the point estimates from Panel A, and we regress this outcome on the treatment dummy using a specification of the form in equation 1. The effect is not significant and is close to zero, indicating that based on their observable characteristics, and absent the intervention, patients in the control and treatment groups could be predicted to be equally likely to default. This result is robust to excluding time to go to the center and distance to the center (two variables available for fewer than 90 percent of patients) from the multivariate regression of default on observables (column 2) and to

including patients enrolled before the beginning of the experiment in that regression (columns 3 and 4).

4.4 Impact on patients' health

We use a comprehensive health section from the patient surveys to test whether the improved adherence to treatment induced measurable improvements in patients' health. Table V, Panel A reports results on physical health, measured by the likelihood to be cured (a dummy equal to 1 for patients indicating that a sputum test or x-ray result confirmed that they are no longer sick or that the OpAsha health worker or a doctor told them they were cured), patients' ranking of their own health on a scale from 1 to 10, and indices indicating difficulties in performing activities (such as drawing water from a well and routine housework), health symptoms (such as cough and respiratory problems), severe health symptoms (if these symptoms were severe), and increased health symptoms (since the patients started taking pills). Panel B reports results on mental and emotional health, including optimism, future-orientation, an index of emotional distress, happiness, and patients' overall satisfaction with their lives on the 1 to 10 scale. All indices are defined to be the equally weighted average of the z scores of their components, following Kling, Liebman, and Katz (2007). The full list of indices' components is indicated in the table's notes. Overall, we do not find any significant impact of the intervention on patient health.

This result is perhaps not surprising, given that the biometric tracking technology focuses on preventing defaults from treatment, a dramatic but relatively rare occurrence. Patients whose health was most likely to improve because of the intervention are the ones who would have interrupted their treatment otherwise. They constitute a small fraction of the overall sample, with default rates at 7.3 percent in the control group.

[Table V about here]

4.5 Impact on provider effort

We now examine whether the positive results identified on TB control outcomes can be explained by increased health worker effort. Table VI, Panel A shows results on health worker attendance, measured using two independent data collection methods: the observation days, during which an independent enumerator would record centers' opening and closing times and health worker presence (in addition

to keeping a log of visiting patients as analyzed in Section 4.2); and the random spot checks. The intervention did not affect whether or not and for how much time the centers were open (columns 1 to 4 and 11 and 12), reflecting the fact that these centers are often located in another facility operating at regular hours irrespective of TB-related activities (see Section 2.3). However, the biometric tracking devices did increase the presence of the health worker at the center. The controlled regressions show a statistically significant 4.5 and 7.1 percentage point increase in the likelihood that the health worker came to the center during observation days and random spot checks, respectively. These effects are positive but not significant in the specification without health worker controls (columns 5 and 6 and 13 and 14). With presence during an average observation day at a fairly high 83.4 percent in the control group, the marginal improvement is sizeable: absence was cut by about 27 percent in observation-day data and 18 percent in spot-check data. Furthermore, observation-day data allow us to compute rigorous attendance durations by subtracting breaks taken by the health worker during the day. The data point to significant improvements in presence time at the center, by about 20 minutes per day or 17 percent (columns 7 and 8). These results echo qualitative findings on health workers' perceptions of their work with the biometric devices. Most note that the installation of the devices increased accountability for them to stay longer at the center to make sure that all patients could come and scan their fingerprints, lest a missed-dose alert would be recorded in the system.

[Table VI about here]

The increase in health worker presence might also be explained by a more rigorous enforcement of work requirements by the NGO management in centers with biometric tracking. Columns 9 and 10 of Table VI, Panel A show that those centers receive significantly more visits from OpASHA staff. The additional visits may correspond to increased enforcement efforts by management, facilitated by the data on health worker and patient attendance that the organization receives from the biometric tracking devices, or to other activities related to the presence of the devices such as technical maintenance. To disentangle these contrasting interpretations, we take advantage of our high-frequency observation-day data and examine whether visits seem to be prompted by past dips in provider attendance or patient adherence. Table A4 in Appendix A shows the results. In centers equipped with a biometric tracking device, the likelihood of a management visit is significantly lower when patient adherence recorded over

the last four observation days is higher (columns 1 and 2). Although less significantly, the same result appears to be true for the extensive margin of health worker presence (columns 3 and 4) but not for the intensive margin (columns 5 and 6). These results suggest that one of the mechanisms by which the intervention improved health worker presence and patient adherence is that the biometric tracking devices facilitated the monitoring of health workers and that their managers acted on the information sent by the devices to OpASHA's server every day.

We now look into effects on the intensity of activities performed by health workers. Table VI, Panel B reports results based on a variety of measurement sources. Using observation-day data, columns 1 and 2 show that biometric tracking induces a 32 percent hike in the frequency of (independently verified) home visits by health workers to patients, an effect significant at the 10 percent level. The finding is corroborated by patient survey data: columns 13 and 14 show that patients assigned to biometric tracking report more visits by health workers to deliver pills at their doorstep.⁶ This result cannot be attributed to a variation in the proportion of mobile health workers across groups because the randomization of biometric equipment was stratified on this variable. This large increase in home visits comes in addition to higher attendance figures at the center, pointing to a sizeable improvement in overall health worker effort to ensure the adherence of patients to their treatment and follow up on missed doses.

The other dimensions of intensity of activities do not seem to be affected by the presence of biometric tracking devices. Health workers do not report intensifying activities aimed at preventing defaults other than increased home visits (columns 3 and 4) or decreasing detection activities (columns 5 and 6). The patient surveys confirm that health workers did not enhance or reduce the personalized service they provide in response to the intervention (columns 7 to 12).

5 Impact on data quality

5.1 Qualitative evidence on misreporting

Salary bonuses linked to performance indicators (here, the monthly number of detections) or general pressure from management may lead health workers to forge performance numbers. Our structured qual-

⁶This result is robust to restricting the sample to patients present at the outset of the experiment as shown in Appendix Table B3, columns 7 and 8.

itative interviews documented that misreporting can take many forms. Most of the qualitative evidence was offered by former or retired health workers, who were more open about this part of their experience than current employees, who generally offered more perfunctory responses.

Health workers report instances where patients were entirely fabricated. Although symptomatic patients are supposed to go to the Designated Microscopy Center themselves to provide sputum and get tested, before being enrolled in the DOTS system, health workers can offer to carry the sputum to the Center. In some instances, they used sputum from one infected patient to create one or more additional cases, which were then registered under fake names or using names of noninfected individuals in the community for greater realism. The ghost patients would then be assigned a TB number, and a box of pills would be delivered to the DOTS center. Health workers could also create fake patients without tinkering with sputum samples by colluding with the lab technician at the Designated Microscopy Center. Qualitative interviews revealed instances when several ground-level agents agreed on a mechanism for artificially inflating detection numbers and splitting the associated monetary incentives. There was reluctance by certain health workers to grant access to patients for survey purposes, a reaction that more forthcoming health workers (especially retired ones) interpret as reflecting the fear that some of their patients would be identified as fabricated.

While health workers had evidence to share on inflated detection numbers, misreporting default data is not as clearly documented in the qualitative interviews. Respondents tended to identify this issue as present in the official TB registers maintained at the hospitals, rather than at the health-worker level. In the regular reporting system without the biometric tracking devices, there is, however, scope for health workers to misreport defaults. It only takes a check in the patient's treatment card to report continued pill intake. Patients who interrupt their treatment can also be marked as having moved or having passed.

Biometric tracking may dramatically reduce the scope of opportunity to misreport data. By requiring fingerprint scans, it reduces the possibility that entirely fabricated patients are enrolled and that defaulting patients are shown as adhering to the treatment. The biometric system does not fully eradicate any possibility of forgery. Qualitative evidence revealed cases in which health workers managed to enroll fabricated patients through one of the methods described above, and then have a relative or neighbor register their fingerprint and scan it every few days over the course of the treatment. It is, however, a

much harder strategy to put in place and maintain over the duration of a full treatment course.

5.2 Measurement

Health workers report the number of detections and defaults for each DOTS center to OpASHA’s management every month. For month i and center j , the reported outcome $\tilde{Y}_{i,j}$ can be broken down as follows:

$$\tilde{Y}_{i,j} = Y_{i,j} + M_{i,j} \quad (3)$$

where $Y_{i,j}$ is the true outcome and $M_{i,j}$ corresponds to misreporting. Let \tilde{Y}_T be the average monthly outcome reported for centers in the treatment group and \tilde{Y}_C the average monthly outcome reported for centers in the control group. The treatment impact estimated on reported data can be decomposed as follows:

$$\tilde{Y}_T - \tilde{Y}_C = (Y_T - Y_C) + (M_T - M_C) \quad (4)$$

In equation 4, $(Y_T - Y_C)$ is the “true” impact, which we computed based on independent data and reported in the previous section, and $(M_T - M_C)$ is the difference in misreporting between centers equipped with biometric devices and control centers. Because the assignment of areas across treatment and control was random, any unintentional dimension of misreporting should be evenly distributed and cancel out in the difference. In addition, the outcomes of interest are little susceptible to random mistakes, as notifying a new patient or concealing a defaulting patient involves several steps over a long period of time. Since we cannot exclude that health workers still find ways to misreport in centers equipped with biometric devices, the quantity $(M_T - M_C)$ is a lower bound for the actual level of misreporting happening in control centers. Yet it should be close in absolute value to M_C , as M_T is expected to be small: as discussed above, the scope for misreporting is much reduced in biometric-equipped centers, where patients need to scan their fingerprints at the time of their detection and several times a week over the entire course of their treatment.

Equation 4 shows that the impact estimated on outcomes in program-generated data can be interpreted as the sum of the true impact and the difference in misreporting between centers in the treatment

and control groups. Therefore, to measure the impact of the intervention on misreporting, we measure the difference between true impact estimates based on outcome measures from our independent data (surveys, observation days, and spot checks) and estimates based on program and government data (center-wise reports, salary slips, and TB registers).

5.3 Impact on misreporting

Table VII reports impact estimates on default as computed using different data sources. Columns 1 and 2 show the negative and significant impact on the likelihood of default measured by independent survey data and correspond to the coefficients previously displayed in Table III (Panel A, columns 5 and 6). Columns 3 and 4 show the impact on defaults as measured in the center-wise data sheets prepared by health workers every month. This primary source of program data shows no significant difference in the fraction of defaults between the centers equipped with a biometric tracking device and those that are not. The contrast between these two findings indicates that health workers underreported defaults in centers without a biometric tracking device and that underreporting decreased by about one-fourth of defaulting patients in centers equipped with a device. Health workers in the control group were able to artificially lower the number of defaults in their reports, with the effect of concealing the actual decrease in default rates in the treatment group. This finding obtained from OpASHA program data holds in data from government TB registers, the official source of information for the TB Control Program: At the end of the reporting chain, patient-level outcomes reported in these registers reflect the underreporting of defaults by health workers without a biometric tracking device (columns 5 and 6).

[Table VII about here]

We use a seemingly unrelated regression framework to test the significance of the difference in point estimates between treatment effects on actual likelihood to default and on default as reported in the program data and in the government registers. Results are displayed at the bottom of Table VII. The tests reject coefficient equality when estimated on TB register data (columns 5 and 6). Although point estimates obtained with reported data (columns 3 and 4) are very similar to the ones using register data, the lower statistical power resulting from using center x month outcomes instead of individual-level data does not allow rejecting coefficient equality (the p-values are 0.111 and 0.118 in the regressions with and

without health worker controls, respectively).

We now turn to measuring effects on misreporting of detections. Table VIII shows the impact of the biometric tracking devices on the number of new cases reported. Columns 1 and 2 in Panel A report the reliable estimates already shown in Table IV, Panel A, which are based on program data cross-checked with patient surveys, and show that the biometric technology did not have a significant impact on the number of detections. Columns 3 and 4 use OpASHA data on the number of new patients per DOTS center x month compiled from the center-wise data sheets. The number of reported detections is significantly lower in biometric-equipped centers than in regular centers. With about 0.74 fewer notifications per month officially reported in equipped centers, and a null impact of the intervention on the number of “true” patients newly enrolled, the tracking technology appears to have curbed overreporting of detections by about one-fifth of all patients declared. Using seemingly unrelated regressions, we can reject coefficient equality at the 1 percent level, in the specification controlling only for strata fixed effects and baseline outcomes, and at the 5 percent level, in the specification including health worker controls.

[Table VIII about here]

Another way to identify fabricated patients is to look at the likelihood that the survey team could verify each reported patient across the treatment and control areas, a strategy similar to the independent auditing surveys conducted by Niehaus and Sukhtankar (2013) and Muralidharan, Niehaus, and Sukhtankar (2016) and the matching-based strategy used by Banerjee et al. (2016) to measure the number of ghost program beneficiaries. Panel B reports survey outcomes for the full sampling frame of patients attempted to be surveyed and enrolled after the beginning of the experiment. In the control group, enumerators were able to verify about 69 percent of all reported detections. This percentage reflects the difficulty of finding addresses and conducting survey operations in urban slums (which applies to areas equipped with a biometric tracking device too), but also the extent of overreporting detection figures, a factor that would be mostly cancelled out by the intervention. Patients were significantly more likely to be found and surveyed in treatment group areas and significantly less likely to refuse to answer (columns 1 and 2).⁷ These results are consistent with the forgery mechanism described in the qualitative evidence in which health

⁷In contrast, Appendix Table A5, Panel A shows no significant difference in the likelihood to verify patients reported in the program data and enrolled before the beginning of the experiment, as would be expected. If anything, we were slightly less likely to successfully verify treatment group patients enrolled before the beginning of the experiment.

workers enroll noninfected individuals as patients. Knowing that enumerators were visiting every one of their patients, health workers would have talked the fake ones into refusing the survey. It also appears quite consistent with the fact that certain health workers resisted survey operations or tried to undermine them over the course of the experiment. Another possible overreporting mechanism consists in making up individuals and registering them under fake identities and addresses. At the time of the survey, those ghost patients are not found by surveyors. The coefficient attached to the survey outcome "not found" is negative, even though nonsignificant, suggesting that the biometric tracking technology reduced scope for this other form of forgery as well.

Finally, Panel C of Table VIII confirms that the misreporting of patients carries over into official government registers. As OpASHA works by delegation of each Indian state to implement the national TB control program, the data it generates form an integral part of the public data. Patients from the treatment group areas found in the official TB registers were significantly more likely to be actual verified patients, suggesting that biometric tracking devices filter out fabricated patients that would otherwise be present in public health records.⁸

Taken together, these results provide robust evidence on the existence of forgery in health records at the level of the primary provider and in government registers. The installation of biometric tracking devices appears to significantly curb the level of misreporting and improve the truthfulness of health records. Accurate and timely data are a critical component of the state capacity to design and implement policy, identify needs and allocate resources accordingly, and respond to crises, making this result particularly important. In addition, this paper provides a direct illustration of the importance of data quality for policymaking: the results from this impact evaluation would have been quite misleading if we had only relied on the existing official data to measure outcomes. As shown in Tables VII and VIII, the study would have reached the conclusion that biometric tracking devices had no impact on the likelihood to default while reducing the number of patients notified by health workers—and may have been used to recommend abandoning the technology. The ability to make evidence-based policy decisions is a component of state capacity that relies heavily on accurate program data.

⁸Again, in contrast, Appendix Table A5, Panel B shows no significant difference in the likelihood that patients found in the TB registers and enrolled before the beginning of the experiment were verified patients.

6 Acceptability, sustainability, and scalability

Increased monitoring may be met with strong resistance by primary providers, to the point of not being sustainable (Banerjee, Duflo, and Glennerster, 2008; Dhaliwal and Hanna, 2017). This section presents evidence on the acceptability of the equipment by health workers and patients, and its sustainability in treatment centers.

Table IX, Panel A presents the impact of the intervention on key measures of health worker well-being. We showed in previous sections that the intervention resulted in fewer overreported detections, even though the actual number of detections remained stable. Monthly salary data obtained from OpASHA confirm that the monetary compensation associated with detections was reduced accordingly: columns 1 and 2 show that the compensation for detection activities significantly decreased, resulting in a 7 percent cut in the total monthly salary of health workers using biometric devices compared to the control group (columns 3 and 4). In addition to a reduced salary, qualitative evidence shows that treated health workers bemoan the loss of discretionary power, e.g., the fact that they can no longer decide to provide several doses of medicine at the same time or give the drugs to a family member for convenience.

However, health workers equipped with biometric devices also enjoy a significantly reduced workload. In the control group, 22 percent of health workers report an excessive workload. The installation of biometric tracking devices reduces this share by half, or even virtually removes this concern—depending on the specification (columns 9 and 10). This reflects the fact that the devices make it easier for health workers to do their work on default prevention. The coefficient attached to the number of challenges faced in ensuring treatment adherence is negative, although non-significant (columns 5 and 6). 82 percent of health workers who worked with biometric tracking found the technology useful or very useful, and 77 percent believe that OpASHA should scale it up. Qualitative evidence shows that health workers are satisfied with having less reporting responsibility since their activity record is entirely accessible by management directly from the server. They also appreciate having all patient details immediately available, including their adherence record, treatment phase, and corresponding dosage. They acknowledge that automatic alerts make the tracking of patients easier, but also regret that scanning fingerprints adds to the number of steps involved in distributing drugs and forces them to stay longer hours at the center to make sure all scheduled patients have had a chance to visit.

[Table IX about here]

Overall, despite imposing additional constraints on both health workers and patients, the installation of biometric tracking devices does not seem to impact negatively the satisfaction level of either. On the health worker side, we do not see any significant impact on their job satisfaction—if anything, the point estimates on a standardized job satisfaction index have a positive sign (columns 11 and 12). This result might reflect the fact that although the job offers less scope for fraudulently seeking compensation, it is overall easier and more rewarding—an idea consistent with the very similar levels of turnover observed in the treatment and control groups (Table A1 in Appendix A). On the patient side, Table IX, Panel B shows that satisfaction levels are remarkably similar across treatment and control groups. Patients across groups are equally satisfied with healthcare received and equally likely to recommend OpASHA to other symptomatic individuals (columns 1 to 4). Consistently, according to health workers, patients across treatment and control groups face a similar number of challenges to complete their treatment (columns 5 and 6). This may result from two opposite impacts: patients who are receiving more frequent home visits may be more satisfied, while those under increased pressure to come to the center in person may be less satisfied. Among patients exposed to the biometric devices, 62.8 percent report a positive or very positive impression of the technology.

The acceptability of biometric tracking devices may account for the durability of the equipment in the centers. The information collected on observation days includes whether the equipment (fingerprint reader, laptop, and USB key) is present in the center and working properly. Taking advantage of the duration of the experiment, we can measure the evolution of technology presence over time. Table IX, Panel C shows that in the first month of the experiment, 92 percent of DOTS centers in the treatment group did have biometric equipment present on the day of observation. Columns 1 and 2 report how each passing week affects this likelihood and show that the presence of technology is remarkably stable over time. The remaining columns suggest that there was a slight reduction in the presence of certain peripheral components over time, but the graphic representation of the evolution (Figure A1 in Appendix A) shows a stabilization at high levels after six months.

Not only is the technology present and functioning over time, the impact on the main outcomes of interest appear to be maintained. If the impact of the technology was driven by a novelty factor, or

if health workers had found ways to circumvent the technology over time, then the positive effects on TB control would only manifest themselves in the early weeks or months. Table A6 in Appendix A shows point estimates on the interaction between treatment and time, in addition to the direct impact of the treatment itself. The time variable is obtained by computing the number of days lapsed between the installation of the biometric device and each patient's treatment start date (for outcomes measured at the patient level) or between the installation of the biometric device and the observation day (for outcomes measured at the patient x observation day level). Coefficients associated to the interacted term are consistently not distinguishable from zero, except an estimate significant at the 10 percent level in one regression out of 22, which provides reassurance that the impact is sustained over time.

Finally, it is worth emphasizing that portable devices were successfully used by mobile health workers, who are included in the sample even though their number was too small for us to conduct separate impact analysis. Mobile health workers' level of satisfaction with the technology was close to the average among treated health workers: 86 percent of those who worked with biometric tracking found the technology useful or very useful, and 71 percent believe that OpASHA should scale it up. This suggests that the technology is not restricted to brick-and-mortar facilities but can also be used in remote settings, rural areas, and other hard-to-reach places.

These findings point to potential significant benefits of scaling up the eCompliance technology evaluated in this paper to improve TB care provision. In 2012, the Government of India launched a web-based centralized database of all TB patients called Nikshay, based on the digitization of treatment cards. By aggregating TB records from across India in one unique dataset, Nikshay serves as a dashboard for the National TB Control Program. Yet, as it relies on the digitization of paper-based documents, Nikshay is not set up for identifying duplicates and establishing unique identity. The quality of the database is entirely reliant on the quality of information on the treatment cards, and the digitization process may even add a layer of potential error in the recording system. Further, the digitization process is implemented with a lag of about one to two months between when the patient misses a dose and when the corresponding information gets entered in the electronic database, thus disabling any quick corrective action to ensure patient adherence to treatment.

On the other hand, biometric identification technology is increasingly available in India, especially

since the national roll-out of Aadhaar, the Government of India's ambitious program to provide a biometrics-based Unique Identification (UID) to all residents of India. Aadhaar is an individual identification number which can be authenticated through fingerprint scan, iris scan, and registered cellphone number and serves as a proof of identity and address, anywhere in India. It currently covers about 1.2 billion individuals, corresponding to 90 percent of the adult population. The objective of the Government of India is to have various social programs use it as a platform for service delivery.

With Nikshay and Aadhaar, the infrastructure is in place for scaling up biometric recognition technology linked in real time to an integrated database of TB care beneficiaries. Our results show that this could massively improve the reliability of records and enable rapid follow-up on patients who miss doses. In addition, it would give the government the information required to make rapid-cycle adjustments to program implementation and hold private provider partners accountable for results, and it would facilitate the implementation of financial incentive schemes or direct benefits transfers to both providers and patients (RNTCP, 2017).

7 Conclusion

This paper provides original evidence on the benefits of using biometric tracking technology to strengthen frontline service delivery and improve the reliability of government data. We exploit the random placement of the technology across four Indian states and use a large variety of data sources to obtain a detailed understanding of impacts and mechanisms and to uncover the extent of misreporting in TB control.

Results point to a sharp reduction in treatment interruptions, greater medication adherence, and a greater compliance to the requirements of the DOTS protocol defined by the WHO. These results are accounted for by a higher level of attendance by health workers and greater effort dedicated to following up with defaulting patients through more frequent home visits. The biometric technology generated these effects by facilitating both the tracking of patients' pill-taking, by health workers, and the monitoring of health workers, by the NGO management. Potential downsides did not materialize: we do not see a reallocation of health worker efforts away from other critical activities, such as early case finding, or a strategic refocusing of care toward more compliant patients. In addition to these impacts on direct TB

control outcomes, the biometric technology dramatically reduces the scope for forging health registers. Triangulating different data sources, we show that the intervention substantially reduced underreporting of treatment interruptions and overreporting of new cases both in program data and in the government registers. The fact that the technology is a productivity-enhancing tool helps explain why it appears well-accepted by both patients and health workers—despite imposing more demands on both and generating a loss in earnings for the latter—and why its impacts are sustained over time. Overall, biometric tracking appears amenable for scale-up through India’s TB Control Program.

The technology studied here measurably improved state capacity at two levels. First, it increased the capacity of a public program to efficiently extend its reach by addressing the principal-agent problem inherent to the decentralization of service provision. It enhanced the quality of implementation of a delivery model based on partnerships with private-sector actors in charge of establishing highly localized points of care. This resulted in a large impact on the likelihood that isolated, vulnerable people receive adequate treatment against a deadly disease. Second, the technology improved the accuracy of public-health data, which increases the capacity of the state to identify needs and allocate resources accordingly. Accurate data pave the way to evidence-driven decisions: in our case, official data affected by misreporting issues would have masked the positive effects of the program and likely led to its abandonment.

In order to enhance state capacity in a durable way, changes need to be accepted and sustainable. The technology we study holds health workers more accountable for their performance while at the same time improving their productivity. This result may hold lessons for strategies to improve state capacity beyond the important case of TB control. Sustainable improvement in the commitment of delivery agents may be facilitated by a simultaneous effort to improve their work conditions and build the state’s capacity. Technology may improve state capacity, not only as a tool to deliver services by means both smarter and faster but also through a positive impact on frontline workers, if they are willing to accept greater scrutiny in exchange for a simplification of their daily tasks.

References

Alchian, Armen A., and Harold Demsetz. "Production, Information Costs, and Economic Organization." *American Economic Review*, 1972, 62 (5): 777–795.

Banerjee, Abhijit, Angus Deaton, and Esther Duflo. "Wealth, Health, and Health Services in Rural Rajasthan." *American Economic Review*, 2004, 94 (2): 326–330.

—, and Esther Duflo. "Addressing Absence." *Journal of Economic Perspectives*, 2006, 20 (1): 117–132.

—, Esther Duflo, and Rachel Glennerster. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association*, 2008, 6 (2–3): 487–500.

— Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy*, 2010, 2 (1):1–30.

—, Esther Duflo, Clement Imbert, Santhosh Mathew, and Rohini Pande. "E-Governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India." NBER Working Paper Series, No. 22803 (2016).

—, Rema Hanna, Jordan Kyle, Benjamin A. Olken, and Sudarno Sumarto. "Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia." *Journal of Political Economy*, 2018, 126 (2): 451–491.

Barnwal, Prabhat. "Curbing Leakage in Public Programs with Direct Benefit Transfers Evidence from India's Fuel Subsidies and Black Markets." Working Paper (2016).

Becker, Gary S., and George J. Stigler. "Law Enforcement, Malfeasance, and Compensation of Enforcers." *Journal of Legal Studies*, 1974, 3 (1): 1–18.

Besley, Timothy, and Torsten Persson. "State Capacity, Conflict, and Development." *Econometrica*, 2010, 78 (1): 1–34.

Bhatnagar, Nupur, Abhishek Sinha, Navkar Samdaria, Aakar Gupta, Shelly Batra, Manish Bhardwaj, and William Thies. "Biometric Monitoring as a Persuasive Technology: Ensuring Patients Visit Health Centers in India's Slums." In Hutchinson, David, Takeo Kanade, and Josef Kittler, *Persuasive*

Technology: Design for Health and Safety (169–180). (Berlin, Germany: Springer, 2012).

Bhattacharya, Kakali. *Fundamentals of Qualitative Research. A Practical Guide*. (Abingdon-on-Thames, UK: Routledge, 2017).

Björkman, Martina, and Jakob Svensson. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *Quarterly Journal of Economics*, 2009, 124 (2): 735–769.

Brambor, Thomas, Agustín Goenaga, Johannes Lindvall, and Jan Teorell. "The Lay of the Land: Information Capacity and the Modern State." *Comparative Political Studies*, 2019.

Callen, Michael, and James D. Long. "Institutional Corruption and Election Fraud. Evidence from a Field Experiment in Afghanistan." *American Economic Review*, 2015, 105 (1): 354–381.

—, Saad Gulzar, Ali Hasanain, and Muhammad Yasir Khan. "The Political Economy of Public Sector Absence: Experimental Evidence from Pakistan." NBER Working Paper, No. 22340 (2018).

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics*, 2008, 90 (3): 414–427.

Chaudhury, Nazmul, and Jeffrey Hammer. "Ghost Doctors: Absenteeism in Rural Bangladeshi Health Facilities." *World Bank Economic Review*, 2004, 18 (3): 423–441.

Cohen, Jessica, and Indrani Saran. "The Impact of Packaging and Messaging on Adherence to Malaria Treatment: Evidence from a Randomized Controlled Trial in Uganda." *Journal of Development Economics*, 2018, 134: 68–95.

Das, Jishnu, and Jeffrey Hammer. "Money for Nothing: The Dire Straits of Medical Practice in Delhi, India." *Journal of Development Economics*, 2007, 83 (1): 1–36.

—, Ada Kwan, Benjamin Daniels, Srinath Satyanarayana, Ramnath Subbaraman, Sofi Bergkvist, Ranendra K. Das, Veena Das, and Madhukar Pai. "Use of Standardised Patients to Assess Quality of Tuberculosis Care: A Pilot, Cross-Sectional Study." *Lancet Infectious Diseases*, 2015, 15 (11): 1305–1313.

DeFulio, Anthony, and Kenneth Silverman. "The Use of Incentives to Reinforce Medication Adherence." *Preventive Medicine*, 2012, 55 (S1): 86–94.

Desrosières, Alain. *The Politics of Large Numbers: A History of Statistical Reasoning*. (Cambridge, MA: Harvard University Press, 2002).

Dhaliwal, Iqbal, and Rema Hanna. "The Devil Is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics*, 2017, 124: 1–21.

Di Tella, Rafael, and Ernesto Schargrotsky. "The Role of Wages and Auditing During a Crackdown on Corruption in the City of Buenos Aires." *Journal of Law and Economics*, 2003, 46 (1): 269–292.

Downey, Laura, Neethi Rao, Lorna Guinness, Miqdad Asaria, Shankar Prinja, Anju Sinha, Rajni Kant, Arvind Pandey, François Cluzeau, and Kalipso Chalkidou. "Identification of Publicly Available Data Sources to Inform the Conduct of Health Technology Assessment in India." *F1000Research*, 2018, 7 (245).

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. "Incentives Work: Getting Teachers to Come to School." *American Economic Review*, 2012, 102 (4): 1241–1278.

—, Michael Greenstone, Rohini Pande, and Nicholas Ryan. "Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *Quarterly Journal of Economics*, 2013, 128 (4): 1499–1545.

Esarey, Justin, and Andrew Menger. "Practical and Effective Approaches to Dealing With Clustered Data." *Political Science Research and Methods*, 2019, 7 (3): 541–559.

Ferraz, Claudio, and Frederico Finan. "Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics*, 2008, 123 (2): 703–745.

Finan, Frederico, Benjamin Olken, and Rohini Pande. "The Personnel Economics of the Developing State." In Banerjee, Abhijit, and Esther Duflo, *Handbook of Field Experiments* (467–513). (Amsterdam, Netherlands: North Holland, 2017).

Gelb, Alan, and Julia Clark. "Identification for Development: The Biometrics Revolution." Center for Global Development Working Paper, No. 315 (2013).

Giné, Xavier, Jessica Goldberg, and Dean Yang. "Credit Market Consequences of Improved Personal Identification: Field Experimental Evidence from Malawi." *American Economic Review*, 2012, 102 (6): 2923–2954.

Giuffrida, Antonio, and David J. Torgerson. "Should We Pay the Patient? Review of Financial Incentives to Enhance Patient Compliance." *British Medical Journal*, 1997, 315 (7110): 703–707.

Goldberg, Jessica, Mario Macis, and Pradeep Chintagunta. "Incentivized Peer Referrals for Tubercu-

losis Screening: Evidence from India." NBER Working Paper Series, No. 25279 (2019).

Henderson, Vernon, Adam Storeygard, and David N. Weil. "A Bright Idea for Measuring Economic Growth." *American Economic Review*, 2011, 101 (3): 194–199.

—, Paul Milgrom. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law Economics & Organization*, 1991, 7: 24–52.

Huddart, Sophia, Thomas Bossuroy, Vincent Pons, Siddharta Baral, Madhukar Pai, and Clara Delavalade. "Knowledge about Tuberculosis and Infection Prevention Behavior: A Nine City Longitudinal Study from India." *PLoS One*, 2018, 13 (10): E0206245.

Ichino, Nahomi, and Matthias Schündeln. "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana." *Journal of Politics*, 2012, 74 (1): 292–307.

Jacob, Brian, and Steven Levitt. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics*, 2003, 118 (3): 843–877.

Kleven, Henrik J., Martin B. Knudsen, Claus T. Kreiner, Søren Pedersen, and Emmanuel Saez. "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica*, 2011, 79 (3): 651–692.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 2007, 75 (1): 83–119.

Koehler-Derrick, Gabriel. "The Informational Origins and Consequences of Colonial Governance." Working Paper, 2019.

Laffont, Jean-Jacques, and David Martimort. *The Theory of Incentives: The Principal-Agent Model*. (Princeton, NJ: Princeton University Press, 2002).

Lee, Melissa M., and Nan Zhang. "Legibility and the Informational Foundations of State Capacity." *The Journal of Politics*, 2017, 79 (1): 118–132.

Lester, Richard T., Paul Ritvo, Edward J. Mills, Antony Kariri, Sarah Karanja, Michael H. Chung, William Jack et al., and Francis A. Plummer. "Effects of a Mobile Pphone Short Message Service on Antiretroviral Treatment Adherence in Kenya (WelTel Kenya1): A Randomised Trial." *Lancet Infectious Diseases*, 2010, 376 (9755): 1838–1845.

Mahapatra, Prasanta, Kenji Shibuya, Alan D. Lopez, Francesca Coullare, Francis C. Notzon, Chalapati Rao, and Simon Szreter. "Civil registration systems and vital statistics: successes and missed opportunities." *The Lancet*, 2007, 370 (9599): 1653–1663.

Martinez, Luis R. "How Much Should We Trust the Dictator's GDP Growth Estimates?" Working Paper (2019).

McDonald, Heather P., Amit X. Garg, and R. Brian Haynes. "Interventions to Enhance Patient Adherence to Medication Prescriptions: Scientific Review." *JAMA*, 2002, 288 (22): 2868–2879.

McMillen, Christian W. *Discovering Tuberculosis: A Global History, 1900 to the Present*. (New Haven, CT: Yale University Press, 2015).

Mikkelsen, Lene, David E. Phillips, Carla AbouZahr, Philip W. Setel, Don de Savigny, Rafael Lozano, and Alan D. Lopez. "A Global Assessment of Civil Registration and Vital Statistics Systems: Monitoring Data Quality and Progress." *The Lancet*, 2015, 386 (10001): 1395–1406.

Mohanam, Manoj, Katherine Hay, and Nachiket Mor. "Quality Of Health Care in India: Challenges, Priorities, and the Road Ahead." *Health Affairs (Project Hope)*, 2016, 35 (10): 1753–1758.

Morse, Janice M. "The Implications of Interview Type and Structure in Mixed-Method Design." In Jaber F. Gubrium, James A. Holstein, Amir B. Marvasti, and Karyn D. McKinney, *The SAGE Handbook of Interview Research: The Complexity of the Craft* (193–206). (2014)

Hutchinson, David, Takeo Kanade, and Josef Kittler, *Persuasive Technology: Design for Health and Safety* (169–180). (Berlin, Germany: Springer, 2012).

Morton, Matthew, Somil Nagpal, Rajeev Sadanandan, and Sebastian Bauhoff. "India's Largest Hospital Insurance Program Faces Challenges in Using Claims Data to Measure Quality." *Health Affairs (Project Hope)*, 2016, 35 (10): 1792–1799.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review*, 2016, 106 (10): 2895–2929.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *American Economic Review*, 2002, 92 (4): 850–873.

Niehaus, Paul, and Sandip Sukhtankar. "Corruption Dynamics: The Golden Goose Effect." *American*

Economic Journal: Economic Policy, 2013, 5 (4): 230–269.

Nieuwlaat, Robby, Nancy Wilczynski, Tamara Navarro, Nicholas Hobson, Rebecca Jeffery, Arun Keepanasseril et al., and R. Brian Haynes. "Interventions for Enhancing Medication Adherence." *Cochrane Database of Systematic Reviews*, 11 (2014): CD000011.

Olken, Benjamin A. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*, 2007, 115 (2): 200–249.

Oomman, Nandini, Garrett Mehl, Matt Berg, and Rachel Silverman. "Modernising vital registration systems: Why now?" *The Lancet*, 2013, 381 (9875): 1336–1337.

Organization for Economic Cooperation and Development. *Strengthening Health Information Infrastructure for Health Care Quality Governance: Good Practices, New Opportunities, and Data Privacy Protection Challenges*. (Paris: OECD, 2013).

Otto, Kate, Meera Shekar, Christopher Herbst, and Rianna Mohammed. "Information and Communication Technologies for Health Systems Strengthening; Opportunities, Criteria for Success, and Innovation for Africa and Beyond." Health, Nutrition, and Population Discussion Paper No. 94943 (World Bank, 2015).

Page, Lucy, and Rohini Pande. "Ending Global Poverty: Why Money Isn't Enough." *Journal of Economic Perspectives*, 2018, 32 (4): 173–200.

Peabody, John W., Jeff Luck, Sharad Jain, Dan Bertenthal, and Peter Glassman. "Assessing the Accuracy of Administrative Data in Health Information Systems." *Medical Care*, 2004, 42 (11):1066–1072.

Phillips, Beth S., Shreya Singhal, Shambhavi Mishra, Fnu Kajal, Sun Yu Cotter, and May Sudhinaraset. "Evaluating Concordance between Government Administrative Data and Externally Collected Data among High-Volume Government Health Facilities in Uttar Pradesh, India." *Global Health Action*, 2019, 12 (1).

Pop-Eleches, Cristian, Harsha Thirumurthy, James P. Habyarimana, Joshua S. G. Zivin, Markus Goldstein, Damien de Walque, Leslie D. Mackeen, Jessica Haberer, Sylvester N. Kimaiyo, John E. Sidle, Duncan Ngare, and David R. Bangsberg. "Mobile Phone Technologies Improve Adherence to Antiretroviral Treatment in a Resource-Limited Setting: A Randomized Controlled Trial of Text Message Reminders." *AIDS*, 2011, 25 (6): 825–834.

Raffler, Pia, Daniel N. Posner and Doug Parkerson. "The Weakness of Bottom-Up Accountability: Experimental Evidence from the Ugandan Health Sector." Working paper (2019).

Raifman, Julia R. G., Heather E. Lanthorn, Slawa Rokicki, and Günther Fink. "The Impact of Text Message Reminders on Adherence to Antimalarial Treatment in Northern Ghana: A Randomized Trial." *PLoS One*, 2014, 9 (10): E109032.

Rasul, Imran, and Daniel Rogger. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *Economic Journal*, 2016, 128 (608): 413–446.

Revised National Tuberculosis Control Programme. "National Strategic Plan for Tuberculosis Elimination 2017–2025." Central TB Division, Directorate General of Health Services, Ministry of Health with Family Welfare, Government of India (2017).

Sabate, Eduardo. *Adherence to Long-Term Therapies: Evidence for Action*. (Geneva, Switzerland: World Health Organization, 2003).

Satyanarayana, Srinath, Ramnath Subbaraman, Priya Shete, Genevieve Gore, Jishnu Das, Adithya Cattamanchi et al., and Madhukar Pai.. "Quality of Tuberculosis Care in India: A Systematic Review." *International Journal of Tuberculosis and Lung Disease*, 2015, 19 (7): 751–763.

Scott, James C. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. (New Haven, CT: Yale University Press, 1998).

Slemrod, Joel, Marsha Blumenthal, and Charles Christian. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics*, 2001, 79 (3): 455–483.

Snidal, Sarah J., Genevieve Barnard, Emmanuel Atuhairwe, and Yanis B. Amor. "Use of eCompliance, an Innovative Biometric System for Monitoring of Tuberculosis Treatment in Rural Uganda." *American Journal of Tropical Medicine and Hygiene*, 2015, 92 (6): 1271–1279.

Sretzer, Simon. "The Right of Registration: Development, Identity Registration, and Social Security—A Historical Perspective." *World Development*, 2007, 35 (1): 67–86.

Tooze, J. Adam. *Statistics and the German State, 1900–1945: The Making of Modern Economic Knowledge*. (Cambridge, UK; New York, NY: Cambridge University Press, 2001).

Van Den Berg, Gerard J., and Bas Van Der Klaauw. "Counseling and Monitoring of Unemployed

Workers: Theory and Evidence from a Controlled Social Experiment." *International Economic Review*, 2006, 47 (3): 895–936.

World Health Organization. "Global Strategy and Targets for Tuberculosis Prevention, Care and Control after 2015" (2014).

World Health Organization. "Global Tuberculosis Report 2018" (2018).

World Health Organization and World Bank. *Tracking Universal Health Coverage: 2017 Global Monitoring Report*. (Washington, DC: World Health Organization, 2017).

Zuniga, Alejandro E.F., Khin Than Win, and Willy Susilo. "Biometrics for Electronic Health Records." *Journal of Medical Systems*, 2010, 34(5): 975–983.

Zurovac, Dejan, Raymond K. Sudoi, Willie S. Akhwale, Moses Ndiritu, Davidson H. Hamer, Alexander K. Rowe, and Robert W. Snow. "The Effect of Mobile Phone Text-Message Reminders on Kenyan Health Workers' Adherence to Malaria Treatment Guidelines: A Cluster Randomised Trial." *Lancet Infectious Diseases*, 2011, 378 (9793): 795–803.

Figure 1. States and Cities in the Sample

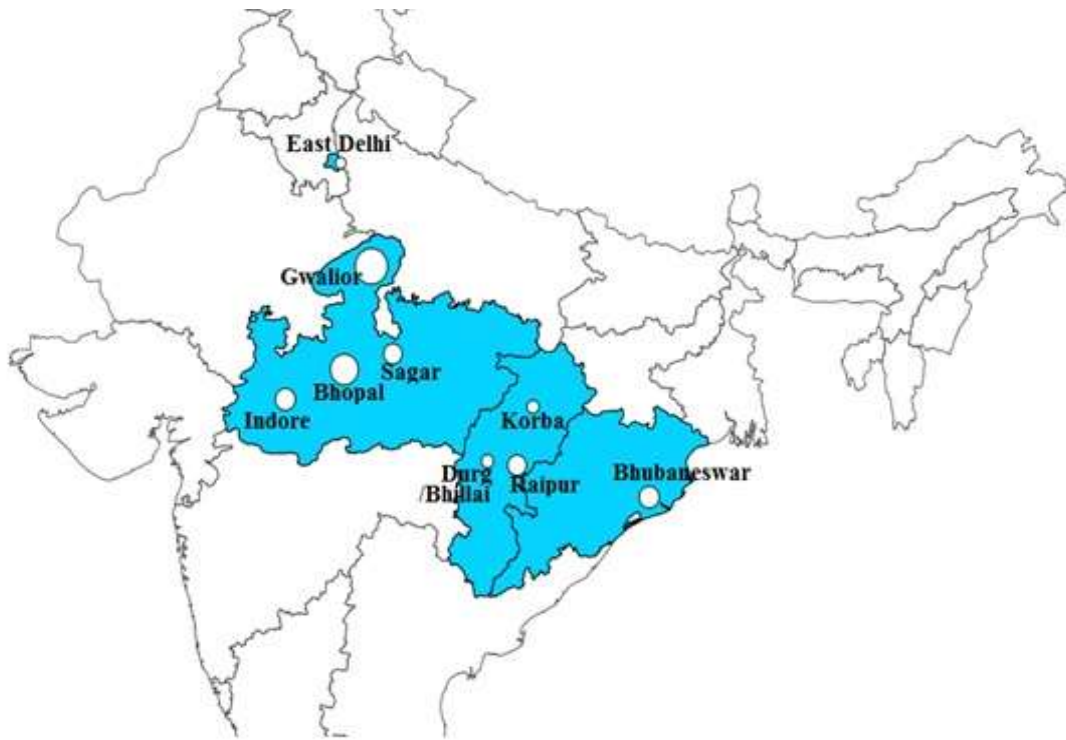


Table I: Health Worker Summary Statistics

	Full sample		Number of obs.	Control group		Treatment group		P-value Treatment = Control (8)
	Mean (1)	SD (2)		Mean (4)	SD (5)	Mean (6)	SD (7)	
<i>Panel A. Health workers' characteristics</i>								
Male	0.741	0.441	85	0.794	0.410	0.710	0.461	0.382
Age	30.9	7.3	85	33.4	7.8	30.4	6.1	0.097
Caste								
General caste	0.400	0.493	85	0.265	0.448	0.516	0.508	0.059
Other Backward Class	0.365	0.484	85	0.382	0.493	0.323	0.475	0.696
Scheduled Caste	0.106	0.310	85	0.118	0.327	0.097	0.301	0.832
Scheduled Tribe	0.094	0.294	85	0.206	0.410	0.032	0.180	0.035
Minority	0.035	0.186	85	0.029	0.171	0.032	0.180	0.945
Religion								
Hindu	0.812	0.393	85	0.765	0.431	0.903	0.301	0.148
Muslim	0.129	0.338	85	0.176	0.387	0.097	0.301	0.412
Other	0.059	0.237	85	0.059	0.239	0.000	0.000	0.075
Highest education level achieved								
Class 12 and below	0.634	0.485	82	0.688	0.471	0.633	0.490	0.684
Tertiary	0.341	0.477	82	0.281	0.457	0.333	0.479	0.671
Other diploma/non-formal education	0.024	0.155	82	0.031	0.177	0.033	0.183	1.000
Work experience								
Any previous work experience	0.753	0.434	85	0.853	0.359	0.645	0.486	0.038
Number of years of work experience	8.7	5.1	81	9.2	5.7	8.7	4.7	0.767
Any previous working in a job related to TB	0.153	0.362	85	0.147	0.359	0.161	0.374	0.910
Any previous experience in the social/NGO sector	0.159	0.367	82	0.206	0.410	0.138	0.351	0.549
Any other income generating activity, in addition to OpASHA	0.083	0.278	84	0.088	0.288	0.129	0.341	0.750
Lives in one of the areas she covers	0.566	0.499	83	0.545	0.506	0.742	0.445	0.080
Household size	5.6	2.9	83	5.2	2.7	6.2	3.4	0.266
Lives alone	0.036	0.187	84	0.000	0.000	0.067	0.254	0.168
Assets								
Has electricity	0.976	0.154	83	1.000	0.000	0.933	0.254	0.104
Has tap water	0.561	0.499	82	0.594	0.499	0.400	0.498	0.112
Has a television	0.904	0.297	83	0.848	0.364	0.933	0.254	0.222
Has a refrigerator	0.542	0.501	83	0.545	0.506	0.467	0.507	0.530
Rents an apartment or house to a third party	0.146	0.356	82	0.118	0.327	0.207	0.412	0.232
Owns her house	0.600	0.493	85	0.559	0.504	0.581	0.502	0.951
Exposure to technology								
Knows how to use a computer	0.565	0.499	85	0.559	0.504	0.516	0.508	0.725
Knows how to use the internet	0.470	0.502	83	0.424	0.502	0.433	0.504	0.962
Has an email account	0.400	0.493	85	0.412	0.500	0.258	0.445	0.168
Has a social networking account	0.329	0.473	85	0.294	0.462	0.258	0.445	0.663
Days spent in the experiment	308.9	141.0	85	349.1	127.7	333.7	137.7	0.655
<i>Panel B. OpASHA center-wise data sheets</i>								
Reported baseline number of detections per month per area	2.692	1.394	65	2.965	1.562	2.393	1.135	0.061
Reported baseline fraction of defaults per month per area	0.034	0.062	65	0.029	0.050	0.039	0.074	0.434
<i>Panel C. Observation days</i>								
Center open	0.984	0.047	57	0.980	0.053	0.989	0.039	0.535
Total time center was open	309.5	183.1	57	307.4	183.5	311.8	186.2	0.966
Health worker present at the center	0.788	0.275	61	0.769	0.264	0.810	0.289	0.520
Total time health worker was present at the center, net of breaks	108.6	60.8	61	110.9	65.5	106.0	56.1	0.714
Visit by OpAsha	0.055	0.114	65	0.067	0.126	0.043	0.100	0.310
Adherence 1: Patient came to center and took or picked up pill	0.519	0.258	65	0.520	0.265	0.517	0.255	0.963
Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit	0.759	0.240	65	0.761	0.238	0.758	0.247	0.843

Notes : In Panel A, the unit of observation is the health worker. For each variable, we report the mean, standard deviation, and number of observations in the full sample (col. 1–3), including the replacement health workers. The means and standard deviations in the control and treatment groups are calculated based on the subsample of health workers present at the beginning of the experiment (col. 4–7). We also report the p-value of the difference with control for strata fixed effects and clustered standard errors (col. 8).

In Panels B and C, the unit of observation is the cluster of areas allocated to a health worker. Baseline detection and default outcomes are computed based on six months of OpASHA center-wise data sheets prior to the experiment. Baseline observation days outcomes are computed based on observations days that took place before the beginning of the experiment.

Table II: Patient Summary Statistics

	Full sample		Number of obs.	Control group		Treatment group		P-value Treatment = Control
	Mean (1)	SD (2)		Mean (4)	SD (5)	Mean (6)	SD (7)	
Male	0.574	0.495	3387	0.535	0.499	0.576	0.495	0.513
Age	33.9	16.4	3385	34.0	16.8	34.3	16.6	0.789
Caste								
Does not know	0.047	0.211	3388	0.040	0.196	0.026	0.159	0.523
General caste	0.176	0.381	3388	0.160	0.367	0.165	0.371	0.566
Other Backward Class	0.357	0.479	3388	0.376	0.485	0.352	0.478	0.621
Scheduled Caste	0.249	0.432	3388	0.229	0.421	0.297	0.458	0.089
Scheduled Tribe	0.083	0.276	3388	0.118	0.323	0.103	0.305	0.158
Minority	0.088	0.283	3388	0.078	0.268	0.058	0.234	0.539
Religion								
Hindu	0.856	0.351	3377	0.869	0.337	0.903	0.297	0.543
Muslim	0.126	0.332	3377	0.112	0.315	0.081	0.274	0.629
Other	0.017	0.131	3377	0.019	0.136	0.016	0.127	0.509
Literacy								
Cannot read or write	0.296	0.457	3390	0.271	0.445	0.301	0.459	0.434
Can read but not write	0.038	0.192	3390	0.021	0.143	0.026	0.159	0.896
Can read and write	0.666	0.472	3390	0.708	0.455	0.673	0.470	0.394
Education								
Pre-primary	0.234	0.424	3381	0.217	0.412	0.234	0.424	0.769
Primary	0.387	0.487	3381	0.375	0.485	0.390	0.488	0.530
Secondary	0.335	0.472	3381	0.367	0.482	0.334	0.473	0.468
Undergraduate and more	0.044	0.205	3381	0.042	0.200	0.042	0.201	0.911
Household size	5.5	2.4	3382	5.6	2.5	5.4	2.3	0.167
Lives alone	0.020	0.141	3392	0.019	0.136	0.016	0.126	0.500
Assets								
Has electricity	0.965	0.184	3309	0.968	0.176	0.935	0.248	0.037
Has tap water	0.537	0.499	3324	0.520	0.500	0.418	0.494	0.465
Has a television	0.822	0.382	3205	0.826	0.379	0.832	0.375	0.464
Has a refrigerator	0.260	0.439	3206	0.276	0.448	0.230	0.422	0.303
Owns her house	0.694	0.461	3315	0.716	0.451	0.738	0.441	0.975
Migration status								
Has always lived in the area	0.502	0.500	3389	0.570	0.496	0.542	0.499	0.305
Has lived there for more than 6 years	0.241	0.428	3389	0.209	0.407	0.258	0.438	0.026
Has lived there for less than 5 years	0.256	0.437	3389	0.221	0.416	0.200	0.401	0.439
Is currently working	0.335	0.472	3117	0.381	0.486	0.406	0.492	0.794
Time to go to the center (in minutes)	11.5	8.3	2429	11.6	8.7	12.4	9.7	0.839
Distance to the center (in miles)	3.0	4.9	2646	3.3	5.4	3.0	4.8	0.727

Notes : The unit of observation is a verified patient. For patients who completed more than one survey, we use data from the first survey completed. For each variable, we first report the mean, standard deviation, and number of observations in the full sample including all patients, whether enrolled before or after the beginning of the experiment (col. 1–3). The means and standard deviations in the control and treatment groups are calculated based on the subsample of patients enrolled before the beginning of the experiment (col. 4–7). We also report the p-value of the difference with control for strata fixed effects and a dummy indicating whether the information comes from an entry or exit survey, and clustered standard errors (col. 8).

Table III: Impact on Treatment Interruption and Adherence

<i>Panel A: Impact on treatment interruption, or default (source: patient surveys)</i>										
	Treatment complete		Treatment ongoing		Default		Died		Outcome unknown	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.003	-0.007	0.022	0.020	-0.018	-0.015	-0.001	0.000	0.000	0.003
	(0.028)	(0.026)	(0.025)	(0.025)	(0.007)**	(0.007)**	(0.005)	(0.005)	(0.012)	(0.011)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes
Observations	2802	2802	2802	2802	2802	2802	2802	2802	2802	2802
R-squared	0.055	0.072	0.064	0.074	0.020	0.033	0.014	0.024	0.054	0.105
Mean in control group	0.507	0.507	0.299	0.299	0.073	0.073	0.016	0.016	0.105	0.105

<i>Panel B: Impact on treatment adherence (source: observation days)</i>								
	Adherence 1: Patient came to the center and took or picked up the pill		Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit		Compliance with treatment verified (as per adherence 1)		Compliance with treatment verified (as per adherence 2)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.139	0.122	0.073	0.056	0.039	0.038	0.024	0.023
	(0.029)***	(0.024)***	(0.031)**	(0.025)**	(0.010)***	(0.010)***	(0.013)*	(0.013)*
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes
Observations	12844	12844	12810	12810	3041	3041	3043	3043
R-squared	0.084	0.131	0.105	0.160	0.037	0.054	0.059	0.072
Mean in control group	0.542	0.542	0.692	0.692	0.046	0.046	0.078	0.078

<i>Panel C: Impact on patient compliance (source: patient surveys)</i>				
	Occasionally sent someone else to get the pills		Ever picked up medicine for one week or more during one visit	
	(1)	(2)	(3)	(4)
Treatment	-0.218	-0.218	-0.079	-0.078
	(0.025)***	(0.025)***	(0.022)***	(0.021)***
Strata fixed effects	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes
Observations	3352	3352	3295	3295
R-squared	0.130	0.149	0.089	0.103
Mean in control group	0.367	0.367	0.283	0.283

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). In Panel A, Panel C, and Panel B, col. 5–8, the unit of observation is a verified patient. We use the full sample of verified patients, whether enrolled before or after the beginning of the experiment. In Panel A, we report these patients' treatment status at the time of their exit survey. Following the official definition by the WHO, we count as default a patient who stopped taking pills before completing four months of treatment or missed doses for a period of two months or more. In Panel C, we use the last survey completed by the patient and control for a dummy equal to 1 if it was the entry survey. In Panel B, col. 1–4, the unit of observation is a verified patient x observation day.

In all regressions, we control for a dummy indicating whether the patient was enrolled after the beginning of the experiment. Patient controls include gender, age, whether the patient belongs to the general caste category, whether she is Hindu, whether she can read and write, her household size, whether she owns her house, whether she has always lived in the area, whether she has lived there for more than six years, and the time it takes her to go to the center.

Table IV: Impact on Detections and Patient Selection

<i>Panel A: Impact on number of verified detections (source: program data cross-checked with patient surveys)</i>						
	Verified detections per center per month					
	(1)	(2)				
Treatment	-0.103 (0.141)	-0.159 (0.188)				
Strata fixed effects	Yes	Yes				
Baseline detection	Yes	Yes				
Health worker controls		Yes				
Observations	1637	1637				
R-squared	0.170	0.182				
Mean in control group	1.762	1.762				

<i>Panel B: Impact on patient selection (patients enrolled after the beginning of the experiment)</i>						
	Control group		Treatment group		P-value	Number of
	Mean	SD	Mean	SD	Treatment = Control	obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.586	0.493	0.576	0.494	0.434	2598
Age	34.4	16.4	33.1	16.1	0.102	2597
Caste						
Does not know	0.048	0.214	0.053	0.224	0.847	2602
General caste	0.181	0.386	0.179	0.383	0.716	2602
Other Backward Class	0.357	0.479	0.352	0.478	0.996	2602
Scheduled Caste	0.231	0.422	0.267	0.442	0.152	2602
Scheduled Tribe	0.078	0.269	0.070	0.255	0.324	2602
Minority	0.104	0.306	0.079	0.270	0.471	2602
Religion						
Hindu	0.832	0.374	0.870	0.337	0.362	2594
Muslim	0.147	0.355	0.117	0.322	0.493	2594
Other	0.021	0.142	0.013	0.114	0.201	2594
Literacy						
Cannot read or write	0.304	0.460	0.294	0.456	0.311	2602
Can read but not write	0.043	0.202	0.044	0.204	0.598	2602
Can read and write	0.653	0.476	0.662	0.473	0.488	2602
Education						
Pre-primary	0.247	0.431	0.226	0.418	0.083	2593
Primary	0.371	0.483	0.410	0.492	0.048	2593
Secondary	0.333	0.471	0.325	0.468	0.910	2593
Undergraduate and more	0.049	0.216	0.039	0.195	0.203	2593
Household size	5.5	2.6	5.4	2.2	0.183	2593
Lives alone	0.025	0.155	0.017	0.128	0.065	2602
Assets						
Has electricity	0.970	0.172	0.966	0.181	0.711	2532
Has tap water	0.562	0.496	0.544	0.498	0.885	2543
Has a television	0.819	0.385	0.822	0.383	0.649	2460
Has a refrigerator	0.264	0.441	0.258	0.438	0.762	2459
Owns her house	0.689	0.463	0.679	0.467	0.712	2538
Migration Status						
Has always lived in the area	0.507	0.500	0.457	0.498	0.217	2600
Has lived there for more than 6 years	0.227	0.419	0.269	0.444	0.079	2600
Has lived there for less than 5 years	0.266	0.442	0.274	0.446	0.840	2600
Is currently working	0.316	0.465	0.321	0.467	0.426	2388
Time to go to the center (in minutes)	11.0	7.4	11.8	8.7	0.281	1812
Distance to the center (in miles)	2.9	4.9	3.0	4.7	0.587	1998

Notes: In Panel A, the unit of observation is a treatment area x month. Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). We count as verified detection a new patient reported by the health worker that we were actually able to find and survey. The baseline outcome is computed on the four months preceding the beginning of the experiment. Health worker controls include gender, age, whether the health worker belongs to the general caste category, whether she is Hindu, whether her level of education is class 12 or below, whether she completed tertiary education, whether she had any previous experience, her household size, and whether she owns her house.

In Panel B, the unit of observation is a verified patient. We use the subsample of verified patients enrolled after the beginning of the experiment and the first survey completed by the patient. For each variable, we report the means and standard deviations in both the control group and the treatment group. We also report the p-value of the difference with control for strata fixed effects and a dummy indicating whether the information comes from an entry or exit survey, and clustered standard errors.

Table V: Impact on Patients' Health

<i>Panel A: Impact on patients' physical health (source: patient surveys)</i>												
	Cured		Patients' ranking of their health		Index of no difficulty in performing activities		Index of health symptoms		Index of severe health symptoms		Index of increased health symptoms	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	0.015	0.012	0.012	-0.010	-0.022	-0.038	-0.008	0.001	-0.035	-0.026	-0.013	-0.011
	(0.025)	(0.023)	(0.089)	(0.086)	(0.030)	(0.029)	(0.032)	(0.031)	(0.026)	(0.024)	(0.018)	(0.019)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes		Yes
Observations	2395	2395	2430	2430	2458	2458	2459	2459	2459	2459	2458	2458
R-squared	0.069	0.095	0.037	0.106	0.048	0.168	0.263	0.296	0.111	0.146	0.005	0.012
Mean in control group	0.400	0.400	7.431	7.431	--	--	--	--	--	--	--	--
<i>Panel B: Impact on patients' emotional health (source: patient surveys)</i>												
	Thinks that household situation will improve in the future		Has a project to improve their life and their family's life		Index of emotional distress		Overall somewhat or very happy		Overall satisfaction with their life			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
Treatment	0.000	0.002	-0.017	-0.017	0.010	0.010	-0.019	-0.020	-0.082	-0.086		
	(0.015)	(0.015)	(0.014)	(0.014)	(0.028)	(0.026)	(0.013)	(0.013)	(0.080)	(0.079)		
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Patient controls		Yes		Yes		Yes		Yes		Yes		
Observations	2077	2077	2375	2375	2454	2454	2435	2435	2426	2426		
R-squared	0.034	0.062	0.088	0.112	0.071	0.102	0.027	0.050	0.194	0.250		
Mean in control group	0.890	0.890	0.120	0.120	--	--	0.942	0.942	7.259	7.259		

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). The unit of observation is a verified patient. We use the full sample of verified patients, whether enrolled before or after the beginning of the experiment, and control for a dummy indicating whether the patient was enrolled after the beginning of the experiment. We report these patients' self-reported health (Panel A) and emotional health (Panel B) outcomes at the time of their exit survey. All outcomes are self-reported. Patient controls as in Table III.

Cured is equal to 1 if the patient thinks they are cured and indicates that a sputum test/x-ray result confirmed that they are no longer sick; that the OpAsha health worker told them they were cured / no longer have to take pills; or that a doctor told them they were cured / no longer have to take pills. Patients' ranking of their health goes from 1 (very poor) to 10 (excellent). Index of no difficulty in performing activities is a standardized index of dummies equal to 1 if the patient hasn't experienced any difficulty in drawing water from a well, routine housework such as cleaning or cooking, routine daily activities, maintaining their usual work schedule and, for children going to school, maintaining their usual class and homework schedule. Index of health symptoms is a standardized index of dummies equal to 1 if the patient has experienced a cough, respiratory problems, other types of body pain, other general health problems or any other health problem since they started taking pills. Index of severe health symptoms is a standardized index of dummies equal to 1 if these symptoms were severe (the cough was painful, very painful, or accompanied with phlegm/blood; the problems prevented them from completing their usual activities, at home or at work/school). Index of increased health symptoms is a standardized index of dummies equal to 1 if these problems have increased since the patient started taking pills. Index of emotional distress is a standardized index of dummies equal to 1 if the patient often or sometimes felt hopeless, that everything is an effort, worthless, envious of others, angry, unable to sleep, prone to crying, never felt relaxed and peaceful, and never felt happy during the past seven days. Patients' overall satisfaction with their life goes from 1 (very bad life) to 10 (very good).

Table VI: Impact on Health Worker Effort

<i>Panel A: Impact on health worker attendance</i>														
	Center open		Total time center was open		Health worker present at the center		Total time health worker was present at the center, net of breaks		Visit by OpAsha		Center open		Health worker present at the center	
	(1)	(2)	(3)	(4)	<i>(source: observation days)</i>		(7)	(8)	(9)	(10)	<i>(source: random spot checks)</i>		(13)	(14)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treatment	0.000 (0.002)	0.001 (0.003)	-10.7 (12.400)	9.2 (14.600)	0.019 (0.022)	0.045 (0.022)*	20.2 (8.2)**	22.3 (8.1)***	0.033 (0.012)***	0.030 (0.011)**	0.001 (0.016)	0.005 (0.021)	0.049 (0.046)	0.071 (0.029)**
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Observations	2,749	2,749	2,414	2,414	3,063	3,063	2,987	2,987	2,855	2,855	916	916	916	916
R-squared	0.002	0.003	0.100	0.200	0.149	0.167	0.200	0.200	0.021	0.031	0.029	0.051	0.154	0.205
Mean in Control Group	0.998	0.998	338.9	338.9	0.834	0.834	117.3	117.3	0.054	0.054	0.945	0.945	0.598	0.598
<i>Panel B: Impact on health worker activities</i>														
	# home visits as % of all patients scheduled to come on that day		Intensity of default activities (index)		Intensity of detection activities (index)		Health worker gives advice related to TB		Frequency of interactions with health worker		Health worker supports patient during treatment		Health worker sometimes or often delivers pills at patient's home	
	<i>(source: observation days)</i>		<i>(source: health worker surveys)</i>				<i>(source: patient surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treatment	0.037 (0.020)*	0.049 (0.026)*	0.031 (0.069)	-0.103 (0.096)	-0.065 (0.056)	-0.071 (0.069)	0.014 (0.016)	0.014 (0.016)	-0.051 (0.063)	-0.045 (0.060)	-0.003 (0.028)	-0.004 (0.028)	0.073 (0.041)*	0.087 (0.035)**
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes								
Baseline control			Yes	Yes	Yes	Yes								
Patient controls								Yes		Yes		Yes		Yes
Observations	2341	2341	134	134	134	134	3280	3280	3193	3193	3293	3293	3295	3295
R-squared	0.111	0.114	0.304	0.380	0.663	0.702	0.051	0.061	0.192	0.203	0.325	0.330	0.335	0.402
Mean in control group	0.115	0.115	--	--	--	--	0.894	0.894	1.749	1.749	0.213	0.213	0.454	0.454

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). Patient and health worker controls as in Tables III and IV. In Panel A, the unit of observation is an observation day (col. 1–8) or a random visit (col. 9–12). In col. 1–4 and 11–14, the sample is restricted to areas with a fixed center. In col. 5–10, the sample includes both areas with a fixed center and those receiving mobile treatment.

In Panel B, col. 1 and 2, the unit of observation is an observation day and the sample is restricted to areas with a fixed center. In col. 3–6, the unit of observation is a health worker x survey (either midline or endline). Intensity of default activities is a standardized index of the health worker's ability to name a patient that they counseled for missing doses or defaulting, the number of methods used to prevent patients from defaulting, the number of days spent visiting ongoing patients over the previous seven days, the number of days spent visiting ongoing patients per week in general, and the number of actions taken when patients miss doses. Intensity of detection activities is a standardized index of the health worker's ability to name a patient that they suspected for TB and sent for sputum testing, the number of sputum samples of TB suspects collected last week, the number of days spent detecting new patients over the previous seven days, the number of days spent detecting new patients per week in general, and the number of methods used to detect new patients. In col. 7–14, the unit of observation is a verified patient. We use the full sample of verified patients, whether enrolled before or after the beginning of the experiment, use the last survey completed by the patient, and control for dummies indicating whether this was the entry survey and whether the patient was enrolled after the beginning of the experiment.

Table VII: Impact on Quality of Reporting on Default

	Actual likelihood to default		Defaults as reported in the program data		Default as reported in the government registers	
	<i>(source: patient surveys)</i>		<i>(source: program data)</i>		<i>(source: government registers)</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.018 (0.007)**	-0.015 (0.007)**	0.000 (0.008)	0.001 (0.008)	0.000 (0.006)	0.000 (0.005)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes				Yes
Baseline outcome			Yes	Yes		
Health worker controls				Yes		
Observations	2802	2802	1832	1832	2913	2913
R-squared	0.020	0.033	0.023	0.031	0.009	0.022
Mean in control group	0.073	0.073	0.045	0.045	0.021	0.021
Test: Treatment = Treatment(1) (or (2))						
P-value	--	--	0.118	0.111	0.060	0.078
F-statistic	--	--	2.51	2.61	3.67	3.22

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). Patient and health worker controls as in Tables III and IV. In all col. but 3 and 4, the unit of observation is a verified patient. We use the full sample of verified patients, whether enrolled before or after the beginning of the experiment, and control for a dummy indicating whether the patient was enrolled after the beginning of the experiment. In col. 3 and 4, the unit of observation is a treatment area x month. The baseline outcome is computed on the six months preceding the beginning of the experiment.

We use a seemingly unrelated regressions framework to compare treatment effects on actual likelihood to default and on default as reported in the program data and in the government registers. We test the null hypothesis that effects on actual and reported defaults are equal and report the corresponding p-value and F-statistic. For instance, in col. 3 (resp. col. 4), we report the p-value and F-statistic corresponding to the test of the null that treatment effects reported in col. 1 and 3 (resp. 2 and 4) are equal.

Table VIII: Impact on Quality of Reporting on Patient Detection

<i>Panel A: Impact on number of verified and reported detections</i>						
	Verified detections <i>(source: program data cross-checked with patient surveys)</i>		Reported detections <i>(source: program data)</i>			
	(1)	(2)	(3)	(4)		
Treatment	-0.103 (0.141)	-0.159 (0.188)	-0.744 (0.243)***	-0.795 (0.260)***		
Strata fixed effects	Yes	Yes	Yes	Yes		
Baseline outcome	Yes	Yes	Yes	Yes		
Health worker controls		Yes		Yes		
Observations	1637	1637	1832	1832		
R-squared	0.170	0.182	0.470	0.489		
Mean in control group	1.762	1.762	3.730	3.730		
Test: Treatment = Treatment(1) (or (2))						
P-value	--	--	0.003	0.014		
F-statistic	--	--	9.29	6.41		
<i>Panel B: Impact on survey outcomes, individual level (sources: program data, patient surveys)</i>						
	Patient verified (found and surveyed)	Survey not completed: Patient refused to answer	Survey not completed: Patient not found	Survey not completed: Health worker warning	Survey not completed: Patient death	Survey not completed: Other reasons
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.058 (0.025)**	-0.020 (0.007)***	-0.011 (0.017)	-0.015 (0.011)	-0.004 (0.007)	-0.008 (0.010)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3647	3647	3647	3647	3647	3647
R-squared	0.057	0.010	0.064	0.025	0.016	0.034
Mean in control group	0.685	0.068	0.106	0.059	0.031	0.049
<i>Panel C. Impact on the fraction of patients reported in the government registers which are verified (sources: government registers, patient surveys)</i>						
	Fraction of verified patients (1)					
Treatment	0.075 (0.028)***					
Strata fixed effects	Yes					
Observations	2998					
R-squared	0.026					
Mean in control group	0.697					

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). In Panel A, the unit of observation is a treatment area x month. We count as verified detection a new patient reported by the health worker that we were actually able to find and survey. The baseline outcome is computed on the four months preceding the beginning of the experiment in col. 1 and 2 and six months in col. 3 and 4 (because more historical data are available). Health worker controls as in Table IV. We use a seemingly unrelated regressions framework to compare treatment effects on the number of verified detections and reported detections. We test the null hypothesis that the two effects are equal and report the corresponding p-value and F-statistic. Specifically, in col. 3 (resp. col. 4), we report the p-value and F-statistic corresponding to the test of the null that treatment effects reported in col. 1 and 3 (resp. 2 and 4) are equal.

In Panel B, the unit of observation is a patient reported in the program data. We use the subsample of patients enrolled after the beginning of the experiment. We report the likelihood that these patients were verified (found and surveyed) and, otherwise, the reason why we could not survey them. We control for a dummy indicating whether the information comes from an entry or exit survey.

In Panel C, the unit of observation is a patient reported in the program data and in the government registers. We use the subsample of patients enrolled after the beginning of the experiment. We report the likelihood that these patients were verified. Indore is excluded as we did not obtain access to the TB register in that city.

Table IX: Sustainability of the Intervention

<i>Panel A: Impact on health worker salary and satisfaction</i>												
	Monthly compensation for detection activities		Monthly salary		# of challenges faced in getting a patient to complete treatment course		# of challenges faced in detecting patients		Excessive workload		Standardized job satisfaction index	
	<i>(source: salary data from OpAsha)</i>				<i>(source: health worker surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	-324.3 (114.4)***	-224.9 (97.1)**	-428.5 (159.7)***	-341.1 (182.7)*	-0.220 (0.164)	-0.238 (0.184)	0.013 (0.162)	-0.167 (0.214)	-0.105 (0.054)*	-0.199 (0.062)***	0.123 (0.093)	0.140 (0.121)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes
Baseline control					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	693	693	936	936	121	121	117	117	118	118	132	132
R-squared	0.400	0.500	0.400	0.500	0.193	0.254	0.207	0.282	0.218	0.255	0.271	0.346
Mean in control group	1321	1321	5791	5791	3.309	3.309	2.746	2.746	0.222	0.222	--	--
<i>Panel B: Impact on patients' satisfaction with healthcare received</i>												
	Satisfied with their interaction with their health worker		Would refer a TB suspect to an OpAsha health worker		# of challenges faced by patients to complete treatment course							
	<i>(source: patient surveys)</i>				<i>(source: health worker surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)						
Treatment	0.005 (0.012)	0.003 (0.012)	0.034 (0.022)	0.033 (0.022)	-0.053 (0.133)	-0.212 (0.204)						
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes						
Patient controls		Yes		Yes								
Health worker controls						Yes						
Baseline control					Yes	Yes						
Observations	2378	2378	2441	2441	119	119						
R-squared	0.024	0.048	0.475	0.482	0.246	0.304						
Mean in control group	0.957	0.957	0.178	0.178	2.500	2.500						
<i>Panel C. Technology presence over time (source: observation days)</i>												
	Center has biometric equipment		Share of all devices (fingerprint reader, laptop, USB key) present		Share of all devices (fingerprint reader, laptop, USB key) present and working							
	(1)	(2)	(3)	(4)	(5)	(6)						
# weeks since beginning of the experiment	-0.002 (0.001)*	-0.002 (0.001)	-0.006 (0.002)***	-0.006 (0.002)***	-0.006 (0.002)***	-0.006 (0.002)***						
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes						
Health worker controls		Yes		Yes		Yes						
Observations	1414	1414	1403	1403	1398	1398						
R-squared	0.309	0.326	0.373	0.389	0.361	0.376						
Mean in first month	0.921	0.921	0.909	0.909	0.899	0.899						

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). Patient and health worker controls as in Tables III and IV. In Panel A, col. 1–4, the unit of observation is a health worker x month. In Panel A, col. 5–10 and Panel B, col. 5 and 6, the unit of observation is a health worker x survey (either midline or endline). The standardized job satisfaction index is a standardized index of the health worker's satisfaction with their compensation, their satisfaction with their work, and whether they recommended the job to someone else in the past six months. In Panel B, col. 1–4, the unit of observation is a verified patient. We use the full sample of verified patients, whether enrolled before or after the beginning of the experiment, report these patients' outcomes at the time of their exit survey, and control for a dummy indicating whether the patient was enrolled after the beginning of the experiment. In Panel C, the unit of observation is an observation day.

APPENDIX FIGURES AND TABLES (FOR ONLINE PUBLICATION)

Appendix A. Additional Results

Figure A1. Technology Presence in Biometric Centers over Time

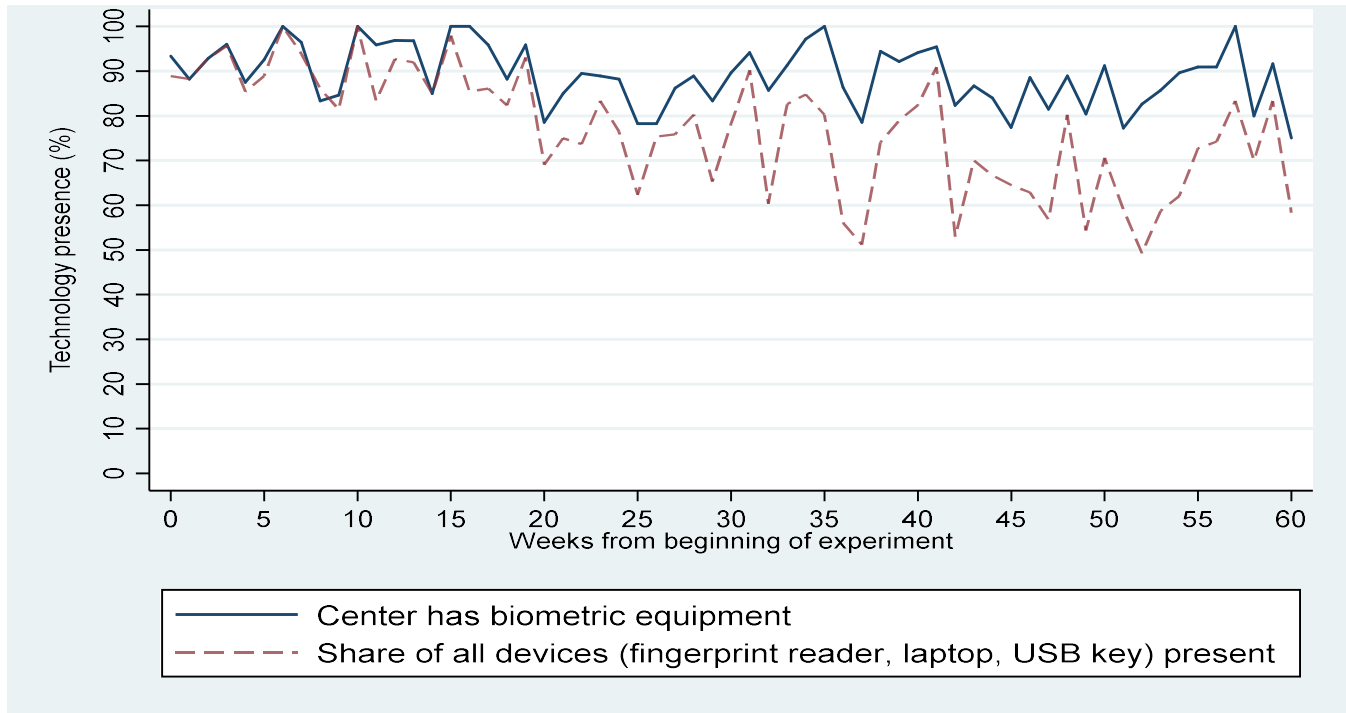


Table A1: Attrition Checks

<i>Panel A: Impact on health worker attrition</i>												
	Treatment		Robust SE		Strata fixed effects		Health worker controls		R-squared		Observations	Mean in control group
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dropped out during intervention	-0.016	0.071	Yes	No	0.213	85	0.273					
	-0.003	0.128	Yes	Yes	0.425	85	0.273					
<i>Panel B: Attritors and replacements' characteristics</i>												
	Attritors						Replacements					
	Control group		Treatment		P-value	Observations	Control group		Treatment		P-value	Observations
	Mean	SD	Mean	SD	T=C		Mean	SD	Mean	SD	T=C	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Male	1.000	0.000	0.700	0.483	0.226	22	0.700	0.483	0.700	0.483	1.000	20
Age	29.3	10.9	30.5	8.3	0.738	22	28.7	7.7	26.3	6.3	0.204	20
Caste												
General caste	0.250	0.452	0.400	0.516	0.567	22	0.400	0.516	0.500	0.527	0.717	20
Other Backward Class	0.500	0.522	0.500	0.527	1.000	22	0.300	0.483	0.500	0.527	0.457	20
Scheduled Caste	0.167	0.389	0.000	0.000	0.115	22	0.200	0.422	0.000	0.000	0.162	20
Scheduled Tribe	0.083	0.289	0.100	0.316	0.904	22	0.000	0.000	0.000	0.000	-	20
Minority	0.000	0.000	0.000	0.000	-	22	0.100	0.316	0.000	0.000	0.248	20
Religion												
Hindu	0.750	0.452	1.000	0.000	0.029	22	0.600	0.516	0.900	0.316	0.126	20
Muslim	0.083	0.289	0.000	0.000	0.260	22	0.100	0.316	0.100	0.316	1.000	20
Other	0.167	0.389	0.000	0.000	0.157	22	0.300	0.483	0.000	0.000	0.023	20
Highest education level achieved												
Class 12 and below	0.636	0.505	0.700	0.483	0.766	21	0.400	0.516	0.700	0.483	0.232	20
Tertiary	0.364	0.505	0.300	0.483	0.766	21	0.600	0.516	0.300	0.483	0.232	20
Other diploma/non-formal education	0.000	0.000	0.000	0.000	-	21	0.000	0.000	0.000	0.000	-	20
Work experience												
Any previous work experience	0.750	0.452	0.700	0.483	0.841	22	0.800	0.422	0.700	0.483	0.656	20
Number of years of work experience	9.091	5.166	10.200	4.341	0.571	21	7.200	4.566	8.700	5.293	0.563	20
Any previous working in a job related to TB	0.083	0.289	0.200	0.422	0.381	22	0.300	0.483	0.000	0.000	0.056	20
Any previous experience in the social/NGO sector	0.000	0.000	0.000	0.000	-	21	0.111	0.333	0.100	0.316	0.939	19
Any other income generating activity, in addition to OpASHA	0.000	0.000	0.100	0.316	0.350	22	0.000	0.000	0.000	0.000	-	19
Lives in one of the areas covered by the centers	0.545	0.522	0.500	0.527	0.862	21	0.400	0.516	0.222	0.441	0.529	19
Household size	5.6	2.5	5.0	3.2	0.645	22	5.3	2.3	5.9	1.9	0.273	20
Lives alone	0.000	0.000	0.100	0.316	0.350	22	0.100	0.316	0.000	0.000	0.248	20
Assets												
Has electricity	1.000	0.000	0.900	0.316	0.351	21	1.000	0.000	1.000	0.000	-	20
Has tap water	0.818	0.405	0.700	0.483	0.595	21	0.900	0.316	0.600	0.516	0.191	20
Has a television	0.818	0.405	0.900	0.316	0.634	21	0.900	0.316	1.000	0.000	0.248	20
Has a refrigerator	0.455	0.522	0.600	0.516	0.450	21	0.500	0.527	0.800	0.422	0.185	20
Rents an apartment or house to a third party	0.083	0.289	0.000	0.000	0.358	20	0.100	0.316	0.111	0.333	0.941	19
Owns her house	0.667	0.492	0.400	0.516	0.219	22	0.600	0.516	0.800	0.422	0.269	20
Exposure to technology												
Knows how to use a computer	0.750	0.452	0.600	0.516	0.434	22	0.500	0.527	0.800	0.422	0.131	20
Knows how to use the internet	0.727	0.467	0.600	0.516	0.534	21	0.500	0.527	0.700	0.483	0.330	20
Has an email account	0.583	0.515	0.300	0.483	0.187	22	0.500	0.527	0.700	0.483	0.330	20
Has a social networking account	0.583	0.515	0.300	0.483	0.139	22	0.500	0.527	0.500	0.527	1.000	20
Days spent in the experiment	138.3	114.3	113.8	108.2	0.638	22	185.1	114.1	218.9	123.0	0.605	20
Reported baseline number of detections per month per area	2.568	2.148	2.770	0.954	0.812	16						
Reported baseline number of defaults per month per area	0.038	0.060	0.014	0.015	0.286	16						
Center open	0.984	0.044	1.000	0.000	1.000	14						
Total time center was open	373.1	235.4	259.7	144.3	0.735	14						
Health worker present at the center	0.772	0.230	0.733	0.406	0.280	15						
Total time health worker was present at the center	96.4	66.5	96.4	73.5	0.374	15						
Visit by OpAsha	0.108	0.182	0.041	0.078	0.400	16						
Adherence 1: Patient came to center and took or picked up pill	0.318	0.279	0.482	0.287	0.374	16						
Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit	0.622	0.303	0.676	0.304	0.492	16						

Notes: In Panel A, the unit of observation is the health worker. ***, **, * indicate significance at 1, 5, and 10 percent, respectively. We include all health workers who were part of the experiment. Health worker controls as in Table IV.

In Panel B, the means and standard deviations in the control and treatment groups are calculated based on the subsample of attritors (health workers who left the sample over the course of the experiment) and replacements (health workers who entered the sample over the course of the experiment) respectively (col. 1–4 and 7–10). We also report the p-value of the difference with clustered standard errors (col. 5 and 11).

In the last ten lines of Panel B, the unit of observation is the cluster of areas allocated to a health worker. The sample includes all clusters in which the initial health worker attrited. Baseline detection and default outcomes are computed based on six months of OpASHA center-wise data sheets prior to the experiment. Baseline observation days outcomes are computed based on observation days that took place before the beginning of the experiment.

Table A2: Availability of Various Data Sources

<i>Panel A: Patient data</i>					
	Entry survey complete (source: patient surveys)	Exit survey complete	Verification data available (source: government registers)	At least one observation day (source: observation days)	Total number of observation days
	(1)	(2)	(3)	(4)	(5)
Treatment	0.024 (0.019)	0.007 (0.017)	0.025 (0.022)	-0.006 (0.014)	0.4 (0.1)**
Strata fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	3393	3393	3057	3393	3393
R-squared	0.476	0.131	0.111	0.064	0.300
Mean in control group	0.607	0.725	0.946	0.896	3.6
<i>Panel B: Healthworker data</i>					
	Baseline survey complete (source: healthworker surveys)	Midline survey complete	Endline survey complete	Number of months of salary data (source: salary data from OpAsha)	
	(1)	(2)	(3)	(4)	
Treatment	0.000 --	-0.038 (0.064)	0.050 (0.080)	-0.2 (0.9)	
Strata fixed effects	Yes	Yes	Yes	Yes	
Observations	85	85	85	85	
R-squared	--	0.282	0.181	0.400	
Mean in control group	1.000	0.818	0.727	11.1	
<i>Panel C: Area-level data</i>					
	At least one observation day (source: observation days)	Total number of observation days	At least one random spot check (source: random spot checks)	Total number of random spot checks	Number of months of program data (source: program data)
	(1)	(2)	(3)	(4)	(5)
Treatment	0.000 --	0.0 (0.7)	0.011 (0.009)	0.9 (0.4)**	-0.3 (0.2)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	122	122	122	122	122
R-squared	--	0.800	0.728	0.600	0.900
Mean in control group	1.000	25.0	0.906	7.1	15.1

Notes : Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). In Panel A, the unit of observation is a verified patient. We use the full sample of verified patients, whether enrolled before or after the beginning of the experiment, and control for a dummy indicating whether the patient was enrolled after the beginning of the experiment. In column 3, we exclude Indore, where we did not obtain access to the TB register. In Panel B, the unit of observation is a healthworker. In Panel C, the unit of observation is a treatment area. While there are 129 areas, the number of observations is 122 because the data for the seven mobile health workers is not area-specific.

Table A3: Impact on Likelihood to Default Predicted Based on Observables

	Default			
	(1)	(2)	(3)	(4)
Male	0.024 (0.016)	0.022 (0.017)	0.021 (0.014)	0.025 (0.016)
Age	0.000 (0.001)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)
Caste: doesn't know	-0.054 (0.052)	0.015 (0.051)	-0.049 (0.054)	0.027 (0.043)
Other Backward Class	0.010 (0.027)	0.018 (0.022)	-0.022 (0.028)	-0.003 (0.020)
Scheduled Caste	-0.001 (0.023)	0.006 (0.016)	-0.019 (0.027)	-0.010 (0.019)
Scheduled Tribe	0.013 (0.039)	0.007 (0.034)	-0.015 (0.036)	0.008 (0.033)
Minority	-0.025 (0.065)	0.011 (0.046)	-0.051 (0.041)	0.008 (0.035)
Muslim	0.014 (0.062)	-0.014 (0.041)	0.003 (0.036)	-0.034 (0.029)
Other religion	-0.091 (0.031)***	-0.090 (0.020)***	-0.078 (0.023)***	-0.018 (0.045)
Can read but not write	-0.101 (0.031)***	-0.079 (0.036)**	-0.119 (0.028)***	-0.052 (0.052)
Can read and write	-0.031 (0.041)	-0.028 (0.036)	-0.056 (0.041)	-0.037 (0.033)
Primary education	0.005 (0.044)	-0.009 (0.044)	0.057 (0.038)	0.032 (0.037)
Secondary education	-0.001 (0.051)	-0.010 (0.045)	0.049 (0.045)	0.024 (0.037)
Undergraduate and more	-0.005 (0.057)	0.007 (0.052)	0.028 (0.045)	0.017 (0.043)
Size of the Household	-0.004 (0.004)	0.001 (0.003)	-0.004 (0.004)	0.000 (0.003)
Lives alone	-0.007 (0.094)	-0.035 (0.055)	-0.034 (0.074)	-0.053 (0.047)
Has tap water	-0.026 (0.023)	-0.036 (0.021)*	-0.026 (0.019)	-0.035 (0.016)**
Has a television	-0.029 (0.033)	0.009 (0.023)	-0.013 (0.023)	0.011 (0.016)
Has a refrigerator	0.039 (0.041)	0.025 (0.032)	0.012 (0.033)	0.015 (0.025)
Owns her house	0.011 (0.027)	0.007 (0.021)	0.011 (0.019)	0.003 (0.016)
Has lived in the area for more than 6 years (but not always)	-0.038 (0.023)	-0.015 (0.028)	-0.020 (0.019)	-0.007 (0.021)
Has lived there for less than 5 years	-0.004 (0.029)	-0.013 (0.026)	0.022 (0.022)	0.000 (0.018)
Is currently working	0.037 (0.027)	0.032 (0.018)*	0.029 (0.018)	0.024 (0.015)
Time to go to the center	-0.001 (0.002)		-0.001 (0.001)	
Distance to the center	0.007 (0.003)***		0.004 (0.003)	
Including patients enrolled before the beginning of the experiment	No	No	Yes	Yes
Observations	593	965	882	1378
R-squared	0.046	0.023	0.031	0.016

Panel B: Impact on likelihood to default predicted based on observables

	Predicted likelihood to default			
	(1)	(2)	(3)	(4)
Treatment	0.003 (0.005)	-0.002 (0.002)	0.006 (0.004)	0.000 (0.002)
Panel A regression includes patients enrolled before the beginning of the experiment	No	No	Yes	Yes
Panel A regression includes time to the center and distance to the center	Yes	No	Yes	No
Strata fixed effects	Yes	Yes	Yes	Yes
Observations	1374	2210	1374	2210
R-squared	0.109	0.078	0.083	0.103
Mean in control group	0.075	0.079	0.070	0.077

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). The unit of observation is a verified patient.

In Panel A, we regress control group verified patients' likelihood to default on patient-level observables. Col. 1 and 2 use the subsample of verified patients enrolled after the beginning of the experiment and col. 3 and 4 use the full sample of verified patients, whether enrolled before or after the beginning of the experiment. Col. 1 and 3 use all patient-level observables shown in Tables II and IV except for "Has electricity," which is equal to 1 for all patients, and col. 2 and 4 drop variables observed for fewer than 90 percent of patients: time to go to the center and distance to the center. Omitted categories are general caste, Hindu, cannot read or write, pre-primary education, and has always lived in the area.

In Panel B, we predict the likelihood to default of all treatment and control group patients enrolled after the beginning of the experiment, based on the coefficients estimated in Panel A, and using all specifications shown in Panel A in turn.

Table A4: Responsiveness of OpAsha Visits to Patients' Treatment Adherence and Health Workers' Presence at the Center

	(1)	(2)	Visit by OpAsha		(5)	(6)
			(3)	(4)		
Patient adherence lagged	0.041 (0.027)	0.031 (0.023)				
Patient adherence lagged * Treatment	-0.103 (0.048)**	-0.079 (0.045)*				
Health worker present at the center lagged			0.032 (0.026)	0.022 (0.024)		
Health worker present at the center lagged * Treatment			-0.081 (0.044)*	-0.066 (0.047)		
Total time health worker was present at the center lagged					0.007 (0.006)	0.002 (0.005)
Total time health worker was present at the center lagged * Treatment					-0.004 (0.006)	0.000 (0.006)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes
Observations	2692	2692	2776	2776	2773	2773
R-squared	0.024	0.032	0.023	0.033	0.022	0.032

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). The unit of observation is an observation day. All regressions control for the direct influence of Treatment. Health worker controls as in Table IV.

Patient adherence measures average adherence of all the center's patients. Individual patient adherence is equal to 1 if the patient came to the center and took or picked up the pill, their relative came to the center and picked up the pill, or the patient or their relative was given the pill during a home visit (in line with the definition of Adherence 2 in Table III). The total time the health worker was present at the center is computed net of breaks and measured in number of hours. The lagged values of patient adherence, health worker presence at the center, and the total time the health worker was present at the center are computed over the four previous observation days.

Table A5: Impact on Quality of Reporting on Patient Detection, Baseline

Panel A: Impact on survey outcomes, individual level (sources: program data, patient surveys)

	Patient verified (found and surveyed)	Survey not completed: Patient refused to answer	Survey not completed: Patient not found	Survey not completed: Health worker warning	Survey not completed: Patient death	Survey not completed: Other reasons
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.025 (0.030)	0.007 (0.014)	0.031 (0.028)	-0.002 (0.012)	-0.005 (0.011)	-0.006 (0.006)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1251	1251	1251	1251	1251	1251
R-squared	0.027	0.026	0.013	0.055	0.020	0.030
Mean in control group	0.637	0.081	0.150	0.044	0.058	0.029

Panel B. Impact on the fraction of patients reported in the government registers which are verified (sources: government registers, patient surveys)

	Likelihood to find the patient in government registers (1)
Treatment	-0.044 (0.034)
Strata fixed effects	Yes
Observations	950
R squared	0.030
Mean in control group	0.676

Notes: Clustered standard errors are in parentheses (***, **, * indicate significance at 1, 5, and 10 percent, respectively). In Panel A, the unit of observation is a patient reported in the program data. We use the subsample of verified patients enrolled before the beginning of the experiment. Other notes as in Table VIII.

Table A6: Changes in Impact on Treatment Interruption and Adherence over Time

<i>Panel A: Impact on treatment interruption, or default (source: patient surveys)</i>										
	Treatment complete		Treatment ongoing		Default		Died		Outcome unknown	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.001	-0.003	0.017	0.013	-0.023	-0.022	-0.003	-0.002	0.011	0.014
	(0.031)	(0.030)	(0.031)	(0.031)	(0.010)**	(0.010)**	(0.004)	(0.004)	(0.012)	(0.011)
Treatment x time	0.003	0.000	0.002	0.003	0.006	0.008	0.002	0.002	-0.013	-0.013
	(0.017)	(0.016)	(0.019)	(0.019)	(0.008)	(0.008)	(0.003)	(0.003)	(0.010)	(0.009)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes
Observations	2802	2802	2802	2802	2802	2802	2802	2802	2802	2802
R squared	0.088	0.107	0.101	0.111	0.020	0.033	0.014	0.025	0.055	0.106
Mean in control group	0.507	0.507	0.299	0.299	0.073	0.073	0.016	0.016	0.105	0.105

<i>Panel B: Impact on treatment adherence (source: observation days)</i>								
	Adherence 1: Patient came to the center and took or picked up the pill		Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit		Compliance with treatment verified (as per adherence 1)		Compliance with treatment verified (as per adherence 2)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.197	0.186	0.118	0.113	0.034	0.036	0.021	0.022
	(0.045)***	(0.040)***	(0.044)***	(0.037)***	(0.012)***	(0.012)***	(0.014)	(0.013)*
Treatment x time	-0.026	-0.028	-0.020	-0.025	0.004	0.002	0.003	0.001
	(0.018)	(0.016)*	(0.019)	(0.017)	(0.007)	(0.007)	(0.008)	(0.008)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes
Observations	12844	12844	12810	12810	3041	3041	3043	3043
R-squared	0.091	0.137	0.109	0.163	0.039	0.056	0.061	0.077
Mean in control group	0.542	0.542	0.692	0.692	0.046	0.046	0.078	0.078

<i>Panel C: Impact on patient compliance (source: patient surveys)</i>				
	Occasionally sent someone else to get the pills		Ever picked up medicine for one week or more during one visit	
	(1)	(2)	(3)	(4)
Treatment	-0.227	-0.226	-0.086	-0.086
	(0.027)***	(0.027)***	(0.024)***	(0.023)***
Treatment x time	0.008	0.007	0.006	0.006
	(0.013)	(0.013)	(0.015)	(0.013)
Strata fixed effects	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes
Observations	3352	3352	3295	3295
R-squared	0.130	0.149	0.090	0.103
Mean in control group	0.367	0.367	0.283	0.283

Notes: We control for time lapsed between the beginning of the experiment and the patient's treatment start date in Panels A and C as well as Panel B, col. 5–8, and for time lapsed between the beginning of the experiment and the observation date, in Panel B, col. 1–4 (in hundreds of days). We interact treatment with time. Other notes as in Table III.

Appendix B. Patients Enrolled Before the Beginning of the Experiment

Table B1: Impact on Treatment Interruption and Adherence (Patients Enrolled Before the Beginning of the Experiment)

<i>Panel A: Impact on treatment interruption, or default (source: patient surveys)</i>								
	Treatment complete		Treatment ongoing		Default		Outcome unknown	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-0.026 (0.045)	-0.024 (0.044)	0.046 (0.042)	0.044 (0.041)	-0.016 (0.018)	-0.017 (0.018)	-0.004 (0.006)	-0.002 (0.007)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes
Observations	788	788	788	788	788	788	788	788
R squared	0.027	0.068	0.031	0.070	0.030	0.053	0.047	0.077
Mean in control group	0.604	0.604	0.321	0.321	0.065	0.065	0.010	0.010
<i>Panel B: Impact on treatment adherence (source: observation days)</i>								
	Adherence 1: Patient came to the center and took or picked up the pill		Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit		Compliance with treatment verified (as per adherence 1)		Compliance with treatment verified (as per adherence 2)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.158 (0.038)***	0.162 (0.034)***	0.041 (0.038)	0.046 (0.035)	-0.009 (0.010)	-0.005 (0.008)	-0.026 (0.011)**	-0.022 (0.011)*
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes
Observations	1812	1812	1819	1819	634	634	636	636
R-squared	0.086	0.136	0.114	0.176	0.027	0.103	0.079	0.137
Mean in control group	0.539	0.539	0.710	0.710	0.031	0.031	0.055	0.055
<i>Panel C: Impact on patient compliance (source: patient surveys)</i>								
	Occasionally sent someone else to get the pills		Ever picked up medicine for one week or more during one visit					
	(1)	(2)	(3)	(4)				
Treatment	-0.222 (0.038)***	-0.215 (0.038)***	-0.103 (0.030)***	-0.107 (0.029)***				
Strata fixed effects	Yes	Yes	Yes	Yes				
Patient controls								
Observations	770	770	764	764				
R-squared	0.113	0.162	0.080	0.128				
Mean in control group	0.414	0.414	0.291	0.291				

Notes: We use the subsample of verified patients enrolled before the beginning of the experiment. In Panel A, we don't report regression results for "Died" because this outcome takes value 0 for all patients in this subsample. Other notes as in Table III.

Table B2: Impact on Patients' Health (Patients Enrolled Before the Beginning of the Experiment)

<i>Panel A: Impact on patients' physical health (source: patient surveys)</i>												
	Cured		Patient's ranking of their health		Index of no difficulty in performing activities		Index of health symptoms		Index of severe health symptoms		Index of increased health symptoms	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	0.037 (0.044)	0.042 (0.040)	-0.062 (0.126)	-0.032 (0.114)	-0.046 (0.057)	-0.052 (0.048)	-0.010 (0.034)	-0.004 (0.031)	-0.055 (0.043)	-0.054 (0.040)	-0.007 (0.027)	-0.011 (0.028)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes		Yes
Observations	770	770	776	776	782	782	783	783	783	783	782	782
R-squared	0.063	0.101	0.043	0.136	0.052	0.196	0.311	0.368	0.159	0.213	0.016	0.044
Mean in control group	0.406	0.406	7.178	7.178	--	--	--	--	--	--	--	--
<i>Panel B: Impact on patients' emotional health (source: patient surveys)</i>												
	Thinks that household situation will improve in the future		Has a project to improve their life and their family's life		Index of emotional distress		Overall somewhat or very happy		Overall satisfaction with their life			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
Treatment	-0.031 (0.024)	-0.022 (0.024)	-0.019 (0.020)	-0.018 (0.020)	-0.043 (0.041)	-0.045 (0.039)	-0.034 (0.022)	-0.033 (0.021)	-0.282 (0.125)**	-0.241 (0.114)**		
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Patient controls		Yes		Yes		Yes		Yes		Yes		
Observations	686	686	772	772	780	780	774	774	775	775		
R-squared	0.057	0.108	0.145	0.182	0.081	0.128	0.051	0.079	0.139	0.209		
Mean in control group	0.864	0.864	0.144	0.144	--	--	0.947	0.947	7.252	7.252		

Notes : We use the subsample of verified patients enrolled before the beginning of the experiment. Other notes as in Table V.

Table B3: Impact on Health Worker Activities as Reported by Patients (Patients Enrolled Before the Beginning of the Experiment)

	Health worker gives advice related to TB		Frequency of interactions with health worker		Health worker supports patient during treatment		Health worker sometimes or often delivers pills at patient's home	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-0.012 (0.023)	-0.013 (0.023)	-0.027 (0.100)	-0.008 (0.101)	-0.017 (0.036)	-0.011 (0.034)	0.076 (0.043)*	0.073 (0.042)*
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes
Observations	759	759	730	730	761	761	761	761
R-squared	0.084	0.114	0.075	0.096	0.421	0.444	0.340	0.363
Mean in control group	0.915	0.915	1.403	1.403	0.227	0.227	0.448	0.448

Notes: We use the subsample of verified patients enrolled before the beginning of the experiment. Other notes as in Table VI.

Table B4: Impact on Quality of Reporting on Default (Patients Enrolled Before the Beginning of the Experiment)

	Actual likelihood to default		Default as reported in the government registers	
	<i>(source: patient surveys)</i>		<i>(source: government registers)</i>	
	(1)	(2)	(3)	(4)
Treatment	-0.016 (0.018)	-0.017 (0.018)	0.008 (0.011)	0.007 (0.010)
Strata fixed effects	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes
Observations	788	788	643	643
R-squared	0.030	0.053	0.011	0.053
Mean in control group	0.065	0.065	0.023	0.023
Test: Treatment = Treatment(1) (or (2))				
P-value	--	--	0.169	0.156
F-statistic	--	--	1.93	2.06

Notes: We use the subsample of verified patients enrolled before the beginning of the experiment. Other notes as in Table VII.

Table B5: Impact on Patients' Satisfaction with Healthcare Received (Patients Enrolled Before the Beginning of the Experiment)

	Satisfied with their interaction with their health worker		Would refer a TB suspect to an OpAsha health worker	
	(1)	(2)	(3)	(4)
Treatment	0.003 (0.014)	0.005 (0.014)	-0.020 (0.024)	-0.019 (0.022)
Strata fixed effects	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes
Observations	752	752	778	778
R-squared	0.032	0.060	0.393	0.419
Mean in control group	0.965	0.965	0.227	0.227

Notes: We use the subsample of verified patients enrolled before the beginning of the experiment. Other notes as in Table IX.

Table B6: Changes in Impact on Treatment Interruption and Adherence over Time (Patients Enrolled Before the Beginning of the Experiment)

<i>Panel A: Impact on treatment interruption, or default (source: patient surveys)</i>								
	Treatment complete		Treatment ongoing		Default		Outcome unknown	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.018 (0.072)	0.018 (0.070)	-0.019 (0.057)	-0.014 (0.055)	0.004 (0.040)	-0.002 (0.040)	-0.004 (0.014)	-0.002 (0.014)
Treatment x time	0.059 (0.079)	0.057 (0.082)	-0.086 (0.068)	-0.078 (0.070)	0.027 (0.043)	0.021 (0.046)	0.000 (0.015)	0.000 (0.014)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes
Observations	788	788	788	788	788	788	788	788
R squared	0.027	0.069	0.034	0.072	0.033	0.056	0.053	0.083
Mean in control group	0.604	0.604	0.321	0.321	0.065	0.065	0.010	0.010
<i>Panel B: Impact on treatment adherence (source: observation days)</i>								
	Adherence 1: Patient came to the center and took or picked up the pill		Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit		Compliance with treatment verified (as per adherence 1)		Compliance with treatment verified (as per adherence 2)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.210 (0.057)***	0.203 (0.055)***	0.118 (0.057)**	0.121 (0.051)**	-0.026 (0.027)	-0.021 (0.026)	-0.059 (0.034)*	-0.053 (0.032)
Treatment x time	-0.058 (0.039)	-0.048 (0.037)	-0.084 (0.037)**	-0.081 (0.033)**	-0.026 (0.026)	-0.023 (0.027)	-0.050 (0.044)	-0.046 (0.044)
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes
Observations	1812	1812	1819	1819	634	634	636	636
R-squared	0.099	0.148	0.124	0.184	0.045	0.117	0.111	0.164
Mean in control group	0.539	0.539	0.710	0.710	0.031	0.031	0.055	0.055
<i>Panel C: Impact on patient compliance (source: patient surveys)</i>								
	Occasionally sent someone else to get the pills		Ever picked up medicine for one week or more during one visit					
	(1)	(2)	(3)	(4)				
Treatment	-0.297 (0.069)***	-0.279 (0.064)***	-0.160 (0.065)**	-0.156 (0.065)**				
Treatment x time	-0.101 (0.084)	-0.086 (0.083)	-0.074 (0.088)	-0.065 (0.084)				
Strata fixed effects	Yes	Yes	Yes	Yes				
Patient controls		Yes		Yes				
Observations	770	770	764	764				
R-squared	0.115	0.165	0.083	0.129				
Mean in control group	0.414	0.414	0.291	0.291				

Notes: We use the subsample of verified patients enrolled before the beginning of the experiment. Other notes as in Table A6.

Appendix C. Wild Cluster Bootstrap Standard Errors

Table C1: Health Worker Summary Statistics (Wild Cluster Bootstrap)

	Control group mean (1)	Treatment group mean (2)	P-value Treatment = Control (3)
<i>Panel A. Health workers' characteristics</i>			
Male	0.794	0.710	0.394
Age	33.4	30.4	0.081
Caste			
General caste	0.265	0.516	0.051
Other Backward Class	0.382	0.323	0.700
Scheduled Caste	0.118	0.097	0.828
Scheduled Tribe	0.206	0.032	0.031
Minority	0.029	0.032	0.920
Religion			
Hindu	0.765	0.903	0.128
Muslim	0.176	0.097	0.374
Other	0.059	0.000	0.070
Highest education level achieved			
Class 12 and below	0.688	0.633	0.682
Tertiary	0.281	0.333	0.663
Other diploma/non-formal education	0.031	0.033	1.000
Work experience			
Any previous work experience	0.853	0.645	0.036
Number of years of work experience	9.2	8.7	0.753
Any previous working in a job related to TB	0.147	0.161	0.902
Any previous experience in the social/NGO sector	0.206	0.138	0.536
Any other income generating activity, in addition to OpASHA	0.088	0.129	0.742
Lives in one of the areas she covers	0.545	0.742	0.092
Household size	5.2	6.2	0.262
Lives alone	0.000	0.067	0.188
Assets			
Has electricity	1.000	0.933	0.182
Has tap water	0.594	0.400	0.102
Has a television	0.848	0.933	0.218
Has a refrigerator	0.545	0.467	0.530
Rents an apartment or house to a third party	0.118	0.207	0.210
Owns her house	0.559	0.581	0.934
Exposure to technology			
Knows how to use a computer	0.559	0.516	0.731
Knows how to use the internet	0.424	0.433	0.949
Has an email account	0.412	0.258	0.177
Has a social networking account	0.294	0.258	0.660
Days spent in the experiment	349.1	333.7	0.659
<i>Panel B. OpASHA center-wise data sheets</i>			
Reported baseline number of detections per month per area	2.965	2.393	0.054
Reported baseline fraction of defaults per month per area	0.029	0.039	0.426
<i>Panel C. Observation days</i>			
Center open	0.980	0.989	0.515
Total time center was open	307.4	311.8	0.961
Health worker present at the center	0.769	0.810	0.519
Total time health worker was present at the center, net of breaks	110.9	106.0	0.710
Visit by OpAsha	0.067	0.043	0.333
Adherence 1: Patient came to center and took or picked up pill	0.520	0.517	0.956
Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit	0.761	0.758	0.837

Notes : We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table I.

Table C2: Patient Summary Statistics (Wild Cluster Bootstrap)

	Control group mean	Treatment group mean	P-value Treatment = Control
	(1)	(2)	(3)
Male	0.535	0.576	0.562
Age	34.0	34.3	0.800
Caste			
Does not know	0.040	0.026	0.568
General caste	0.160	0.165	0.628
Other Backward Class	0.376	0.352	0.691
Scheduled Caste	0.229	0.297	0.128
Scheduled Tribe	0.118	0.103	0.210
Minority	0.078	0.058	0.590
Religion			
Hindu	0.869	0.903	0.626
Muslim	0.112	0.081	0.696
Other	0.019	0.016	0.577
Literacy			
Cannot read or write	0.271	0.301	0.469
Can read but not write	0.021	0.026	0.930
Can read and write	0.708	0.673	0.424
Education			
Pre-primary	0.217	0.234	0.756
Primary	0.375	0.390	0.568
Secondary	0.367	0.334	0.526
Undergraduate and more	0.042	0.042	0.926
Household size	5.6	5.4	0.191
Lives alone	0.019	0.016	0.566
Assets			
Has electricity	0.968	0.935	0.042
Has tap water	0.520	0.418	0.524
Has a television	0.826	0.832	0.536
Has a refrigerator	0.276	0.230	0.424
Owns her house	0.716	0.738	0.985
Migration status			
Has always lived in the area	0.570	0.542	0.365
Has lived there for more than 6 years	0.209	0.258	0.050
Has lived there for less than 5 years	0.221	0.200	0.490
Is currently working	0.381	0.406	0.801
Time to go to the center (in minutes)	11.6	12.4	0.844
Distance to the center (in miles)	3.3	3.0	0.732

Notes : We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table II.

Table C3: Impact on Treatment Interruption and Adherence (Wild Cluster Bootstrap)

<i>Panel A: Impact on treatment interruption, or default (source: patient surveys)</i>										
	Treatment complete		Treatment ongoing		Default		Died		Outcome unknown	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.003	-0.007	0.022	0.020	-0.018	-0.015	-0.001	0.000	0.000	0.003
P-value	0.957	0.836	0.490	0.535	0.021	0.045	0.808	0.979	1.000	0.802
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes
Observations	2802	2802	2802	2802	2802	2802	2802	2802	2802	2802
Mean in control group	0.507	0.507	0.299	0.299	0.073	0.073	0.016	0.016	0.105	0.105
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on treatment adherence (source: observation days)</i>										
	Adherence 1: Patient came to the center and took or picked up the pill		Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit		Compliance with treatment verified (as per adherence 1)		Compliance with treatment verified (as per adherence 2)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Treatment	0.139	0.122	0.073	0.056	0.039	0.038	0.024	0.023		
P-value	0.000	0.000	0.035	0.044	0.002	0.000	0.145	0.154		
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Patient controls		Yes		Yes		Yes		Yes		
Observations	12844	12844	12810	12810	3041	3041	3043	3043		
Mean in control group	0.542	0.542	0.692	0.692	0.046	0.046	0.078	0.078		
Number replications	5000	5000	5000	5000	5000	5000	5000	5000		
<i>Panel C: Impact on patient compliance (source: patient surveys)</i>										
	Occasionally sent someone else to get the pills		Ever picked up medicine for one week or more during one visit							
	(1)	(2)	(3)	(4)						
Treatment	-0.218	-0.218	-0.079	-0.078						
P-value	0.000	0.000	0.002	0.002						
Strata fixed effects	Yes	Yes	Yes	Yes						
Patient controls		Yes		Yes						
Observations	3352	3352	3295	3295						
Mean in control group	0.367	0.367	0.283	0.283						
Number replications	5000	5000	5000	5000						

Notes: We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table III.

Table C4: Impact on Detections and Patient Selection (Wild Cluster Bootstrap)

Panel A: Impact on number of verified detections (source: program data cross-checked with patient surveys)

	Verified detections per center per month	
	(1)	(2)
Treatment	-0.103	-0.159
P-value	0.522	0.527
Strata fixed effects	Yes	Yes
Baseline detection	Yes	Yes
Health worker controls		Yes
Observations	1637	1637
Mean in control group	1.762	1.762
Number replications	5000	5000

Panel B: Impact on patient selection (patients enrolled after the beginning of the experiment)

	Control group mean (1)	Treatment group mean (2)	P-value Treatment = Control (3)	Number of obs. (4)
Male	0.586	0.576	0.463	2598
Age	34.4	33.1	0.171	2597
Caste				
Does not know	0.048	0.053	0.879	2602
General caste	0.181	0.179	0.735	2602
Other Backward Class	0.357	0.352	1.000	2602
Scheduled Caste	0.231	0.267	0.200	2602
Scheduled Tribe	0.078	0.070	0.370	2602
Minority	0.104	0.079	0.542	2602
Religion				
Hindu	0.832	0.870	0.376	2594
Muslim	0.147	0.117	0.498	2594
Other	0.021	0.013	0.257	2594
Literacy				
Cannot read or write	0.304	0.294	0.346	2602
Can read but not write	0.043	0.044	0.637	2602
Can read and write	0.653	0.662	0.510	2602
Education				
Pre-primary	0.247	0.226	0.109	2593
Primary	0.371	0.410	0.092	2593
Secondary	0.333	0.325	0.910	2593
Undergraduate and more	0.049	0.039	0.278	2593
Household size	5.5	5.4	0.234	2593
Lives alone	0.025	0.017	0.109	2602
Assets				
Has electricity	0.970	0.966	0.726	2532
Has tap water	0.562	0.544	0.900	2543
Has a television	0.819	0.822	0.700	2460
Has a refrigerator	0.264	0.258	0.804	2459
Owns her house	0.689	0.679	0.730	2538
Migration Status				
Has always lived in the area	0.507	0.457	0.267	2600
Has lived there for more than 6 years	0.227	0.269	0.118	2600
Has lived there for less than 5 years	0.266	0.274	0.849	2600
Is currently working	0.316	0.321	0.449	2388
Time to go to the center (in minutes)	11.0	11.8	0.314	1812
Distance to the center (in miles)	2.9	3.0	0.652	1998

Notes: We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table IV.

Table C5: Impact on Patients' Health (Wild Cluster Bootstrap)

<i>Panel A: Impact on patients' physical health (source: patient surveys)</i>												
	Cured		Patients' ranking of their health		Index of no difficulty in performing activities		Index of health symptoms		Index of severe health symptoms		Index of increased health symptoms	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	0.015	0.012	0.012	-0.010	-0.022	-0.038	-0.008	0.001	-0.035	-0.026	-0.013	-0.011
P-value	0.568	0.645	0.904	0.912	0.484	0.235	0.840	0.976	0.228	0.320	0.528	0.588
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes		Yes
Observations	2395	2395	2430	2430	2458	2458	2459	2459	2459	2459	2458	2458
Mean in control group	0.400	0.400	7.431	7.431	--	--	--	--	--	--	--	--
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on patients' emotional health (source: patient surveys)</i>												
	Thinks that household situation will improve in the future		Has a project to improve their life and their family's life		Index of emotional distress		Overall somewhat or very happy		Overall satisfaction with their life			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
Treatment	0.000	0.002	-0.017	-0.017	0.010	0.010	-0.019	-0.020	-0.082	-0.086		
P-value	0.996	0.922	0.283	0.287	0.747	0.733	0.195	0.174	0.368	0.344		
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Patient controls		Yes		Yes		Yes		Yes		Yes		
Observations	2077	2077	2375	2375	2454	2454	2435	2435	2426	2426		
Mean in control group	0.890	0.890	0.120	0.120	--	--	0.942	0.942	7.259	7.259		
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000		

Notes : We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table V.

Table C6: Impact on Health Worker Effort (Wild Cluster Bootstrap)

<i>Panel A: Impact on health worker attendance</i>														
	Center open		Total time center was open		Health worker present at the center		Total time health worker was present at the center, net of breaks		Visit by OpAsha		Center open		Health worker present at the center	
	<i>(source: observation days)</i>				<i>(source: random spot checks)</i>									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treatment	0.000	0.001	-10.7	9.2	0.019	0.045	20.2	22.3	0.033	0.030	0.001	0.005	0.049	0.071
P-value	0.794	0.784	0.448	0.605	0.469	0.110	0.033	0.035	0.011	0.028	0.974	0.850	0.322	0.025
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Observations	2,749	2,749	2,414	2,414	3,063	3,063	2,987	2,987	2,855	2,855	916	916	916	916
Mean in Control Group	0.998	0.998	338.9	338.9	0.834	0.834	117.3	117.3	0.054	0.054	0.945	0.945	0.598	0.598
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on health worker activities</i>														
	# home visits as % of all patients scheduled to come on that day		Intensity of default activities (index)		Intensity of detection activities (index)		Health worker gives advice related to TB		Frequency of interactions with health worker		Health worker supports patient during treatment		Health worker sometimes or often delivers pills at patient's home	
	<i>(source: observation days)</i>		<i>(source: health worker surveys)</i>				<i>(source: patient surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treatment	0.037	0.049	0.031	-0.103	-0.065	-0.071	0.014	0.014	-0.051	-0.045	-0.003	-0.004	0.073	0.087
P-value	0.123	0.150	0.662	0.354	0.281	0.352	0.473	0.456	0.483	0.515	0.937	0.911	0.114	0.030
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Baseline control			Yes	Yes	Yes	Yes								
Patient controls								Yes		Yes		Yes		Yes
Observations	2341	2341	134	134	134	134	3280	3280	3193	3193	3293	3293	3295	3295
Mean in control group	0.115	0.115	--	--	--	--	0.894	0.894	1.749	1.749	0.213	0.213	0.454	0.454
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000

Notes: We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table VI.

Table C7: Impact on Quality of Reporting on Default (Wild Cluster Bootstrap)

	Actual likelihood to default		Defaults as reported in the program data		Default as reported in the government registers	
	<i>(source: patient surveys)</i>		<i>(source: program data)</i>		<i>(source: government registers)</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.018	-0.015	0.000	0.001	0.000	0.000
P-value	0.021	0.045	0.992	0.936	0.969	0.946
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes				Yes
Baseline outcome			Yes	Yes		
Health worker controls				Yes		
Observations	2802	2802	1832	1832	2913	2913
Mean in control group	0.073	0.073	0.045	0.045	0.021	0.021
Number replications	5000	5000	5000	5000	5000	5000
Test: Treatment = Treatment(1) (or (2))						
P-value	--	--	0.159	0.168	0.109	0.142
F-statistic	--	--				

Notes: We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Unlike in Table VII, we cannot compare treatment effects on actual likelihood to default and on default as reported in the program data and in the government registers using a seemingly unrelated regressions framework as the covariance matrix stored by the `cgmwildboot` Stata command cannot be used in any post-estimation inference (see the Stata help file for `cgmwildboot`). Other notes as in Table VII.

Table C8: Impact on Quality of Reporting on Patient Detection (Wild Cluster Bootstrap)

<i>Panel A: Impact on number of verified and reported detections</i>						
	Verified detections <i>(source: program data cross-checked with patient surveys)</i>		Reported detections <i>(source: program data)</i>			
	(1)	(2)	(3)	(4)		
Treatment	-0.103	-0.159	-0.744	-0.795		
P-value	0.522	0.527	0.012	0.019		
Strata fixed effects	Yes	Yes	Yes	Yes		
Baseline outcome	Yes	Yes	Yes	Yes		
Health worker controls		Yes		Yes		
Observations	1637	1637	1832	1832		
Mean in control group	1.762	1.762	3.730	3.730		
Number replications	5000	5000	5000	5000		
Test: Treatment = Treatment(1) (or (2))						
P-value	--	--	0.006	0.042		
F-statistic	--	--				
<i>Panel B: Impact on survey outcomes, individual level (sources: program data, patient surveys)</i>						
	Patient verified (found and surveyed)	Survey not completed: Patient refused to answer	Survey not completed: Patient not found	Survey not completed: Health worker warning	Survey not completed: Patient death	Survey not completed: Other reasons
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.058	-0.020	-0.011	-0.015	-0.004	-0.008
P-value	0.043	0.010	0.554	0.189	0.621	0.493
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3647	3647	3647	3647	3647	3647
Mean in control group	0.685	0.068	0.106	0.059	0.031	0.049
Number replications	5000	5000	5000	5000	5000	5000
<i>Panel C. Impact on the fraction of patients reported in the government registers which are verified (sources: government registers, patient surveys)</i>						
	Fraction of verified patients (1)					
Treatment	0.075					
P-value	0.020					
Strata fixed effects	Yes					
Observations	2998					
Mean in control group	0.697					
Number replications	5000					

Notes: We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Unlike in Table VIII, we cannot compare treatment effects on the number of verified detections and reported detections using a seemingly unrelated regressions framework as the covariance matrix stored by the `cgmwildboot` Stata command cannot be used in any post-estimation inference (see the Stata help file for `cgmwildboot`). Other notes as in Table VIII.

Table C9: Sustainability of the Intervention (Wild Cluster Bootstrap)

<i>Panel A: Impact on health worker salary and satisfaction</i>												
	Monthly compensation for detection activities		Monthly salary		# of challenges faced in getting a patient to complete treatment course		# of challenges faced in detecting patients		Excessive workload		Standardized job satisfaction index	
	<i>(source: salary data from OpAsha)</i>				<i>(source: health worker surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	-324.3	-224.9	-428.5	-341.1	-0.220	-0.238	0.013	-0.167	-0.105	-0.199	0.123	0.140
P-value	0.009	0.074	0.019	0.148	0.211	0.255	0.941	0.465	0.077	0.005	0.231	0.302
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes
Baseline control					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	693	693	936	936	121	121	117	117	118	118	132	132
Mean in control group	1321	1321	5791	5791	3.309	3.309	2.746	2.746	0.222	0.222	--	--
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on patients' satisfaction with healthcare received</i>												
	Satisfied with their interaction with their health worker		Would refer a TB suspect to an OpAsha health worker		# of challenges faced by patients to complete treatment course							
	<i>(source: patient surveys)</i>				<i>(source: health worker surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)						
Treatment	0.005	0.003	0.034	0.033	-0.053	-0.212						
P-value	0.704	0.796	0.164	0.169	0.693	0.435						
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes						
Patient controls		Yes		Yes								
Health worker controls						Yes						
Baseline control					Yes	Yes						
Observations	2378	2378	2441	2441	119	119						
Mean in control group	0.957	0.957	0.178	0.178	2.500	2.500						
Number replications	5000	5000	5000	5000	5000	5000						
<i>Panel C. Technology presence over time (source: observation days)</i>												
	Center has biometric equipment		Share of all devices (fingerprint reader, laptop, USB key) present		Share of all devices (fingerprint reader, laptop, USB key) present and working							
	(1)	(2)	(3)	(4)	(5)	(6)						
# weeks since beginning of the experiment	-0.002	-0.002	-0.006	-0.006	-0.006	-0.006						
Center fixed effects	0.088	0.152	0.001	0.001	0.001	0.000						
Health worker controls	Yes	Yes	Yes	Yes	Yes	Yes						
Observations	1414	1414	1403	1403	1398	1398						
Mean in first month	0.921	0.921	0.909	0.909	0.899	0.899						
Number replications	5000	5000	5000	5000	5000	5000						

Notes: We use the wild cluster bootstrap procedure proposed by Cameron, Colin, Gelbach, and Miller (2008) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table IX.

Appendix D. Pairs Cluster Bootstrap Standard Errors

Table D1: Health Worker Summary Statistics (Pairs Cluster Bootstrap)

	Control group mean (1)	Treatment group mean (2)	P-value Treatment = Control (3)
<i>Panel A. Health workers' characteristics</i>			
Male	0.794	0.710	0.455
Age	33.4	30.4	0.108
Caste			
General caste	0.265	0.516	0.069
Other Backward Class	0.382	0.323	0.823
Scheduled Caste	0.118	0.097	0.976
Scheduled Tribe	0.206	0.032	0.048
Minority	0.029	0.032	1.108
Religion			
Hindu	0.765	0.903	0.163
Muslim	0.176	0.097	0.464
Other	0.059	0.000	0.157
Highest education level achieved			
Class 12 and below	0.688	0.633	0.784
Tertiary	0.281	0.333	0.778
Other diploma/non-formal education	0.031	0.033	1.143
Work experience			
Any previous work experience	0.853	0.645	0.049
Number of years of work experience	9.2	8.7	0.891
Any previous working in a job related to TB	0.147	0.161	1.051
Any previous experience in the social/NGO sector	0.206	0.138	0.614
Any other income generating activity, in addition to OpASHA	0.088	0.129	0.874
Lives in one of the areas she covers	0.545	0.742	0.103
Household size	5.2	6.2	0.327
Lives alone	0.000	0.067	0.130
Assets			
Has electricity	1.000	0.933	0.112
Has tap water	0.594	0.400	0.127
Has a television	0.848	0.933	0.257
Has a refrigerator	0.545	0.467	0.613
Rents an apartment or house to a third party	0.118	0.207	0.272
Owns her house	0.559	0.581	1.106
Exposure to technology			
Knows how to use a computer	0.559	0.516	0.839
Knows how to use the internet	0.424	0.433	1.127
Has an email account	0.412	0.258	0.195
Has a social networking account	0.294	0.258	0.752
Days spent in the experiment	349.1	333.7	0.772
<i>Panel B. OpASHA center-wise data sheets</i>			
Reported baseline number of detections per month per area	2.965	2.393	0.075
Reported baseline fraction of defaults per month per area	0.029	0.039	0.513
<i>Panel C. Observation days</i>			
Center open	0.980	0.989	0.611
Total time center was open	307.4	311.8	1.117
Health worker present at the center	0.769	0.810	0.618
Total time health worker was present at the center, net of breaks	110.9	106.0	0.819
Visit by OpAsha	0.067	0.043	0.399
Adherence 1: Patient came to center and took or picked up pill	0.520	0.517	1.124
Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit	0.761	0.758	0.982

Notes : We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table I.

Table D2: Patient Summary Statistics (Pairs Cluster Bootstrap)

	Control group mean	Treatment group mean	P-value Treatment = Control
	(1)	(2)	(3)
Male	0.535	0.576	0.640
Age	34.0	34.3	0.940
Caste			
Does not know	0.040	0.026	0.642
General caste	0.160	0.165	0.717
Other Backward Class	0.376	0.352	0.768
Scheduled Caste	0.229	0.297	0.142
Scheduled Tribe	0.118	0.103	0.256
Minority	0.078	0.058	0.703
Religion			
Hindu	0.869	0.903	0.689
Muslim	0.112	0.081	0.774
Other	0.019	0.016	0.658
Literacy			
Cannot read or write	0.271	0.301	0.542
Can read but not write	0.021	0.026	1.034
Can read and write	0.708	0.673	0.494
Education			
Pre-primary	0.217	0.234	0.888
Primary	0.375	0.390	0.647
Secondary	0.367	0.334	0.612
Undergraduate and more	0.042	0.042	1.061
Household size	5.6	5.4	0.245
Lives alone	0.019	0.016	0.649
Assets			
Has electricity	0.968	0.935	0.151
Has tap water	0.520	0.418	0.638
Has a television	0.826	0.832	0.602
Has a refrigerator	0.276	0.230	0.551
Owns her house	0.716	0.738	1.120
Migration status			
Has always lived in the area	0.570	0.542	0.426
Has lived there for more than 6 years	0.209	0.258	0.066
Has lived there for less than 5 years	0.221	0.200	0.555
Is currently working	0.381	0.406	0.929
Time to go to the center (in minutes)	11.6	12.4	0.983
Distance to the center (in miles)	3.3	3.0	0.842

Notes : We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table II.

Table D3: Impact on Treatment Interruption and Adherence (Pairs Cluster Bootstrap)

<i>Panel A: Impact on treatment interruption, or default (source: patient surveys)</i>										
	Treatment complete		Treatment ongoing		Default		Died		Outcome unknown	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.003	-0.007	0.022	0.020	-0.018	-0.015	-0.001	0.000	0.000	0.003
P-value	0.948	0.846	0.510	0.560	0.037	0.068	0.815	0.988	0.993	0.807
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes
Observations	2802	2802	2802	2802	2802	2802	2802	2802	2802	2802
Mean in control group	0.507	0.507	0.299	0.299	0.073	0.073	0.016	0.016	0.105	0.105
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on treatment adherence (source: observation days)</i>										
	Adherence 1: Patient came to the center and took or picked up the pill		Adherence 2: Adh. 1 + Relative came to the center and picked up the pill, or patient or relative was given pill during a home visit		Compliance with treatment verified (as per adherence 1)		Compliance with treatment verified (as per adherence 2)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Treatment	0.139	0.122	0.073	0.056	0.039	0.038	0.024	0.023		
P-value	0.000	0.000	0.052	0.062	0.002	0.002	0.145	0.159		
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Patient controls		Yes		Yes		Yes		Yes		
Observations	12844	12844	12810	12810	3041	3041	3043	3043		
Mean in control group	0.542	0.542	0.692	0.692	0.046	0.046	0.078	0.078		
Number replications	5000	5000	5000	5000	5000	5000	5000	5000		
<i>Panel C: Impact on patient compliance (source: patient surveys)</i>										
	Occasionally sent someone else to get the pills		Ever picked up medicine for one week or more during one visit							
	(1)	(2)	(3)	(4)						
Treatment	-0.218	-0.218	-0.079	-0.078						
P-value	0.000	0.000	0.005	0.005						
Strata fixed effects	Yes	Yes	Yes	Yes						
Patient controls		Yes		Yes						
Observations	3352	3352	3295	3295						
Mean in control group	0.367	0.367	0.283	0.283						
Number replications	5000	5000	5000	5000						

Notes: We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table III.

Table D4: Impact on Detections and Patient Selection (Pairs Cluster Bootstrap)

Panel A: Impact on number of verified detections (source: program data cross-checked with patient surveys)

	Verified detections per center per month	
	(1)	(2)
Treatment	-0.103	-0.159
P-value	0.514	0.551
Strata fixed effects	Yes	Yes
Baseline detection	Yes	Yes
Health worker controls		Yes
Observations	1637	1637
Mean in control group	1.762	1.762
Number replications	5000	5000

Panel B: Impact on patient selection (patients enrolled after the beginning of the experiment)

	Control group mean (1)	Treatment group mean (2)	P-value Treatment = Control (3)	Number of obs. (4)
Male	0.586	0.576	0.475	2598
Age	34.4	33.1	0.168	2597
Caste				
Does not know	0.048	0.053	0.874	2602
General caste	0.181	0.179	0.747	2602
Other Backward Class	0.357	0.352	0.997	2602
Scheduled Caste	0.231	0.267	0.207	2602
Scheduled Tribe	0.078	0.070	0.367	2602
Minority	0.104	0.079	0.505	2602
Religion				
Hindu	0.832	0.870	0.405	2594
Muslim	0.147	0.117	0.523	2594
Other	0.021	0.013	0.269	2594
Literacy				
Cannot read or write	0.304	0.294	0.381	2602
Can read but not write	0.043	0.044	0.652	2602
Can read and write	0.653	0.662	0.527	2602
Education				
Pre-primary	0.247	0.226	0.138	2593
Primary	0.371	0.410	0.094	2593
Secondary	0.333	0.325	0.923	2593
Undergraduate and more	0.049	0.039	0.313	2593
Household size	5.5	5.4	0.224	2593
Lives alone	0.025	0.017	0.106	2602
Assets				
Has electricity	0.970	0.966	0.739	2532
Has tap water	0.562	0.544	0.909	2543
Has a television	0.819	0.822	0.696	2460
Has a refrigerator	0.264	0.258	0.794	2459
Owns her house	0.689	0.679	0.741	2538
Migration Status				
Has always lived in the area	0.507	0.457	0.305	2600
Has lived there for more than 6 years	0.227	0.269	0.145	2600
Has lived there for less than 5 years	0.266	0.274	0.869	2600
Is currently working	0.316	0.321	0.467	2388
Time to go to the center (in minutes)	11.0	11.8	0.330	1812
Distance to the center (in miles)	2.9	3.0	0.651	1998

Notes: We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table IV.

Table D5: Impact on Patients' Health (Pairs Cluster Bootstrap)

<i>Panel A: Impact on patients' physical health (source: patient surveys)</i>												
	Cured		Patients' ranking of their health		Index of no difficulty in performing activities		Index of health symptoms		Index of severe health symptoms		Index of increased health symptoms	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	0.015	0.012	0.012	-0.010	-0.022	-0.038	-0.008	0.001	-0.035	-0.026	-0.013	-0.011
P-value	0.605	0.663	0.912	0.922	0.491	0.247	0.832	0.984	0.233	0.344	0.523	0.601
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes		Yes		Yes		Yes		Yes		Yes
Observations	2395	2395	2430	2430	2458	2458	2459	2459	2459	2459	2458	2458
Mean in control group	0.400	0.400	7.431	7.431	--	--	--	--	--	--	--	--
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on patients' emotional health (source: patient surveys)</i>												
	Thinks that household situation will improve in the future		Has a project to improve their life and their family's life		Index of emotional distress		Overall somewhat or very happy		Overall satisfaction with their life			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
Treatment	0.000	0.002	-0.017	-0.017	0.010	0.010	-0.019	-0.020	-0.082	-0.086		
P-value	0.985	0.903	0.294	0.301	0.755	0.742	0.195	0.183	0.365	0.346		
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Patient controls		Yes		Yes		Yes		Yes		Yes		
Observations	2077	2077	2375	2375	2454	2454	2435	2435	2426	2426		
Mean in control group	0.890	0.890	0.120	0.120	--	--	0.942	0.942	7.259	7.259		
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000		

Notes : We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table V.

Table D6: Impact on Health Worker Effort (Pairs Cluster Bootstrap)

<i>Panel A: Impact on health worker attendance</i>														
	Center open		Total time center was open		Health worker present at the center		Total time health worker was present at the center, net of breaks		Visit by OpAsha		Center open		Health worker present at the center	
	<i>(source: observation days)</i>				<i>(source: random spot checks)</i>									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treatment	0.000	0.001	-10.7	9.2	0.019	0.045	20.2	22.3	0.033	0.030	0.001	0.005	0.049	0.071
P-value	0.929	0.806	0.515	0.631	0.533	0.145	0.044	0.054	0.020	0.047	1.128	0.851	0.403	0.068
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Observations	2,749	2,749	2,414	2,414	3,063	3,063	2,987	2,987	2,855	2,855	916	916	916	916
Mean in Control Group	0.998	0.998	338.9	338.9	0.834	0.834	117.3	117.3	0.054	0.054	0.945	0.945	0.598	0.598
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on health worker activities</i>														
	# home visits as % of all patients scheduled to come on that day		Intensity of default activities (index)		Intensity of detection activities (index)		Health worker gives advice related to TB		Frequency of interactions with health worker		Health worker supports patient during treatment		Health worker sometimes or often delivers pills at patient's home	
	<i>(source: observation days)</i>		<i>(source: health worker surveys)</i>				<i>(source: patient surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treatment	0.037	0.049	0.031	-0.103	-0.065	-0.071	0.014	0.014	-0.051	-0.045	-0.003	-0.004	0.073	0.087
P-value	0.178	0.189	0.678	0.420	0.292	0.395	0.470	0.454	0.514	0.538	0.948	0.926	0.121	0.036
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Baseline control			Yes	Yes	Yes	Yes								
Patient controls								Yes		Yes		Yes		Yes
Observations	2341	2341	134	134	134	134	3280	3280	3193	3193	3293	3293	3295	3295
Mean in control group	0.115	0.115	--	--	--	--	0.894	0.894	1.749	1.749	0.213	0.213	0.454	0.454
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000

Notes: We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table VI.

Table D7: Impact on Quality of Reporting on Default (Pairs Cluster Bootstrap)

	Actual likelihood to default		Defaults as reported in the program data		Default as reported in the government registers	
	<i>(source: patient surveys)</i>		<i>(source: program data)</i>		<i>(source: government registers)</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.018	-0.015	0.000	0.001	0.000	0.000
P-value	0.037	0.068	1.001	0.911	0.954	0.973
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Patient controls		Yes				Yes
Baseline outcome			Yes	Yes		
Health worker controls				Yes		
Observations	2802	2802	1832	1832	2913	2913
Mean in control group	0.073	0.073	0.045	0.045	0.021	0.021
Number replications	5000	5000	5000	5000	5000	5000
Test: Treatment = Treatment(1) (or (2))						
P-value	--	--	0.191	0.229	0.115	0.141
F-statistic	--	--				

Notes: We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Unlike in Table VII, we cannot compare treatment effects on actual likelihood to default and on default as reported in the program data and in the government registers using a seemingly unrelated regressions framework as the covariance matrix stored by the clusterbs Stata command cannot be used in any post-estimation inference (see the Stata help file for clusterbs). Other notes as in Table VII.

Table D8: Impact on Quality of Reporting on Patient Detection (Pairs Cluster Bootstrap)

<i>Panel A: Impact on number of verified and reported detections</i>						
	Verified detections <i>(source: program data cross-checked with patient surveys)</i>		Reported detections <i>(source: program data)</i>			
	(1)	(2)	(3)	(4)		
Treatment	-0.103	-0.159	-0.744	-0.795		
P-value	0.514	0.551	0.008	0.032		
Strata fixed effects	Yes	Yes	Yes	Yes		
Baseline outcome	Yes	Yes	Yes	Yes		
Health worker controls		Yes		Yes		
Observations	1637	1637	1832	1832		
Mean in control group	1.762	1.762	3.730	3.730		
Number replications	5000	5000	5000	5000		
Test: Treatment = Treatment(1) (or (2))						
P-value	--	--	0.012	0.057		
F-statistic	--	--				
<i>Panel B: Impact on survey outcomes, individual level (sources: program data, patient surveys)</i>						
	Patient verified (found and surveyed)	Survey not completed: Patient refused to answer	Survey not completed: Patient not found	Survey not completed: Health worker warning	Survey not completed: Patient death	Survey not completed: Other reasons
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.058	-0.020	-0.011	-0.015	-0.004	-0.008
P-value	0.049	0.014	0.541	0.203	0.634	0.493
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3647	3647	3647	3647	3647	3647
Mean in control group	0.685	0.068	0.106	0.059	0.031	0.049
Number replications	5000	5000	5000	5000	5000	5000
<i>Panel C. Impact on the fraction of patients reported in the government registers which are verified (sources: government registers, patient surveys)</i>						
	Fraction of verified patients (1)					
Treatment	0.075					
P-value	0.030					
Strata fixed effects	Yes					
Observations	2998					
Mean in control group	0.697					
Number replications	5000					

Notes: We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Unlike in Table VIII, we cannot compare treatment effects on the number of verified detections and reported detections using a seemingly unrelated regressions framework as the covariance matrix stored by the clusterbs Stata command cannot be used in any post-estimation inference (see the Stata help file for clusterbs). Other notes as in Table VIII.

Table D9: Sustainability of the Intervention (Pairs Cluster Bootstrap)

<i>Panel A: Impact on health worker salary and satisfaction</i>												
	Monthly compensation for detection activities		Monthly salary		# of challenges faced in getting a patient to complete treatment course		# of challenges faced in detecting patients		Excessive workload		Standardized job satisfaction index	
	<i>(source: salary data from OpAsha)</i>				<i>(source: health worker surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	-324.3	-224.9	-428.5	-341.1	-0.220	-0.238	0.013	-0.167	-0.105	-0.199	0.123	0.140
P-value	0.016	0.138	0.024	0.199	0.225	0.302	0.943	0.521	0.106	0.033	0.221	0.328
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes
Baseline control					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	693	693	936	936	121	121	117	117	118	118	132	132
Mean in control group	1321	1321	5791	5791	3.309	3.309	2.746	2.746	0.222	0.222	--	--
Number replications	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
<i>Panel B: Impact on patients' satisfaction with healthcare received</i>												
	Satisfied with their interaction with their health worker		Would refer a TB suspect to an OpAsha health worker		# of challenges faced by patients to complete treatment course							
	<i>(source: patient surveys)</i>				<i>(source: health worker surveys)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)						
Treatment	0.005	0.003	0.034	0.033	-0.053	-0.212						
P-value	0.735	0.818	0.177	0.182	0.728	0.496						
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes						
Patient controls		Yes		Yes								
Health worker controls						Yes						
Baseline control					Yes	Yes						
Observations	2378	2378	2441	2441	119	119						
Mean in control group	0.957	0.957	0.178	0.178	2.500	2.500						
Number replications	5000	5000	5000	5000	5000	5000						
<i>Panel C. Technology presence over time (source: observation days)</i>												
	Center has biometric equipment		Share of all devices (fingerprint reader, laptop, USB key) present		Share of all devices (fingerprint reader, laptop, USB key) present and working							
	(1)	(2)	(3)	(4)	(5)	(6)						
# weeks since beginning of the experiment	-0.002	-0.002	-0.006	-0.006	-0.006	-0.006						
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes						
Health worker controls		Yes		Yes		Yes						
Observations	1414	1414	1403	1403	1398	1398						
Mean in first month	0.921	0.921	0.909	0.909	0.899	0.899						
Number replications	5000	5000	5000	5000	5000	5000						

Notes: We use the pairs cluster bootstrap procedure proposed by Esarey and Menger (2019) to allow for correlation of the error terms at the level of the cluster and report the corresponding p-value. We use 5,000 bootstrap iterations. Other notes as in Table IX.