Stanford | King Center on Global Development

Data and Policy Decisions: Experimental Evidence from Pakistan

Michael Callen

Saad Gulzar

Syed Ali Hasanain

Muhammad Yasir Khan

Arman B. Rezaee

August, 2020

Working Paper No. 1077

NBER WORKING PAPER SERIES

DATA AND POLICY DECISIONS: EXPERIMENTAL EVIDENCE FROM PAKISTAN

Michael Callen
Saad Gulzar
Syed Ali Hasanain
Muhammad Yasir Khan
Arman B. Rezaee

Working Paper 27678
http://www.nber.org/papers/w27678

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2020

This paper combines two previous papers circulated with the titles “The Political Economy of Public Sector Absence: Experimental Evidence from Pakistan” and “Personalities and Public Sector Absence: Evidence from a Health Experiment in Pakistan”. We thank Farasat Iqbal for championing and implementing the project and Asim Fayaz and Zubair Bhatti for designing the smartphone monitoring program. Support is generously provided by the International Growth Centre (IGC) political economy program, the IGC Pakistan Country Office, and the University of California Office of the President Lab Fees Research Program Grant #235855. Callen was supported by grant #FA9550-09-1-0314 from the Air Force Office of Scientific Research. We thank Erlend Berg, Eli Berman, Leonardo Bursztyn, Ali Cheema, Melissa Dell, Ruben Enikolopov, Barbara Geddes, Naved Hamid, Gordon Hanson, Michael Kremer, Asim Ijaz Khwaja, Craig McIntosh, Ijaz Nabi, Aprajit Mahajan, Monica Martinez-Bravo, Benjamin A. Olken, Gerard Padro-i-Miquel, Karthik Muralidharan, Rohini Pande, Daniel N. Posner, Ronald Rogowski, Jacob N. Shapiro, Christopher Woodruff, Oliver Vanden Eynde, David Yanagizawa-Drott, Ekaterina Zhuravskaya and various seminar participants for insightful comments. Excellent research assistance was provided by Muhammad Zia Mehmood and Haseeb Ali. We thank Ali Cheema and Farooq Naseer for kindly sharing their data on election outcomes. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Data and Policy Decisions: Experimental Evidence from Pakistan
Michael Callen, Saad Gulzar, Syed Ali Hasanain, Muhammad Yasir Khan, and Arman B. Rezaee
NBER Working Paper No. 27678
August 2020
JEL No. C93,D02,D73,K42,O17

ABSTRACT

We evaluate a program in Pakistan that equips government health inspectors with a smartphone app which channels data on rural clinics to senior policy makers. The system led to rural clinics being inspected 104% more often after 6 months, but only 43.8% more often after a year, with the latter estimate not attaining significance at conventional levels. There is also no clear evidence that the increase in inspections led to increases in general staff attendance. In addition, we test whether senior officials act on the information provided by the system. Focusing only on districts where the app is deployed, we find that highlighting poorly performing facilities on a dashboard viewed by supervisors raises doctor attendance by 75%. Our results indicate that technology may be able to mobilize data to useful effect, even in low capacity settings

Michael Callen
London School of Economics
Houghton Street
London, WC2A 2AE
United Kingdom
and NBER
m.j.callen@lse.ac.uk

Saad Gulzar
Stanford University
616 Serra Street
Encina Hall West, Room 100
Stanford, CA 94305-6044
gulzar@stanford.edu

Syed Ali Hasanain
Lahore University of Management Sciences
Street 29 Lahore 54792, Pakistan
hasanain@lums.edu.pk

Muhammad Yasir Khan
University of California at Berkeley
2220 Piedmont Ave
Student Services Building 545
Berkeley, CA 94720
yasir.khan@berkeley.edu

Arman B. Rezaee
University of California, Davis
1149 Social Sciences & Humanities Building
One Shields Avenue
Davis, CA 95616
abrezaee@ucdavis.edu

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/1329

1 Introduction

Information technology is providing governments across the globe with greater access to data to inform policymaking. Technologies like smartphones and tablets make it simple and cheap to collect, compile, and visualize specialized and timely information relevant to a range of decisions. Increasingly, policymakers no longer need to rely on the collection and aggregation of geographically disparate paper records to understand how their government is operating. They can have this information instantly, and presented in the form that best suits their needs.

But will policymakers use these data? Will activating such technologies improve the quality of service provision? These are complex questions where the capabilities of government personnel, the specific government organization, and the broader political and institutional environment all potentially impact the answers.

To provide evidence on these questions, we conduct a randomized controlled evaluation of a smartphone monitoring program in Punjab, Pakistan. The program—officially termed 'Monitoring the Monitors'—equips government inspectors with a smartphone application that collects data and feeds it to an online dashboard system. This provides real-time information on rural public health clinics in Punjab, Pakistan, aggregated into simple charts and tables for the review of senior health officials.[1] It also includes several fail-safes to ensure accurate reporting: reports are geo-stamped and time-stamped and all staff reported present must be photographed with the inspector. In this environment, irregular inspections (in our baseline, only 23% of facilities had received their required monthly inspection) and doctor absence (doctors were present at 24% of facilities during regular operating hours in our baseline) are serious issues. The smartphone system supplanted the previous paper-based system for collecting operational data on public health facilities, which rarely functioned.

The evaluation spans 35 of 36 districts in Punjab.[2] Punjab is a province of 100 million

[1] The data include staff attendance, availability of medicine, patient visits, vaccines provided, cleanliness, and so on.

[2] One district was withheld to pilot the system.

people, with many citizens utilizing public health services. The experiment involved 117 inspectors, 35 senior officers, 2,496 rural health clinics (of which, we sample 850), and took place across 240 different parliamentary assembly constituencies. This scale allows us to cluster randomize at a relevant policy unit – the district.[3] This setting also provides variation to begin to speculatively examine whether individual and institutional constraints are relevant to the performance of the system.

Beyond improving the flow of data, the system also changed the behavior of inspectors, at least temporarily. In independent audits conducted six months after the survey, the inspection rate increased from 25.5% to 51.9% ($p < 0.01$). After a year of operation, inspection rates were 33.8% in the treatment districts and 23.5% in control districts, with this difference not being significant at conventional levels.[4]

Theoretically, if clinic staff base their beliefs about the likelihood they will be inspected on how often they are actually inspected, this increase in inspections could change staff attendance. However, the data generally do not indicate that increases in inspections resulted in better staff attendance. Of the seven categories of clinic staff, we only find highly speculative evidence that doctors may have increased attendance. However, this result is highly sensitive to the choice of specification, and should be interpreted accordingly.

We also built into the experiment a feature that allows us to examine whether providing senior officials with data changes their behavior.[5] Specifically, if more than three of the seven health workers that are supposed to staff a rural clinic are absent during a health inspection, we 'flagged' a facility as underperforming by highlighting it in red on the dashboard.[6] We

[3]Indeed, at the conclusion of the evaluation, the system was scaled to cover the province and continues to operate.

[4]The standard error of the difference clustered at the district level is 7.9 percentage points and so not significant at conventional levels. With 35 clusters, asymptotic reference distributions may not be valid, so we also use Fisher exact tests which do not appeal to asymptotic distributions. The Fisher exact p-value of the difference is 0.065.

[5]We do not directly observe corrective actions taken by senior officials. Rather, we look at whether doctor attendance improves when senior officials learn about absence, which is a possible outcome of corrective actions.

[6]We selected the threshold of three as this gave us the largest mass of inspections and so afforded the most statistical power.

test for effects on officials' behavior by examining whether doctor attendance increases in flagged facilities.[7] We find that flagging increases doctor attendance from 23.6% to 41.3% (standard error of difference = 8.2 percentage points), while it has no effect on the attendance of other, less senior, clinic staff. We conceive of this as creating a regression discontinuity that allows us to study whether data changes the behavior of senior policymakers. We discuss the identifying assumptions required for this to be causal and subject it to an extensive set of validity tests. We interpret this as evidence that policymakers use data when making decisions. We report results from a battery of robustness checking the causal interpretation of this result in 4.4.1.

Last, we investigate two dimensions of heterogeneity for flagging effects. First, we find that attendance responds more to flagging when the information is channeled to senior health officials with normatively better personality measures as measured using the Big Five Personality Test and the Perry Public Sector Motivation battery. Second, we find smaller effects in less politically competitive parliamentary constituencies. While speculative, these results suggest that both intrinsic and extrinsic forces are important factors in determining the extent to which senior bureaucrats might act on new information.

Our study principally concerns the potential for information technology and data to increase accountability and improve policy (Duflo et al., 2012; Blum and Pande, 2015; Callen and Long, 2015; Callen et al., 2016; Nealer et al., 2017), but relates to three additional literatures. The first regards whether monitoring of government workers improves attendance (Banerjee and Duflo, 2006; Banerjee et al., 2008; Chaudhury et al., 2006; Dhaliwal and Hanna, 2017; Muralidharan et al., 2019, 2020). The second broadly studies incentives in bureaucracies in developing countries (Ashraf et al., 2014, 2015; Bertrand et al., 2017; Dal Bó et al., 2013; Deserranno, 2017; Finan et al., 2017; Gulzar and Pasquale, 2017; Khan et al., 2016; Rogger and Rasul, forthcoming; Xu, 2017; Habyarimana et al., 2020). The final involves experiments at scale (Muralidharan et al., 2016; Muralidharan and Niehaus, 2017).

[7] We discuss instruments available to officials to encourage attendance in section 2.

The paper proceeds as follows. Section 2 provides background on the 'Monitoring the Monitors' program. Section 3 describes the data and the experiment. Section 4 provides results, and section 5 concludes.

2 Background

2.1 Public Health Services in Punjab

In Punjab, public health services are provided by the Department of Health, which is headed by the Secretary of Health. The smartphone intervention we describe in this paper works by facilitating flows of information through the existing chain of command, so describing that chain is fundamental to characterizing the reform.

This provincial Department comprises 36 District Health Departments, each headed by an Executive District Officer, hereafter referred to as a 'senior health official'. Senior health officials report directly to the Secretary. Performance by senior health officials is commonly rewarded with appointment to a higher office. Senior health officials are, in turn, each supported by several Deputy District Officers, hereafter referred to as 'health inspectors', typically one for each sub-district (there are, on average, 3.4 sub-districts per district). Figure 1 depicts this administrative hierarchy.

Health inspectors are charged with inspecting all of the health facilities in their sub-district at least once every month (see Appendix Section B for further details of health inspector duties). There are several ways that senior officials can compel or encourage inspectors to do their jobs. Conversations with several senior officials reveal that they typically begin simply by having a conversation with a problematic inspector. The next step is to refer the matter to a senior provincial-level official in charge of general administration. This can result in a formal inquiry and ultimately in pay cuts. Recent research in the same context provides direct evidence of how structured interactions between senior health officials and

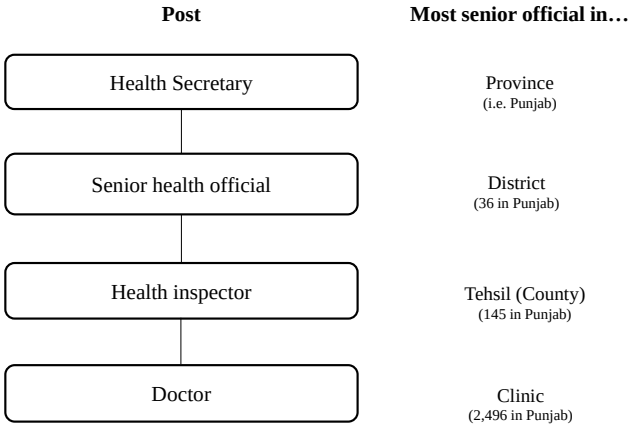


Figure 1: Organization Chart for the Punjab Department of Health

clinic staff can improve performance (Khan, 2020).[8]

While their primary job is to collect data, health inspectors do have the authority to directly punish absent clinic staff by issuing a 'show-cause notice', which requires staff to explain their absence to senior health officials. They can also suspend and deny pay to contract staff, including doctors. In severe cases of persistent absence, health inspectors can transfer staff to less-desirable locations, but they cannot terminate employment. Unlike their superiors, health inspectors rarely ascend to higher leadership positions.

Senior health officials do not typically sanction facility-level staff directly. Rather, they will send verbal communications through the chain of command via the inspector. The next step is to call for a formal explanation for absence. After that, the matter is referred to a provincial-level official who can recommend pay cuts or transfers.

There are five classifications of health facilities; we focus on the frontline tier, called Basic Health Units, hereafter referred to as 'health clinics'. Each health clinic is headed

[8]The exercise in this paper encourages frontline community health workers to watch a video of the senior health official describing the mission of the Health Department and requires them to take part in sessions reflecting on that mission. This is cross-randomized with a performance bonus. Both increase performance, though financial incentives are less effective in the presence of the mission treatment.

by a Medical Officer, henceforth 'doctor'.[9] These doctors are of particular interest for this study. Doctors are general practitioners who have completed five years of medical school, and are therefore the most trained health professionals in rural areas (see Appendix Section C for details on doctor hiring practices). While more senior doctors are paid more, they have essentially the same portfolio of responsibilities. Very few doctors rise through the ranks to become health inspectors: compared to the 2,496 Medical Officer posts in clinics, there are only 123 such senior positions.

2.2 Pre-existing Paper-Based Monitoring System

During their required monthly inspections, health inspectors are required to collect information on a standard paper form. This form records utilization, resource availability, and worker absence. We provide this form in Appendix E. Once collected, forms are brought to a central district facility, manually entered into a spreadsheet, and aggregated into a monthly report for senior health officials.

This inspection system affords only limited visibility into inspectors' activities to senior officials. Compounding this problem, senior health officials have only two weak means of sanctioning an inspector: issuing a verbal reprimand or, in serious cases, sending a written request for investigation to provincial authorities. The investigation process is long, highly bureaucratic, and, anecdotally, prone to interference by elected politicians.

2.3 'Monitoring the Monitors' Smartphone Monitoring Program

We partnered with the Department of Health to design and experimentally evaluate the 'Monitoring the Monitors' program. This program replaced the existing paper-based monitoring system with an Android-based smartphone application, which collects the same data as the paper forms and transmits them instantly to a central online dashboard for the Sec-

[9] Additional clinic staff include a Dispenser, a Lady Health Worker, a Health Technician, a School Health and Nutrition Supervisor, a Computer Operator, and a Midwife. Officially, health clinics are open from 8am to 2pm, Monday through Saturday.

retary of Health and senior health officials.[10] The dashboard provides summary statistics, charts, and graphs in a format designed in collaboration with senior health officials. Inspections are geotagged, timestamped, and require photos of the inspector and all health clinic staff marked present to check for reliability. The geotagging and timestamping features are designed to ensure inspectors visit health clinics, while the staff photos are intended to ensure that the digital reports of staff attendance are accurate.

Figure 2, Panel A, depicts the view of the dashboard that the Secretary of Health sees when first logging on. It presents a bar chart giving the number of health clinic inspections conducted in a district as a proportion of number of inspections assigned that month, allowing the Secretary to compare performance across districts. Panel B provides an alternate view available to senior health officials—a summary spreadsheet where each row corresponds to a different health clinic inspection that occurred in a senior health official's district.

3 Data and Experiment

3.1 Data

To measure the impacts of our smartphone monitoring on health clinic inspections and doctor attendance, we collected primary data on a representative sample of 850 (34%) of the 2,496 health clinics in Punjab.[11] All districts in Punjab except Khanewal are represented in our data.[12] To our knowledge, this is the first representative survey of health clinics in Punjab. Figure 3 provides a map of the health clinics in our experimental sample along with district boundaries.

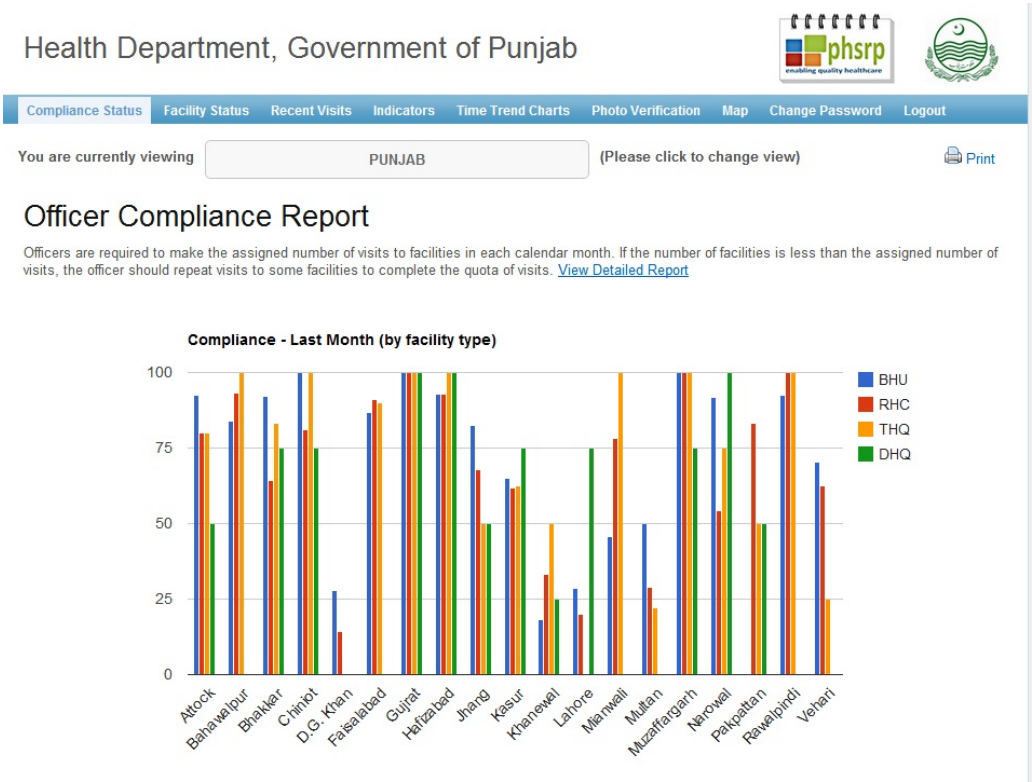
Enumerators made three unannounced visits to these 850 health clinics: one before smart-

[10] Appendix F provides the training manual for the mobile application provided to inspectors and Appendix G provides the training manual provided to senior health officials to assist them in using the dashboard.

[11] Health clinics were selected randomly using an Equal Probability of Selection (EPS) design, stratified on district and distance from the district headquarters. Our estimates of performance are thus self-weighting, and no sampling corrections are used in the analysis.

[12] The smartphone technology was piloted in Khanewal.

Panel A: Summary of Health Clinic Inspection Compliance by District



Panel B: Summary of Health Clinic Inspections within a District

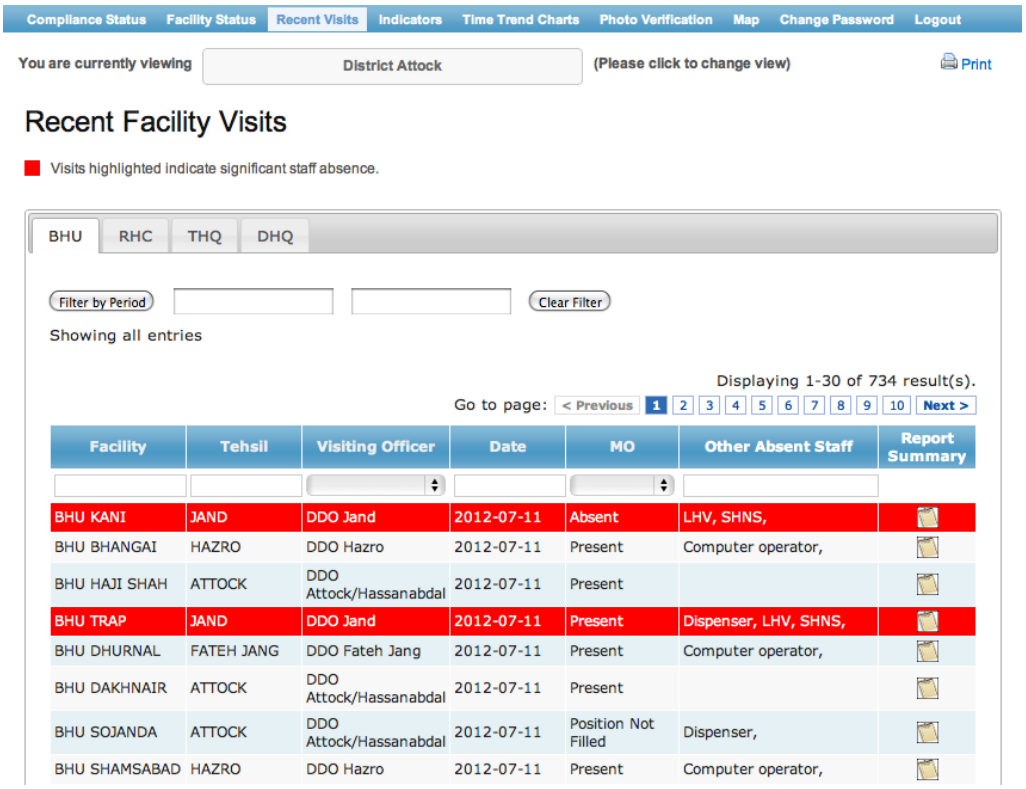


Figure 2: Online Dashboard Screenshots

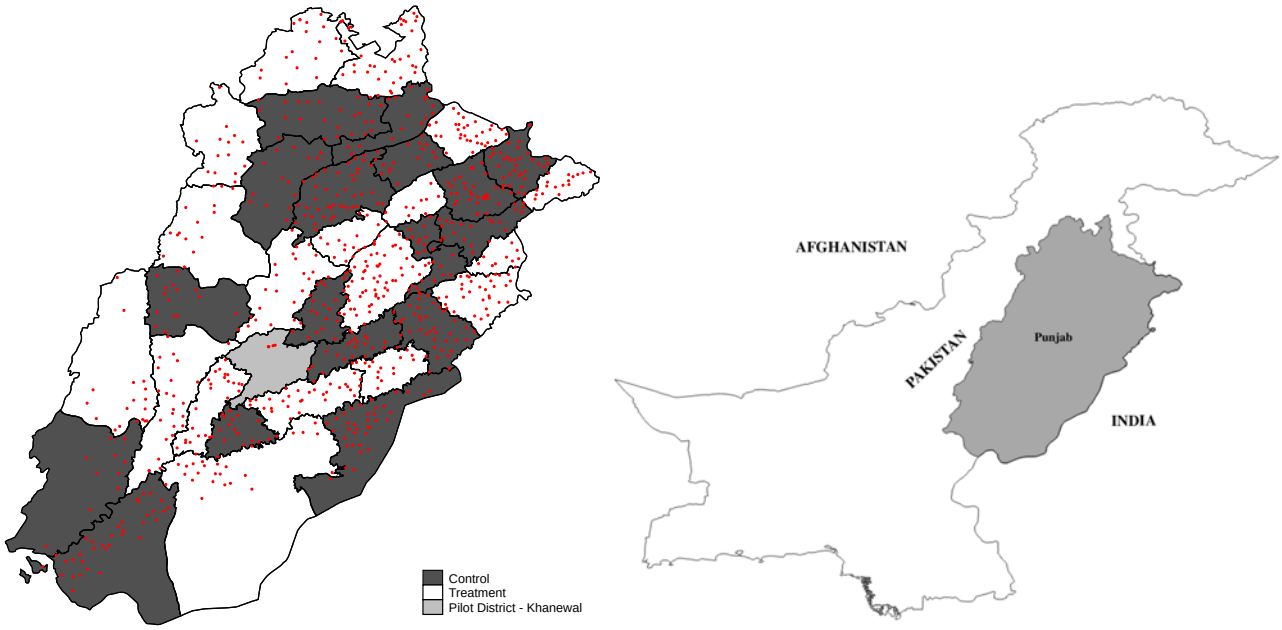


Figure 3: Locations of Health Clinics Surveyed

Notes: Borders demarcate districts in Punjab.

phone monitoring began, in November 2011, and two after smartphone monitoring began (in treatment districts), in June and October 2012.

During these unannounced visits, enumerators collected the same information that health inspectors record—information on health clinic utilization, resource availability, and worker absence—as well as information on the occurrence of health clinic inspections themselves. Enumerators physically verified health clinic staff presence, filling out an attendance sheet at the end of their visit and in private for doctors as well as dispensers, lady health workers, health/medical technicians, school health and nutrition supervisors, and midwives.[13] Summary statistics from unannounced visits at baseline are presented in Appendix Table A1.

Health inspectors record visits by signing paper registers maintained at the health facility. Enumerators measured whether a health inspection occurred in the prior month by interviewing facility staff and verifying the register record. In some cases enumerators were

[13]Doctors are officially required to be present and see patients at the health clinic. An unannounced visit therefore captures the official work assigned to doctors. We did not capture data on computer technicians as they are rarely assigned.

unable to confidently verify whether or not an inspection had occurred in the prior month. We treat such cases as missing data for analysis and verify in Appendix Table A1 (considering a dummy variable equal to one if the enumerator could not verify the last health inspection) that such cases do not correlate with treatment assignment at baseline and in Appendix Table A2 that we do not have treatment-driven attrition wave by wave. Additional information about data collection can be found in Appendix Section D.

We also conducted face-to-face time use surveys with all health inspectors in Punjab between February and March 2013.

3.2 Experiment

Our experimental sample comprises 35 of the 36 districts in Punjab. We randomly implemented the smartphone program in 18 of the 35 remaining districts. We randomized at the district level for two reasons. First, the intervention channels information about health inspections to district-level senior health officials. Second, all inspectors in a district are required to attend monthly meetings and so interact frequently, while these relations are much weaker across districts. District-level randomization therefore makes sense in terms of the design of program and also reduces concerns about contamination.

We stratified treatment on baseline health clinic staff attendance, the number of clinics in a district (to ensure a roughly even number of clinics in treatment and control), and whether or not the district was being run by the World Bank-led Public Sector Reform Program (PSRP). Figure 3 depicts control and treatment districts. We then re-randomized to achieve balance on a set of variables related to health clinic staff attendance, the frequency of health inspections, and the quality of service provision.[14]

[14]Specifically, we randomized using the 'big stick' approach, whereby we redrew our treatment assignment until the minimum p-value from difference in means tests between health clinics in treatment and control districts for a pre-specified set of balance variables was greater than some threshold. In our case we selected a threshold of 0.21. We balanced on the share of assigned health clinic staff who were present during the baseline unannounced visit, whether the health clinic had been inspected in the previous month by its health inspector, whether the health clinic had been inspected in the previous month by its senior health official, the number of antenatal visits recorded on the health clinic register in the previous month, the log of the

Appendix Table A1 reports balance between clinics in treatment and control districts at the baseline. The only measure showing imbalance is an indicator for doctor presence during our unannounced visit.[15] While stratifying on the share of staff present at baseline achieved balance for five of the six categories of staff that are supposed to be present at health clinics, it did not do so for doctors. We therefore use a difference-in-differences specification to estimate impacts of the program on doctors.[16]

4 Results

We now present results from our experimental evaluation of the 'Monitoring the Monitors' program. We estimate treatment effects on health inspection rates and on health inspector time use with the following specification:

$$Y_{itds} = \alpha + \beta Treatment_d + \delta_t + \gamma_s + \varepsilon_{itds}, \tag{1}$$

where Y_{itds} is either a dummy equal to one for health clinics inspected in the previous month or a measure of health inspector time use. i refers to the clinic, t to the survey wave, d refers to the district, and s to the randomization stratum that the district was in. $Treatment_d$ is a dummy variable equal to 1 for treated districts, δ_t are survey wave fixed effects, and γ_s are randomization strata fixed effects. These regressions use only post-treatment data (survey waves 2 and 3). We cluster all standard errors at the district level.

We estimate treatment effects on health clinic staff attendance and assignment using three specifications, all meant to account for baseline imbalances in our primary outcomes (which are only present and statistically significant for our measure of doctor absence). First,

number of polio vaccines administered at the health clinic in the previous month, whether the health clinic's doctor claimed a connection to their local parliamentarian, the tenure of the health clinic's doctor, the log of the population in the health clinic's catchment area, and the log of the distance of the health clinic to the district's headquarters.

[15] Treatment districts at baseline have 17.1 percentage point higher doctor attendance (p-value = 0.003).

[16] While it would be ideal, we cannot check for balance in pre-trends in doctor attendance, a necessary assumption for our difference-in-differences estimation to be unbiased, because we only have one pre-treatment observation for each clinic.

we use a difference-in-differences specification:

$$Y_{itds} = \alpha + \beta_1 Treatment_d + \beta_2 Post_t + \beta_3 Treat_d * Post_{dt} + \delta_t + \gamma_s + \theta_i + \varepsilon_{itds}, \quad (2)$$

where Y_{itds} is now either a dummy for whether a health clinic staff member is present during our announced visits or a dummy for whether a staff member is currently assigned to work at a health clinic at the point of our unannounced visits, and δ_t are survey wave fixed effects, γ_s are randomization strata fixed effects and θ_i are clinic fixed effects. Second, we use our primary treatment effect specification, equation 1, with adding Y_{i0ds}, the baseline level of the outcome, to the set of covariates in the regression. Third, we estimate:

$$Y_{itds} = \alpha + \beta Inspected_d + \delta_t + \gamma_s + \varepsilon_{itds} \quad (3)$$

where we instrument for whether a facility was inspected in the prior month using our treatment assignment.[17]

For all of our primary treatment effects, we also present results from regressions with additional controls and/or sample selection criteria as robustness checks. We point out any cases where results are sensitive to such decisions.

4.1 Approach to Inference

With only 35 districts in our sample, the asymptotic reference distributions for our test statistics may be invalid. We therefore report Fisher exact p-values (Fisher, 1935) which do not require a limiting distribution (Gerber and Green, 2012). This test assumes a null of no treatment effect for any unit. We perform this test by creating a set of artificial treatment assignments that satisfy the balancing requirements of the assignment protocol for actual treatment. For each treatment assignment, a corresponding artificial treatment effect is

[17] Tables A4 and A5 therefore report the first stage of this regression.

generated. The effect estimated using the actual treatment assignment is then compared against the 1,000 artificial treatment effect estimates. The p-value is the share of artificial treatment effects that have a larger magnitude than the actual treatment effect.

4.2 Impact on Health Inspectors

Table 1 presents estimates of the program's impact on the rate of inspections. We find that health clinics in treatment districts were 18.1 percentage points more likely to be inspected in the previous month during the treatment period. This represents a 74 percent increase in inspection rates in treatment districts relative to control districts. Breaking this up into the two waves of post-treatment data collection, we find comparable effects, though there is evidence that the effect of treatment had attenuated by October 2012, a year after the introduction of the program.[18] In Table 1 we report simple means to facilitate inference on both means directly and on comparisons both between groups and across time. Appendix Tables A4 and A5 report corresponding treatment effect estimates from a range of specifications (including with strata fixed effects) which are consistent with the raw means, except, importantly, the wave three treatment effect is not significant at conventional levels with both randomization strata and wave fixed effects included (exact test p-value of 0.19).

We examine whether the additional time for inspections come at the costs of other tasks. Appendix Table A3 reports the monitoring program's impact on the time use of health inspectors. We do not find significant evidence that health inspectors in treatment districts are spending less time on other tasks after treatment, while we do find significant increases in time inspecting health clinics. However, we treat these results speculatively, as we only have a sample of 117 inspectors and a noisy measure of time use. We do, for example, find a negative but insignificant coefficient on time spent monitoring hospitals that could account for 60 percent of the increased time monitoring clinics.

[18]The p-value corresponding to the test of equal treatment effects in wave 2 and in wave 3 is 0.08.

Table 1: The Effect of Smartphone Monitoring on the Rate of Inspections

	Treatment (1)	Control (2)	Difference (3)	P-value Mean Diff. (4)	P-value Exact Test (5)
Facility Inspected in the Previous Month (=1)	0.426	0.245	0.181	0.010	0.001
	(0.048)	(0.046)	(0.066)		
# of Observations	759	764			
Facility Inspected (=1), Wave 2 Only (June 2012)	0.519	0.255	0.264	0.002	0.003
	(0.063)	(0.048)	(0.079)		
# of Observations	366	373			
Facility Inspected (=1), Wave 3 Only (October 2012)	0.338	0.235	0.103	0.200	0.065
	(0.053)	(0.059)	(0.079)		
# of Observations	393	391			

Notes: This table reports unconditional average treatment effects of the 'Monitoring the Monitors' program on the rate of health clinic inspections. The unit of observation in is the health clinic. The data come from primary unannounced surveys after the treatment was launched (wave 2 and 3). The dependent variable is an indicator variable that equals 1 if an inspector visited a clinic within a month prior to the survey, and 0 otherwise. The regression reports differences between treatment and control clinics. P-values reported in column (4) are for the difference in mean between columns (1) and (2) (i.e. that the treatment had no impact). Column (5) reports the Fisher Exact p-values. Standard errors clustered at the district level are reported in parentheses. Results conditional on randomization strata fixed effects, using a difference-in-differences specification, with and without baseline controls, can be found in Appendix Tables A4 and A5.

4.3 The Impact on Health Clinic Staff

Our estimates indicate that the program increased health inspections; this could increase health clinic staff attendance (an explicit goal of the program). We report estimates of program impact on staff attendance in Table 2. Panels A and B report impacts using a difference-in-differences specification. Panel C reports results using an average treatment effects regression controlling for baseline attendance.[19] Panel D reports a two-stage least squares estimate of the impact of a health facility being inspected on staff attendance, instrumenting for inspection with treatment.[20] We find mixed results here—we find no impacts of monitoring on staff assignment in Panels A and B and we find large, positive, and significant impacts on doctor attendance in Panels C and D. These results are not robust

[19]We do not present results not controlling for baseline attendance as we have a large imbalance in doctor attendance between treatment and control districts at baseline, as reported in Appendix Table A1.

[20]In this setting, inspections could impact staff attendance through the following channels. First, staff see an inspection take place directly, and therefore shift their beliefs about the probability of an inspection in the future. Second, they learn that an inspection took place through colleagues indirectly, and shift their beliefs. Third, they learn that an inspection took place because they are contacted about their absence during the inspection by a supervisor. There was no effort to publicize the reform to clinic staff and we expect that most staff learned about the reform when the inspector arrived at their facility with a smartphone and compelled all present staff to take a picture in front of the facility to confirm the attendance data. If the program affects attendance through channels other than inspections, then the exclusion restriction is not satisfied.

across specifications, and should be viewed accordingly.[21]

In contrast to doctors, other staff do not appear to respond to treatment regardless of specification. Speculatively, there are at least three reasons for this. First, attendance rates for clinic staff who are not doctors, with the exception of school and health nutrition supervisors, are substantially higher, and so there is less room for improvement. Second, doctors are also in charge of facilities so might be more responsive to the changes in incentives created by the program. Third, some staff, like the School Health and Nutrition Supervisor and Lady Health Workers carry out most of their work outside of clinics, and so may be less affected by these reports.

Note our measures of staff attendance do not condition on staff assignment to facilities as staff assignment itself could be impacted by treatment. We test for treatment effects on doctor assignment directly in Appendix Table A8, with the same set of panels as our attendance results. We again find mixed results, with weak evidence that doctor assignment is indeed higher, and by a similar magnitude as the increase in doctor attendance, in treatment districts. This would suggest that the mechanism for increased doctor attendance is through senior health officials assigning more doctors to work in rural health clinics. This is very plausible, possibly due to increased monitoring of senior health officials' district performance through the web dashboard or simply due to these officials having easier access to information about doctor assignment in their district.[22]

[21] Appendix tables A6 and A7 present results on attendance by survey wave. While our results become more imprecise, we do not see qualitatively different results across waves. This loss of precision also means our first-stage becomes too weak to conduct IV analysis by wave.

[22] We attempted to obtain detailed doctor assignment records from the Health Department on multiple occasions and it was clear there was no regular system for keeping track of doctor assignment during the period of our experiment. Hence, we are not able to conclusively identify what is driving the increased assignment of doctors. It is possible that this increase is achieved by moving doctors from control to treatment districts. However, as described in Appendix Section C, it is unlikely that doctors are moved across districts under ordinary circumstances.

Table 2: The Effect of Smartphone Monitoring on Clinic Staff Attendance

	Doctor (1)	Dispenser (2)	Lady Health Worker (3)	Health Tech (4)	School Health Nutri. Supervisor (5)	Midwife (6)
Panel A: Difference in Differences with Survey Wave FE						
Monitoring	-0.021	-0.067	-0.080	0.032	0.011	0.055
	(0.043)	(0.061)	(0.053)	(0.058)	(0.072)	(0.065)
	[0.613]	[0.970]	[0.997]	[0.491]	[0.522]	[0.446]
Mean in Controls	0.227	0.736	0.594	0.413	0.318	0.535
# Districts	35	35	35	35	35	35
# Observations	2422	2457	2457	2458	2448	2448
R-Squared	0.081	0.062	0.034	0.047	0.040	0.052
Panel B: Difference in Differences with Survey Wave + Facility FEs						
Monitoring	-0.010	-0.067	-0.075	0.028	0.013	0.055
	(0.044)	(0.059)	(0.054)	(0.059)	(0.072)	(0.064)
	[0.645]	[0.968]	[0.942]	[0.465]	[0.701]	[0.386]
Mean in Controls	0.228	0.736	0.596	0.413	0.318	0.537
# Districts	35	35	35	35	35	35
# Observations	2408	2446	2446	2447	2437	2437
R-Squared	0.533	0.464	0.447	0.552	0.398	0.512
Panel C: Control for Baseline Attendance with Survey Wave FE						
Monitoring	0.102**	-0.010	-0.037	-0.006	0.008	-0.040
	(0.040)	(0.040)	(0.034)	(0.041)	(0.046)	(0.067)
	[0.096]	[0.837]	[0.871]	[0.743]	[0.447]	[0.796]
Mean in Controls	0.215	0.746	0.601	0.427	0.301	0.533
# Districts	35	35	35	35	35	35
# Observations	1548	1579	1579	1581	1568	1566
R-Squared	0.126	0.076	0.036	0.115	0.073	0.114
Panel D: IV with Control for Baseline Attendance and Survey Wave FE						
Inspected	0.640**	-0.100	-0.310	-0.046	0.020	-0.223
	(0.256)	(0.251)	(0.205)	(0.258)	(0.283)	(0.383)
	[0.022]	[0.593]	[0.986]	[0.588]	[0.517]	[0.706]
Mean in Controls	0.275	0.765	0.571	0.422	0.304	0.521
# Districts	35	35	35	35	35	35
# Observations	1437	1468	1468	1470	1458	1458

Notes: This table reports average treatment effects of the 'Monitoring the Monitors' program on the staff attendance. The unit of observation is the clinic, and data come from primary unannounced surveys after the treatment was launched. The dependent variable is an indicator for whether a doctor was present at the clinic during an announced visit. All models include randomization block fixed effects. Standard errors clustered at the district level are in parentheses. Square brackets report the Fisher Exact p-values. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.4 Do Policy Makers Use Data on Staff Attendance?

In the 18 districts of Punjab that received the 'Monitoring the Monitors' treatment, the system aggregated and presented inspection report data to senior health officials through an online dashboard. This dashboard was also visible to the Health Secretary and the Director General of Health for Punjab.

To test whether senior health officials would act on these data, we introduced a manipulation to the dashboard that made certain health clinic inspection reports salient. Specifically,

we highlighted in red (hereon 'flagged') inspection reports that reported three or more staff (of seven generally) as absent during a health inspector's visit to the clinic.[23] Figure 2 Panel B provides an example of a dashboard view with some facilities flagged in red. We then test for changes in senior health official action through subsequent staff attendance.

The exact formula for this arbitrary threshold was not known to anyone but the research team. This approach creates a sharp discontinuity and permits measurement of the impacts on subsequent staff attendance. Our identifying assumption requires that facilities just below the cutoff (those with two staff absent during a health inspector's visit) and facilities just above the cutoff (those with three staff absent) share potential outcomes in the absence of the flagging. There are several reasons that this identifying assumption may not hold. For example, there could be mean reversion. If facilities with two staff absent revert to the mean of having three staff absent, or if facilities having three staff absent revert to a mean of having two staff absent (or both), our specification would pick this up as an effect of flagging. In Section 4.4.1 we report four tests of this identifying assumption, all of which are consistent with a causal interpretation of our estimates of the impact of flagging on subsequent attendance.

Our data have limitations, however, which we also acknowledge in Section 4.4.1. Perhaps the greatest limitation is the fact that the data are limited to treatment districts and furthermore to facilities that had a health inspection (and thus flagging or not flagging) within a window before one of our survey visits to measure attendance. Thus, while we can conduct some placebo tests to test whether our flagging was 'as-if' random for those facilities right below and above the cutoff, we are limited in how well we can verify parallel pre-trends and control for facility absence history. Results should be interpreted accordingly.

More specifically, we examine whether this manipulation within treated districts affected

[23] We deliberately selected three as the cutoff for flagging as the majority of reports indicated that two or three staff were absent, affording the greatest statistical power.

subsequent doctor absence in our primary data with the following specification:

$$Present\ Survey_{jt} = \alpha + \beta_1 Flagged_{jt-1} + \delta_t + \eta_{jt} \tag{4}$$

$Present\ Survey_{jt}$ is a dummy variable equal to 1 if the doctor j was absent during an unannounced visit by our enumerator in wave t, $Flagged_{it-1}$ is a dummy variable that equals 1 if the facility was flagged in red on the dashboard in a window of time prior to the primary survey wave t. For our primary analysis, we restrict to those facilities with either two or three staff absent, those on either side of the sharp discontinuity, though we report all possible bandwidths in Appendix Table A9. The estimate of the flagging effect is not significant at conventional levels for bandwidths that include data away from the discontinuity.

Selecting the window for which $Flagged_{it-1}$ is non-missing (i.e. 0 or 1) involves trade-offs and provides us substantial discretion in the analysis. We expect it would take at least a few days for senior officials to acknowledge and act on data from the dashboard. Practically, they need to receive the report that a clinic is underperforming, and then communicate their dissatisfaction to clinic personnel. Senior officials have a number of tools to reprimand clinic staff, with verbal warnings being by far the most common. Staff then need to react and change their behavior. This process will require at least a few days to play out, but should not take more than a few weeks. Also, practically, if the window for which $Flagged_{it-1}$ is defined is too long, virtually every facility will become flagged, limiting the variation with which to estimate effects. For transparency, we present estimates of the effect on doctor attendance for a broad number of potential time windows.

Specifically, we run the regression from equation (4) 750 times, varying the window for which we define a clinic as flagged prior to a primary unannounced visit to a clinic along two dimensions—we vary the length of the window being used along the x-axis and the delay from when a clinic is highlighted in red to when the window begins along the y-axis (so for example, a length of 25 and delay of 15 corresponds to considering a clinic as flagged if it

Table 3: Effect of Flagging Underperformance on the Dashboard

	Doctor	Dispenser	LHW	Health Tech	SHNS	Midwife
	(1)	(2)	(3)	(4)	(5)	(6)
Flagged	0.177**	0.076	-0.115	-0.066	-0.037	-0.125
	(0.082)	(0.083)	(0.090)	(0.092)	(0.093)	(0.091)
Unflagged Mean	0.236	0.676	0.703	0.405	0.324	0.568
# Clinics	112	116	116	116	116	116
# Reports	130	136	136	136	136	136
R-Squared	0.298	0.258	0.190	0.264	0.176	0.242
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports on the effect on subsequent doctor attendance of flagging on an online dashboard the fact that a clinic had three or more staff absent to a senior policymaker. Clinics were flagged in red on an online dashboard if three or more of the seven staff were absent in one or more health inspections of the clinic 11 to 25 days prior to an unannounced visit by our survey enumerators. The data sample limits to facility reports in which either two or three staff were absent (the threshold to trigger the underreporting red flag). In addition, the sample in all columns is limited to 'Monitoring the Monitors' treatment districts due to the necessity of the web dashboard for flagging clinics. All regressions include survey wave fixed effects. LHWs are Lady Health Workers and SHNS are School Health and Nutrition Supervisors. Standard errors clustered at the clinic level are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

was highlighted in red anytime 15 to 40 days prior to an unannounced visit).[24]

We observe a positive, robust, and significant treatment effect of flagging on doctor attendance across a wide range of windows, depicted in the top left panel of Appendix Figure A1. At our preferred window,[25] reported in Table 3, doctor attendance subsequent to flagging on the dashboard increases by 17.7 percentage points or about 75 percent.[26] Conversations with government partners suggest that the most likely driver of this effect is verbal reprimands from senior health officials to doctors in charge of clinics.

Like our primary treatment results on staff attendance, attendance for other staff does not respond to dashboard flagging. This is likely due to the same reasons cited above: attendance for other staff is much higher at baseline compared to doctors; doctors are also in charge of

[24]This analysis necessarily involves testing the same null of $\beta_1 = 0$ in many closely related specifications. We do not adjust inference for multiple tests.

[25]We chose this window to allow for the interpretation of a specific point estimate. It is a window we find plausible and it has a coefficient and significance level that is in the middle of those reported in Appendix Figure A1.

[26]Note that this positive result cannot be directly compared to the estimated average treatment effect of 'Monitoring the Monitors' on doctor attendance. To detect the effect of flagging we are limiting our sample to treatment facilities only (those that could be flagged on the dashboard) that were right below or above the staff absence cutoff for flagging.

the facilities so it makes sense that they are the ones to respond to facility monitoring; and some staff, like the School Health and Nutrition Supervisor and Lady Health Workers, have their primary job outside of the specific facility where attendance was audited.

4.4.1 Validity Tests for 'Flagging' Results

This section presents four validity tests for the flagging results. First, an alternative explanation for the effect of the dashboard could be simple mean reversion. If absence is unusually high in one month, it might drop in the next month because of mean reversion, whether or not it was flagged. We check this in Appendix Figure A2 where we use placebo thresholds between one and two and three and four absences. We find no evidence of mean reversion at placebo thresholds. Note that the first placebo test compares facilities recording two absences (the actual control group) against the group recording one absence (who are also not flagged on the dashboard). The second placebo test involves comparing facilities recording three absences (the actual 'treatment' group) against those with four recorded absences (who are also flagged on the dashboard). If a tendency for facilities with either two or three absences to revert to the mean were driving the results, with no role for flagging, we would expect the first placebo test to yield a negative estimate, and the second placebo test to yield a positive estimate. In the small number of cases where our placebo regressions find significance, the signs are the opposite of what mean reversion would predict (and consistent with a true positive effect at the 2/3 cutoff).

Second, we present in Appendix Figure A3 a placebo test in which we test whether flagging on the dashboard predicts increases in attendance recorded in survey visits prior to the flag. That is, we define a facility as flagged if it was highlighted red on the online dashboard in varying windows *after* our survey visits. It is not possible for dashboard flags to causally impact outcomes in the past. This exercise therefore tests for a time series process whereby either facilities with three absences tend toward a mean of having only two absences, or for a process whereby facilities with two absences tend toward a mean of having three

absences (or both). Again, we find no clear evidence of this pattern of mean reversion. An important caveat, however, is that these tests have low power. Using the analogous placebo window to our preferred specification above, for example, while the coefficient on the flagging placebo is -0.004, its 95% confidence interval is from -0.25 to 0.25.

Third, Appendix Table A10 presents two results controlling for the possible time dynamics of inspections, flaggings, and doctor attendance. For example, we might think a facility that is flagged multiple times in a row sees different outcomes than one that is flagged and then not flagged or flagged and then not inspected. Column 1 repeats our main flagging result. Then, in column 2, we restrict the sample to cases in which the flagging that occurred prior to our survey visit was the first time a facility was ever flagged. Treatment in this sample is straightforward to interpret since it does not require accounting for prior flags.[27] Taking this approach shrinks the sample as it removes facilities that were flagged multiple times before our survey visit. Estimates in this much smaller sample comprising only 130 reports are similar in magnitude to our result in column 1, suggesting doctors respond to flagging even when it is the first time. To additionally control for dynamic effects, in column 3, we add fixed effects for the number of prior inspections in our treatment period (all of which must have been cases where two or less staff was absence for the flagging visit to be the first). Results are again similar. While this does not rule out dynamic effects, it suggests flagging can have an effect for facilities with a range of inspection histories.

Fourth, a standard practice to check the validity of regression discontinuity designs is to examine whether fixed or predetermined variables are smooth across the discontinuity (Calonico et al., 2014). The analogue here is to check for balance on predetermined variables between facilities with either three absences or two absences. Appendix Table A11 presents balance across facilities in our main specification in Table 3 coded as flagged versus facilities

[27]In general, at any given point in time, different facilities will have been flagged different numbers of times. With sufficient data, flexibly controlling for the comprehensive treatment history would allow recovery of the causal (and recovery of heterogeneous treatment effects given different histories). With only three survey visits and being restricted to the smartphone monitoring districts, we lack to the data to pursue this approach.

coded as unflagged. We do find imbalance on four of 18 variables at the 10% level, though not in baseline doctor attendance, our primary outcome variable.[28] Finally, in Appendix Table A12 we re-run our primary flagging result controlling for all variables for which there is baseline imbalance, and find that results are similar.

Fifth, it is worth noting that flagging effects materialize approximately 10-20 days after the reports are first filed on the online dashboard. Throughout the project, we interacted with bureaucrats using the dashboard. They indicated very consistently that they would review the dashboard approximately once a week. A 10 day response is consistent with a manager checking about once a week and then demanding action from their subordinates. We therefore did not expect that attendance would respond immediately. That we only find effects in an intermediate window of a reasonable length is consistent with our explanation of how the dashboard was used: actions by senior officials and subsequent responses by doctors require time rather than through simple mean reversion. These results can be found on the top left panel of Appendix Figure A1.

4.5 Heterogeneity of Results

Whether increased monitoring and data affect service delivery reasonably depends not only on the abilities of government personnel but also on the broader political environment. We now examine two sources of heterogeneity, taking advantage of the scale of our experiment.

First, a growing literature documents the role of individuals in public service delivery across the developing world, summarized in Finan et al. (2017). We examine whether personality characteristics of the health workers in our study predict their response to monitoring and to information. To do this, we measured personality characteristics—the Big Five Personality Index and the Perry Public Service Motivation Index—of all of the doctors in our sample clinics and the universe of health inspectors and senior health officials in Punjab.

[28] Identification, in this case, only requires parallel trends prior to flagging. If we had reliable time series data on attendance prior to flagging, we could test this directly. Unfortunately, we only have at most two pre-flagging survey audits.

In Callen et al. (2017a) we explore this dimension of heterogeneity in great detail. As an example of the role of personality characteristics in this setting, we present heterogeneity of our flaggings results in Table A13. We find that personality characteristics systematically predict responses by senior health officials to our dashboard experiment, as measured by future doctor attendance in flagged facilities: a one standard deviation increase in the Big 5 Personality Index of senior bureaucrats increases the effect of flagging on the likelihood that doctor is present at the clinic during a subsequent unannounced visit by 28 percentage points. We see a similar, though smaller and statistically insignificant, point estimate for the interaction with Public Service Motivation. Of course, these results are merely predictive as we do not have identifying variation in personality traits.

Second, we explore whether measures of political competition predict flagging effects. The bureaucrats we worked with to create the program felt strongly that the program would break down when politicians interfered with senior officials' attempts to sanction their subordinates. Indeed, in our surveys, senior health officials report that politicians routinely interfere in this way.[29] In Callen et al. (2017b) we match each clinic in our sample to a provincial assembly constituency and examine in detail the extent to which these political moderators affect the efficacy of the Monitoring the Monitors program. As an example, we find that the treatment effect of flagging on doctor attendance varies by the degree of competitiveness in the previous election. Table A14 presents the results. We find that while flagging increases subsequent attendance by 35.9 percentage points in the most politically competitive third of constituencies, flagging has no apparent effect in the least competitive third. In addition, flagging works better on doctors who do not report a direct connection with a local politician. Indeed, the point estimates, though noisy, suggest the program may have negatively affected attendance of connected doctors.

[29]Based on interviews with all senior health officials in Punjab, we find that 44% report a politician interfering in their decision to sanction an underperforming employee during the previous year.

5 Discussion and Conclusion

A fundamental objective of policy research is to convey facts and data to policy makers. We find that providing senior health officials with information on low staff attendance causes them to take corrective action, indicating that data can change how policy makers behave, even in settings with weak institutions. This suggests two additional general lines of inquiry.

First, our test of whether information affects policy decisions is direct. For our senior officials, who are tasked with responsibilities that affect millions, ensuring the functioning of frontline service facilities is a priority. A natural question is whether we would see the same response to a more complicated object, like causal estimates of program effect, cost-benefit analyses, or more general forms of research evidence.[30] The 'Monitoring the Monitors' program provides proof-of-concept that technology can mobilize data to real effect. This also suggests, but by no means proves conclusively, that the massive investments being made by governments, technologists, researchers, philanthropists, and aid organizations to promote evidence-based policy can make a difference in developing countries.

Second, whether data and evidence impact policy likely depends on the characteristics of policy makers, and the political and institutional environment in which they operate. The scale of our experiment provides enough variation to examine preliminarily whether these factors matter. We find some evidence, albeit highly speculative, that these considerations are relevant. The personalities of both inspectors and senior health officials predict how the 'Monitoring the Monitors' program impacted them. Similarly, the degree of local political competition predicts where the program will work best. These results are correlational, and should be treated with appropriate skepticism, but do suggest that the effect of data on policy decision might depend critically on context.

The 'Monitoring the Monitors' program cost 17,800 USD to set up and 510 USD per month to operate.[31] While the results of the program are mixed, given this low cost, we

[30] Recent research finds that providing information to politicians can similarly impact their decisions (Hjort et al., 2019).

[31] The set up costs included 4,470 USD to develop the app and 13,330 USD for smartphones.

would expect it to pass a cost-effectiveness test. The government of Punjab scaled the program up to cover the entire province at the conclusion of the study. This investment by Punjab and others like it have driven a revolution in the amount of data that can quickly and cheaply be accessed for policy decisions. This trend is only likely to accelerate with the rise of remote sensing, digital trace (e.g., cell phone call and mobile money transaction records), smartphones, and other research innovations. A key lesson from this exercise is that, appropriately channeled, these data streams can improve policy outcomes.

References

Ashraf, Nava, Oriana Bandiera, and Kelsey Jack, "No Margin, No Mission? A Field Experiment on Incentives for Public Services Delivery," *Journal of Public Economics*, 2014, *120*, 1–17.

_ , _ , and Scott Lee, "Do-gooders and Go-getters: Career Incentives, Selection, and Performance in Public Service Delivery.," *Working Paper*, 2015.

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

Banerjee, Abhijit V., 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*, 04-05 2008, *6* (2-3), 487–500.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu, "The Glittering Prizes: Career Incentives and Bureaucratic Performance," *Working Paper*, 2017.

Blum, Florian and Rohini Pande, "Data poverty makes it harder to fix real poverty. That's why the UN should push countries to gather and share data.," July 2015.

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

_ , Clark Gibson, Danielle Jung, and James Long, "Improving Electoral Integrity with Information and Communications Technology," *Journal of Experimental Political Science*, 2016, *3* (1), 4–17.

_ , Saad Gulzar, Ali Hasanain, Yasir Khan, and Arman Rezaee, "Personalities and Public Sector Performance: Evidence from a Health Experiment in Pakistan," *NBER Working Paper #21180*, 2017.

_ , _ , _ , _ , **and** _ , "The Political Economy of Public Sector Absence," *NBER Working Paper #22340*, 2017.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust data-driven inference in the regression-discontinuity design," *The Stata Journal*, 2014, *14* (4), 909–946.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers, "Missing in Action: Teacher and Health Worker Absence in Developing Countries," *Journal of Economic Perspectives*, Winter 2006, *20* (1).

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi, "Strengthening State Capabilities: The Role of Financial Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *Quarterly Journal of Economics*, 2013, *128* (3), 1169 – 1218.

Deserranno, Erika, "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda," 2017.

Dhaliwal, Iqbal and Rema Hanna, "Deal with the Devil: The Successes and Limitations of Bureaucratic Reform in India," *Journal of Development Economics*, 2017.

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

Finan, Frederico, Benjamin A Olken, and Rohini Pande, *The Personnel Economics of the State*, North Holland, 2017.

Fisher, Ronald, *The Design of Experiments*, Edinburgh: Oliver and Boyd, 1935.

Gerber, Alan S. and Donald P. Green, *Field Experiments: Design, Analysis, and Interpretation*, New York, NY: Norton New York, 2012.

Gulzar, Saad and Benjamin J. Pasquale, "Politicians, Bureaucrats, and Development: Evidence from India," *American Political Science Review*, 2017, *111* (1), 162–183.

Habyarimana, James, Stuti Khemani, and Thiago Scot, "The Importance of Political Selection in the Performance of Government Bureaucracies," Technical Report 2020.

Hjort, Jonas, Diana Moreira, Gautam Rao, and Juan Francisco Santini, "How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities," Technical Report 2019.

Khan, Adnan, Asim Ijaz Khwaja, and Benjamin A. Olken, "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *Quarterly Journal of Economics*, 2016, *131* (1), 219–271.

Khan, Muhammad Yasir, "Organizational Mission, Financial Rewoards, and the Performace of Public Sector Personnel," Technical Report, UC Berkeley 2020.

Muralidharan, Karthik and Paul Niehaus, "Experimentation at Scale," 11 2017, *31*, 103–124.

_ , _ , and Sandip Sukhtankar, "Building state capacity: Evidence from biometric smartcards in India," *The American Economic Review*, 2016, *106* (10), 2895–2929.

_ , _ , and _ , "Identity verification standards in welfare programs: experimental evidence from India," Technical Report, UC San Diego 2020.

_ , _ , _ , and Jeff Weaver, "Improving last-mile service delivery using phone-based monitoring," Technical Report, UC San Diego June 2019.

Nealer, Erin, Charles Rice, and Erol Yayboke, "Harnessing the Data Revolution to Achieve the Sustainable Development Goals: Enabling Frogs to Leap," *Center for Stategic and International Studies*, 2017.

Rogger, Daniel and Imran Rasul, "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service," *Economic Journal*, forthcoming.

Xu, Guo, "The Costs of Patronage: Evidence from the British Empire," *Working Paper*, 2017.

APPENDIX: FOR ONLINE PUBLICATION ONLY

A Additional Tables and Figures

A.1 Tables

Table A1: Randomization Verification

	Conventional Monitoring (=1)	Smartphone Monitoring (=1)	Difference	P-value	Control Observations	Treatment Observations
Health clinic open during visit (=1)	0.926	0.930	-0.004	0.911	420	428
	[0.262]	[0.256]	(0.033)			
Number of Staff Assigned to the Health Clinic	5.121	5.285	-0.164	0.185	420	428
	[0.925]	[0.940]	(0.121)			
Number of Staff Present	2.719	2.883	-0.164	0.370	420	428
	[1.514]	[1.637]	(0.181)			
Doctor Present (=1)	0.237	0.408	-0.171	0.003	409	414
	[0.426]	[0.492]	(0.054)			
Health/Medical Technician Present (=1)	0.396	0.349	0.047	0.392	409	413
	[0.490]	[0.477]	(0.055)			
Dispenser Present (=1)	0.706	0.777	-0.071	0.259	408	413
	[0.456]	[0.417]	(0.062)			
School Health and Nutrition Supervisor Present (=1)	0.347	0.341	0.006	0.922	406	413
	[0.477]	[0.475]	(0.060)			
Lady Health Worker Present (=1)	0.581	0.634	-0.054	0.298	408	413
	[0.494]	[0.482]	(0.051)			
Midwife Present (=1)	0.536	0.478	0.058	0.206	405	412
	[0.499]	[0.500]	(0.045)			
Health Inspector Has Visited in the Last Month (=1)	0.229	0.219	0.010	0.855	332	320
	[0.421]	[0.414]	(0.055)			
Enumerator Could Not Verify Last Health Inspection (=1)	0.210	0.252	-0.043	0.467	420	428
	[0.407]	[0.435]	(0.058)			
Senior Health Official Has Visited in the Last Month (=1)	0.041	0.038	0.003	0.883	246	290
	[0.198]	[0.191]	(0.018)			
Number of Antenatal Visits in the Last Month	54.255	59.115	-4.859	0.601	274	288
	[54.344]	[49.144]	(9.217)			
Number of Polio Vaccinations in the Last Month	480.226	649.748	-169.522	0.542	305	326
	[1130.822]	[1665.478]	(274.994)			
Doctor Connected to Local Parliamentarian (=1)	0.330	0.273	0.057	0.542	91	176
	[0.473]	[0.447]	(0.092)			
Doctor's Tenure (in months)	111.732	106.025	5.707	0.779	82	161
	[98.782]	[90.214]	(20.156)			
Population of Health Clinic's Catchment Area	22.159	24.243	-2.084	0.187	416	426
	[6.934]	[8.161]	(1.546)			
Distance of the Health Clinic to the District's HQ (kms)	49.521	49.464	0.058	0.992	420	401
	[29.334]	[30.817]	(5.563)			

Notes: This table checks balance between treatment and control clinics. The unit of observation is the clinic (basic health unit). The first ten rows report data from the baseline survey of health facilities which involved making unannounced visits to facilities in November, 2011. The last four rows report data based on the February 2008 parliamentary election. The political competition index is a Herfindahl index computed as the sum of squared candidate vote shares in each provincial assembly constituency. Variable standard deviations are reported in brackets. Standard errors are reported in parentheses.

Table A2: Health Inspection Measurement Attrition

	Conventional Monitoring (=1)	Smartphone Monitoring (=1)	Difference	P-value	Control Observations	Treatment Observations
Enumerator could not verify last health inspection (Wave 1)	0.210	0.252	-0.043	0.467	420	428
	[0.407]	[0.435]	(0.058)			
Enumerator could not verify last health inspection (Wave 2)	0.116	0.145	-0.029	0.518	422	428
	[0.321]	[0.352]	(0.044)			
Enumerator could not verify last health inspection (Wave 3)	0.078	0.082	-0.004	0.908	424	428
	[0.268]	[0.274]	(0.034)			

Notes: This table checks balance between in our enumerators' ability to verify the last health inspection between treatment and control clinics, survey wave by wave. The unit of observation is the clinic (basic health unit). Variable standard deviations are reported in brackets. Standard errors are reported in parentheses.

Table A3: The Effect of Smartphone Monitoring on Inspector Time Use

	Inspecting			Management in Head Office			Other Work				Total	
	Clinics	Hospitals	Total (1)+(2)	Clinics	Hospitals	Total (4)+(5)	Immunization Campaigns	Official Meetings	Other Work	Total (8)+(9)	Work (3)+(6)+(10)	Breaks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Smartphone Monitoring	23.668**	-13.616	10.052	-8.025	6.495	-1.530	3.367	36.679*	-4.297	35.749	44.271	-2.821
	(9.562)	(11.618)	(11.419)	(8.810)	(9.823)	(12.257)	(24.837)	(20.898)	(33.460)	(33.915)	(30.900)	(5.554)
Exact p-value	[0.034]	[0.891]	[0.251]	[0.482]	[0.324]	[0.261]	[0.489]	[0.194]	[0.685]	[0.302]	[0.197]	[0.379]
Control Mean	49.444	47.963	97.407	37.500	40.417	77.917	96.713	66.944	62.500	226.157	401.481	24.444
# Districts	35	35	35	35	35	35	35	35	35	35	35	35
# Observations	117	117	117	117	117	117	117	117	117	117	117	117

Notes: This table reports average treatment effects of the 'Monitoring the Monitors' program on the time use of inspectors. The unit of observation is the inspector, and data comes from surveys of inspectors. The dependent variable in each column is the reported number of minutes spent on the category listed in the column header. All regressions include randomization block FEs. Standard errors clustered at the district level are in parentheses. Square brackets report the Fisher Exact p-values. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A4: The Effect of Smartphone Monitoring on Inspections

	Wave2+3 (1)	Wave 2 (2)	Wave 3 (3)	Wave2+3 (4)	Wave 2 (5)	Wave 3 (6)	Diff-in-Diff (7)	Diff-in-Diff (8)
Smartphone Monitoring	0.181***	0.264***	0.103	0.161***	0.261***	0.058	0.181***	0.210***
	(0.066)	(0.079)	(0.079)	(0.056)	(0.063)	(0.073)	(0.066)	(0.065)
Exact p-value	[0.002]	[0.004]	[0.066]	[0.009]	[0.014]	[0.191]	[0.002]	[0.000]
Control Mean	0.245	0.255	0.235	0.245	0.255	0.235	0.240	0.239
# Districts	35	35	35	35	35	35	35	35
# Observations	1523	739	784	1523	739	784	2175	2117
Fixed Effects	No	No	No	Block	Block	Block	Wave	Wave+Facility

Notes: This table reports average treatment effects of the 'Monitoring the Monitors' program on the probability of a clinic being inspected. The unit of observation is the clinic, and data come from primary unannounced surveys after the treatment was launched. The dependent variable is an indicator for whether an inspector visited the facility visit. Standard errors clustered at the district level are in parentheses. Square brackets report the Fisher Exact p-values. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A5: The Effect of Smartphone Monitoring on Inspections (controlling for baseline level)

	Wave2+3 (1)	Wave 2 (2)	Wave 3 (3)	Wave2+3 (4)	Wave 2 (5)	Wave 3 (6)	Wave2+3 (7)
Smartphone Monitoring	0.207***	0.290***	0.129	0.184***	0.304***	0.060	0.183***
	(0.062)	(0.077)	(0.077)	(0.054)	(0.064)	(0.072)	(0.055)
Exact p-value	[0.002]	[0.004]	[0.027]	[0.008]	[0.012]	[0.155]	[0.008]
Control Mean	0.249	0.261	0.238	0.249	0.261	0.238	0.249
# Districts	35	35	35	35	35	35	35
# Observations	1192	580	612	1192	580	612	1192
Fixed Effects	No	No	No	Block	Block	Block	Wave+Block
Control for baseline level	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports average treatment effects of the 'Monitoring the Monitors' program on the probability of a clinic being inspected while controlling for baseline inspection levels. The unit of observation is the clinic, and data come from primary unannounced surveys after the treatment was launched. The dependent variable is an indicator for whether an inspector visited the facility visit. Standard errors clustered at the district level are in parentheses. Square brackets report the Fisher Exact p-values. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A6: The Effect of Smartphone Monitoring on Clinic Staff Attendance - Wave 2

	Doctor	Dispenser	Lady Health Worker	Health Tech	School Health Nutri. Supervisor	Midwife
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Difference in Differences						
Monitoring	-0.044	-0.082	-0.102*	-0.037	0.034	0.014
	(0.058)	(0.060)	(0.056)	(0.063)	(0.086)	(0.067)
	[0.807]	[0.961]	[0.996]	[0.863]	[0.385]	[0.634]
Mean in Controls	0.247	0.735	0.614	0.439	0.333	0.554
# Districts	35	35	35	35	35	35
# Observations	1596	1632	1632	1633	1628	1624
R-Squared	0.090	0.062	0.038	0.044	0.034	0.048
Panel B: Difference in Differences with Facility FEs						
Monitoring	-0.035	-0.087	-0.103*	-0.037	0.031	0.009
	(0.060)	(0.057)	(0.057)	(0.064)	(0.083)	(0.068)
	[0.704]	[0.991]	[0.957]	[0.843]	[0.582]	[0.663]
Mean in Controls	0.242	0.737	0.615	0.447	0.334	0.559
# Districts	35	35	35	35	35	35
# Observations	1498	1566	1566	1568	1558	1550
R-Squared	0.639	0.556	0.552	0.646	0.538	0.641
Panel C: Control for Baseline Attendance						
Monitoring	0.114***	-0.023	-0.035	-0.063	0.015	-0.073
	(0.037)	(0.039)	(0.032)	(0.044)	(0.073)	(0.068)
	[0.059]	[0.860]	[0.761]	[0.943]	[0.423]	[0.864]
Mean in Controls	0.250	0.764	0.648	0.486	0.319	0.572
# Districts	35	35	35	35	35	35
# Observations	749	783	783	784	779	775
R-Squared	0.179	0.062	0.037	0.134	0.061	0.138

Notes: This table reports average treatment effects of the 'Monitoring the Monitors' program on the staff attendance in survey wave 2. The unit of observation is the clinic, and data come from primary unannounced surveys after the treatment was launched. The dependent variable is an indicator for whether a doctor was present at the clinic during an announced visit. All regressions include randomization strata FEs. Standard errors clustered at the district level are in parentheses. Square brackets report the Fisher Exact p-values. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A7: The Effect of Smartphone Monitoring on Clinic Staff Attendance - Wave 3

	Doctor (1)	Dispenser (2)	Lady Health Worker (3)	Health Tech (4)	School Health Nutri. Supervisor (5)	Midwife (6)
Panel A: Difference in Differences						
Monitoring	-0.001	-0.055	-0.058	0.098	-0.012	0.095
	(0.055)	(0.070)	(0.062)	(0.063)	(0.082)	(0.068)
	[0.493]	[0.880]	[0.961]	[0.097]	[0.693]	[0.308]
Mean in Controls	0.212	0.722	0.568	0.379	0.318	0.517
# Districts	35	35	35	35	35	35
# Observations	1649	1646	1646	1647	1639	1641
R-Squared	0.088	0.076	0.048	0.050	0.054	0.048
Panel B: Difference in Differences with Facility FEs						
Monitoring	0.008	-0.045	-0.054	0.090	0.001	0.088
	(0.055)	(0.069)	(0.064)	(0.065)	(0.086)	(0.067)
	[0.546]	[0.859]	[0.894]	[0.210]	[0.737]	[0.280]
Mean in Controls	0.211	0.720	0.570	0.381	0.320	0.515
# Districts	35	35	35	35	35	35
# Observations	1598	1592	1592	1594	1578	1582
R-Squared	0.625	0.560	0.557	0.618	0.533	0.593
Panel C: Control for Baseline Attendance						
Monitoring	0.094*	0.002	-0.040	0.052	0.000	-0.007
	(0.053)	(0.055)	(0.043)	(0.050)	(0.050)	(0.073)
	[0.134]	[0.667]	[0.854]	[0.460]	[0.524]	[0.719]
Mean in Controls	0.182	0.729	0.554	0.369	0.284	0.495
# Districts	35	35	35	35	35	35
# Observations	799	796	796	797	789	791
R-Squared	0.146	0.102	0.054	0.118	0.126	0.108

Notes: This table reports average treatment effects of the 'Monitoring the Monitors' program on the staff attendance in survey wave 3. The unit of observation is the clinic, and data come from primary unannounced surveys after the treatment was launched. The dependent variable is an indicator for whether a doctor was present at the clinic during an announced visit. All regressions include randomization strata FEs. Standard errors clustered at the district level are in parentheses. Square brackets report the Fisher Exact p-values. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A8: The Effect of Smartphone Monitoring on Clinic Staff Assignment

	Doctor	Dispenser	Lady Health Worker	Health Tech	School Health Nutri. Supervisor	Midwife
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Survey Wave FE						
Monitoring	0.007	-0.022	-0.036	-0.055	-0.012	-0.060
	(0.048)	(0.016)	(0.024)	(0.052)	(0.117)	(0.068)
	[0.400]	[0.877]	[0.976]	[0.875]	[0.326]	[0.761]
Mean in Controls	0.532	0.979	0.938	0.749	0.876	0.855
# Districts	35	35	35	35	35	35
# Observations	2422	2457	2457	2458	2479	2483
R-Squared	0.141	0.029	0.038	0.090	0.208	0.142
Panel B: Survey Wave + Facility FEs						
Monitoring	0.023	-0.024	-0.033	-0.051	-0.010	-0.061
	(0.048)	(0.016)	(0.024)	(0.053)	(0.118)	(0.068)
	[0.441]	[0.878]	[0.952]	[0.965]	[0.380]	[0.840]
Mean in Controls	0.533	0.979	0.937	0.748	0.876	0.854
# Districts	35	35	35	35	35	35
# Observations	2408	2446	2446	2447	2473	2477
R-Squared	0.677	0.521	0.551	0.648	0.558	0.504
Panel C: Baseline Control with Survey Wave FE						
Monitoring	0.091***	-0.016	-0.015	-0.067	0.019	-0.057
	(0.031)	(0.013)	(0.011)	(0.054)	(0.071)	(0.058)
	[0.084]	[0.788]	[0.761]	[0.908]	[0.182]	[0.722]
Mean in Controls	0.517	0.987	0.942	0.738	0.811	0.782
# Districts	35	35	35	35	35	35
# Observations	1548	1579	1579	1581	1623	1627
R-Squared	0.310	0.078	0.094	0.193	0.217	0.063
Panel D: IV with Baseline Controls and Survey Wave FE						
	(1)	(2)	(3)	(4)	(5)	(6)
Inspected	0.626**	-0.088	-0.084	-0.424	0.148	-0.326
	(0.294)	(0.079)	(0.077)	(0.284)	(0.378)	(0.354)
	[0.463]	[0.336]	[0.520]	[0.520]	[0.551]	[0.600]
Mean in Controls	0.597	0.979	0.938	0.701	0.792	0.746
# Districts	35	35	35	35	35	35
# Observations	1437	1468	1468	1470	1509	1513

Notes: This table reports average treatment effects of the 'Monitoring the Monitors' program on the staff assignment. The unit of observation is the clinic, and data come from primary unannounced surveys after the treatment was launched. The dependent variable is an indicator for whether a staff member was assigned at the clinic during an unannounced visit. All regressions include randomization strata FEs. Standard errors clustered at the district level are in parentheses. Square brackets report the Fisher Exact p-values. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A9: Flagging Alternative Bandwidths

	Doctor Present in Unannounced Visit (=1)			
	(1)	(2)	(3)	(4)
Flagged	0.177**	0.098	0.075	0.079
	(0.082)	(0.062)	(0.054)	(0.054)
Unflagged Mean	0.236	0.238	0.281	0.281
# Clinics	112	199	266	268
# Reports	130	241	373	376
R-Squared	0.298	0.218	0.160	0.156
# of Staff Absent	2 or 3	1 to 4	0 to 5	0 to 7

Notes: This table reports on the effect on subsequent doctor attendance of flagging on an online dashboard. Clinics were flagged in red on an online dashboard if three or more of the seven staff were absent in one or more health inspections of the clinic 11 to 25 days prior to an unannounced visit by our survey enumerators. Column 1 restricts to the Discontinuity sample, or facilities with reports in which either two or three staff were absent (the threshold to trigger the underreporting red flag). Column 2 restricts to reports with one to four staff absent. Column 3 restricts to reports with zero to five staff absent. Column 4 does not restrict by the number of staff absent. In addition, the sample in all columns is limited to Monitoring the Monitor treatment districts due to the necessity of the web dashboard for flagging clinics. All regressions include district and survey wave fixed effects. Standard errors clustered at the clinic level reported in parentheses. $p < 0.1$, $^{*}p < 0.05$, $^{***}p < 0.01$.

Table A10: First Flagging Effects

	Doctor Present in Unannounced Visit (=1)		
	(1)	(2)	(3)
Flagged	0.177**	0.224*	0.273*
	(0.082)	(0.118)	(0.148)
Unflagged Mean	0.236	0.208	0.208
# Clinics	112	66	66
# Reports	130	72	72
R-Squared	0.298	0.229	0.372
District Fixed Effects	Yes	Yes	Yes
Restricted to First Flagging	No	Yes	Yes
Number of previous inspections FEs	No	No	Yes
Sample	Discontinuity	Discontinuity	Discontinuity

Notes: This table reports on the effect on subsequent doctor attendance of flagging on an online dashboard the fact that a clinic had three or more staff absent to a senior health official. Clinics were flagged in red on an online dashboard if three or more of the seven staff were absent in one or more health inspections of the clinic 11 to 25 days prior to an unannounced visit by our survey enumerators. All columns restrict to the Discontinuity sample, or facilities with reports in which either two or three staff were absent (the threshold to trigger the underreporting red flag). Column 1 reports results identical to Table 3 column 1. Column 2 restricts the sample to cases in which the flagging that occurred prior to our survey visit was the first time a facility was ever flagged. Column 3 further includes fixed effects for the number of previous health inspections recorded in the online dashboard for a facility before the beginning of the flagging window. In addition, the sample in all columns is limited to Monitoring the Monitor treatment districts due to the necessity of the web dashboard for flagging clinics. All regressions include district and survey wave fixed effects. Standard errors clustered at the clinic level reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A11: Flagging Balance

	Not flagged	Flagged	Difference	P-value	Not flagged Observations	Flagged Observations
Health clinic open during visit (=1)	0.980	0.903	0.077	0.084	50	62
	[0.141]	[0.298]	(0.047)			
Number of Staff Assigned to the Health Clinic	5.340	5.000	0.340	0.045	50	62
	[0.798]	[1.024]	(0.191)			
Number of Staff Present	3.400	2.694	0.706	0.047	50	62
	[1.245]	[1.500]	(0.340)			
Doctor Present (=1)	0.408	0.379	0.029	0.674	49	58
	[0.497]	[0.489]	(0.091)			
Health/Medical Technician Present (=1)	0.347	0.345	0.002	0.984	49	58
	[0.481]	[0.479]	(0.132)			
Dispenser Present (=1)	0.796	0.828	-0.032	0.725	49	58
	[0.407]	[0.381]	(0.107)			
School Health and Nutrition Supervisor Present (=1)	0.531	0.362	0.169	0.196	49	58
	[0.504]	[0.485]	(0.131)			
Lady Health Worker Present (=1)	0.755	0.586	0.169	0.102	49	58
	[0.434]	[0.497]	(0.106)			
Midwife Present (=1)	0.633	0.379	0.253	0.016	49	58
	[0.487]	[0.489]	(0.084)			
Health Inspector Has Visited in the Last Month (=1)	0.250	0.333	-0.083	0.450	40	54
	[0.439]	[0.476]	(0.131)			
Enumerator Could Not Verify Last Health Inspection (=1)	0.200	0.129	0.071	0.417	50	62
	[0.404]	[0.338]	(0.082)			
Senior Health Official Has Visited in the Last Month (=1)	0.026	0.022	0.005	0.904	38	46
	[0.162]	[0.147]	(0.035)			
Number of Antenatal Visits in the Last Month	59.308	55.275	4.033	0.699	39	40
	[45.801]	[44.848]	(11.894)			
Number of Polio Vaccinations in the Last Month	649.632	555.170	94.461	0.840	38	47
	[1689.583]	[1493.729]	(486.732)			
Doctor Connected to Local Parliamentarian (=1)	0.211	0.227	-0.017	0.901	19	22
	[0.419]	[0.429]	(0.140)			
Doctor's Tenure (in months)	106.786	99.000	7.786	0.783	14	20
	[77.134]	[89.016]	(34.294)			
Population of Health Clinic's Catchment Area	24.815	24.385	0.429	0.763	50	62
	[8.249]	[8.702]	(1.824)			
Distance of the Health Clinic to the District's HQ (kms)	51.571	54.800	-3.229	0.476	49	60
	[33.125]	[27.417]	(7.628)			

Notes: This table checks balance between flagged and not flagged clinics considering those that were inspected for the online dashboard 11-25 days prior to one of our survey visit (the same sample as the primary flagging results in Table 3). The unit of observation is the clinic (basic health unit). All data is from the baseline survey of health facilities which involved making unannounced visits to facilities in November, 2011. Variable standard deviations are reported in brackets. Standard errors are reported in parentheses.

Table A12: Flagging Effects Controlling for Baseline Imbalances

	Doctor Present in Unannounced Visit (=1)	
	(1)	(2)
Flagged	0.177**	0.175**
	(0.082)	(0.082)
Unflagged Mean	0.236	0.246
# Clinics	112	107
# Reports	130	124
R-Squared	0.298	0.358
District Fixed Effects	Yes	Yes
Controls for Baseline Imbalances	No	Yes
Sample	Discontinuity	Discontinuity

Notes: This table reports on the effect on subsequent doctor attendance of flagging on an online dashboard the fact that a clinic had three or more staff absent to a senior policymaker. Clinics were flagged in red on an online dashboard if three or more of the seven staff were absent in one or more health inspections of the clinic 11 to 25 days prior to an unannounced visit by our survey enumerators. The Discontinuity sample limits to facility reports in which either two or three staff were absent (the threshold to trigger the underreporting red flag). In addition, the sample in all columns is limited to Monitoring the Monitor treatment districts due to the necessity of the web dashboard for flagging clinics. Controls for Baseline Imbalances include: Health clinic open during visit (=1), Number of Staff Assigned to the Health Clinic, Number of Staff Present, and Midwife Present (=1). All regressions include district and survey wave fixed effects. Standard errors clustered at the clinic level reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A13: Heterogeneous Flagging Effects by Senior Health Official Personality

	Doctor Present in Unannounced Visit (=1)		
	(1)	(2)	(3)
Flagged	0.177**	0.156*	0.210**
	(0.082)	(0.084)	(0.090)
Flagged x Senior Health Official Big Five Index		0.282*	
		(0.150)	
Flagged x Senior Health Official PSM Index			0.165
			(0.126)
Mean of dependent variable	0.308	0.267	0.267
# Observations	130	120	120
# Clinics	112	103	103
R-Squared	0.298	0.265	0.255
District Fixed Effects	Yes	Yes	Yes
Sample	Discontinuity	Discontinuity	Discontinuity

Notes: This table reports on the effect on subsequent doctor attendance of flagging on an online dashboard the fact that a clinic had three or more staff absent to a senior policymaker. Clinics were flagged in red on an online dashboard if three or more of the seven staff were absent in one or more health inspections of the clinic 11 to 25 days prior to an unannounced visit by our survey enumerators. The Discontinuity sample limits to facility reports in which either two or three staff were absent (the threshold to trigger the underreporting red flag). In addition, the sample in all columns is limited to Monitoring the Monitor treatment districts due to the necessity of the web dashboard for flagging clinics.The Big Five and PSM indices. The Big Five and PSM indices are z-score averages of the five and six traits within the Big Five and PSM respectively, elicited through responses to statements that represent the trait on a five point Likert scale, in which 1 corresponds to disagree strongly, 2 to disagree a little, 3 to neutral, 4 to agree a little, and 5 to agree strongly. Likert responses are given the same direction. Standard errors clustered at the clinic level reported in parentheses. All regressions include district and survey wave fixed effects and condition on a doctor being posted. All regressions include survey wave fixed effects. Standard errors clustered at the clinic level reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A14: Heterogeneous Flagging Effects by Political Competition

	Doctor Present in Unannounced Visit (=1)		
	(1)	(2)	(3)
Flagged	0.177**		
	(0.082)		
Flagged x Doctor Does Not Know Politician			0.201*
			(0.109)
Flagged x Doctor Knows Politician			-0.250
			(0.249)
Flagged x High Competition		0.359***	
		(0.118)	
Flagged x Med Competition		0.004	
		(0.165)	
Flagged x Low Competition		-0.087	
		(0.134)	
Constant	0.191***	0.209*	0.588**
	(0.061)	(0.117)	(0.233)
Flagged x High Comp = Flagged x Med Comp (p-value)		0.090	
Flagged x High Comp = Flagged x Low Comp (p-value)		0.014	
Flagged x Doctor Does Not Know = Flagged x Doctor Knows (p-value)			0.072
# Clinics	112	112	80
# Reports	130	130	91
R-Squared	0.298	0.352	0.347
District Fixed Effects	Yes	Yes	Yes
Sample	Discontinuity	Discontinuity	Discontinuity

Notes: This table reports on the effect on subsequent doctor attendance of flagging on an online dashboard the fact that a clinic had three or more staff absent to a senior policymaker. Clinics were flagged in red on an online dashboard if three or more of the seven staff were absent in one or more health inspections of the clinic 11 to 25 days prior to an unannounced visit by our survey enumerators. The Discontinuity sample limits to facility reports in which either two or three staff were absent (the threshold to trigger the underreporting red flag). In addition, the sample in all columns is limited to Monitoring the Monitor treatment districts due to the necessity of the web dashboard for flagging clinics. All regressions include survey wave fixed effects. Standard errors clustered at the clinic level reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.2 Figures

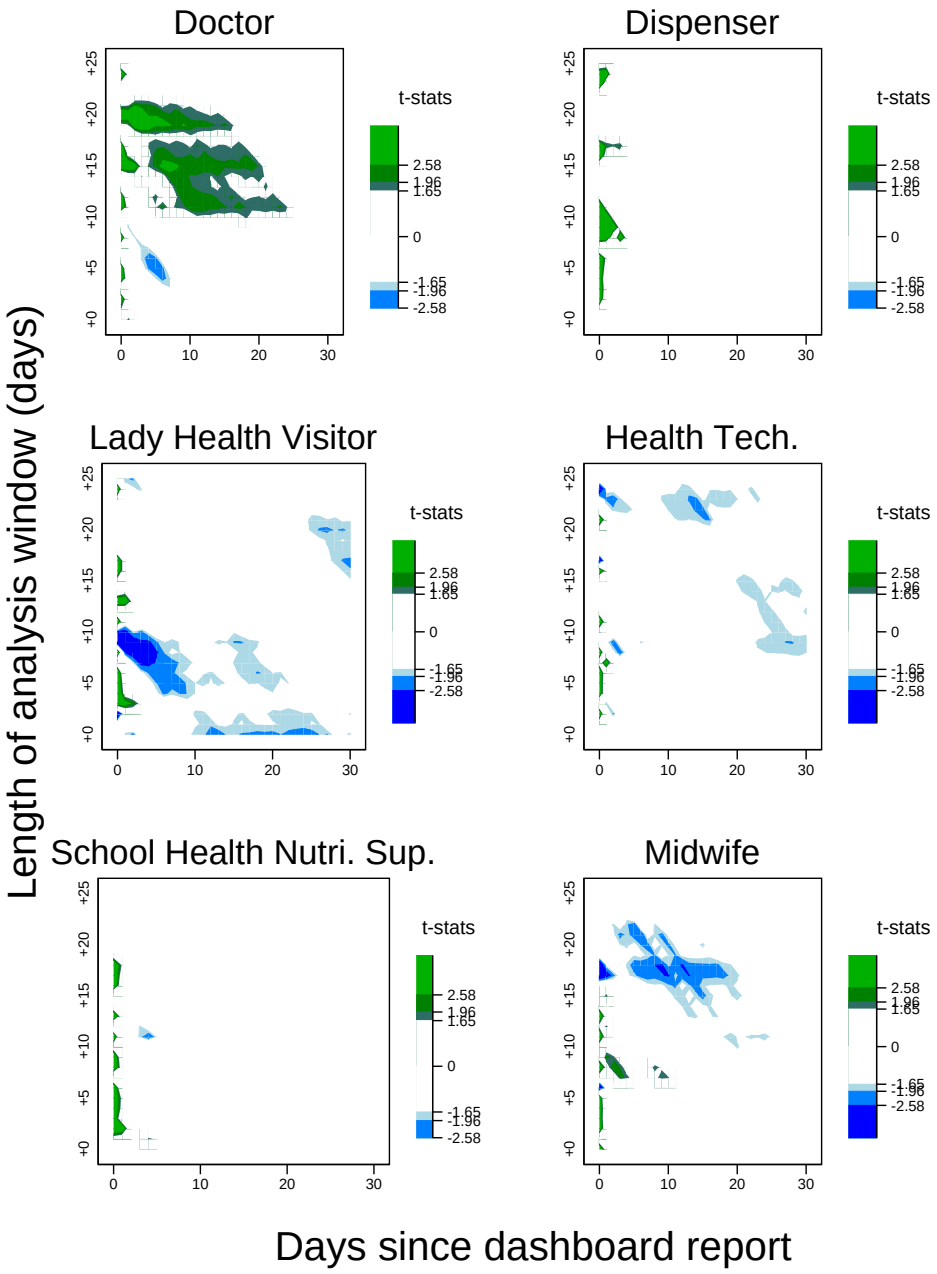


Figure A1: Absence after Flagging

Notes: This figure presents robustness of flagging results in Table 3 to the window of time prior to an unannounced visit that a clinic being highlighted in red on the dashboard is considered flagged. Each figure reports t-stats from 1300 hypothesis tests analogous to that conducted in Table 3 column (3) that Flagged = 0, varying the window for which we define a clinic as flagged prior to a primary unannounced visit to a clinic along two dimensions—we vary the length of the window being used along the x-axis and the delay from when a clinic is highlighted in red to when the window begins along the y-axis (so for example, a length of 30 and delay of 15 corresponds to considering a clinic as flagged if it was highlighted in red anytime 15 to 45 days prior to an unannounced visit).

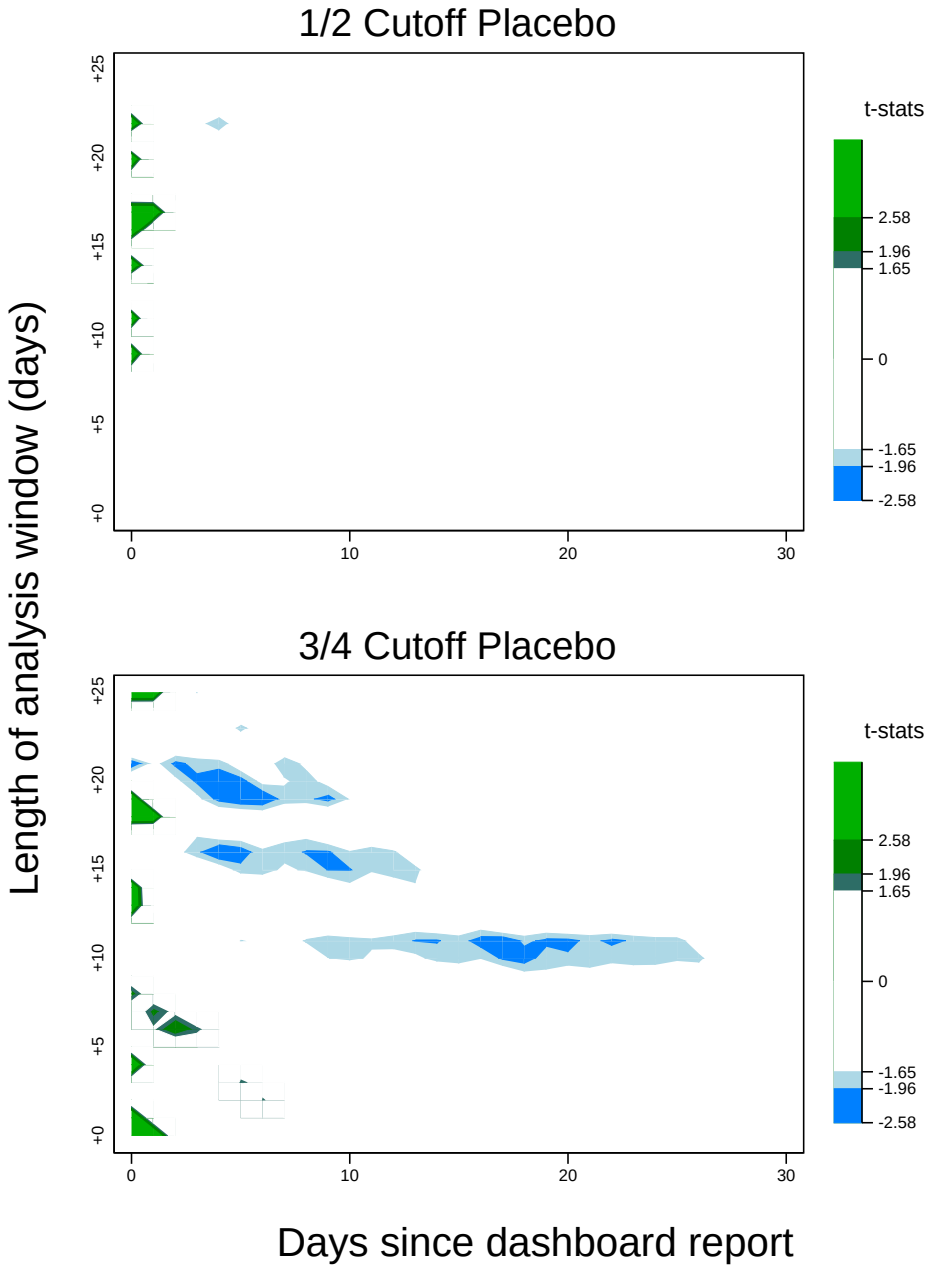


Figure A2: Absence after Flagging Placebos

Notes: This figure presents placebo tests for the flagging results in Table 3, coding the data as if the cutoff for a health clinic being highlighted in red was for more than 1 staff absent or more than 3 staff absent, rather than the true cutoff that was used (more than 2 staff absent). Each figure reports t-stats from 1300 hypothesis tests analogous to that conducted in Table 3 column (3) that Flagged = 0, with these placebo definitions of flagging, varying the window for which we define a clinic as flagged prior to a primary unannounced visit to a clinic along two dimensions—we vary the length of the window being used along the x-axis and the delay from when a clinic is highlighted in red to when the window begins along the y-axis (so for example, a length of 30 and delay of 15 corresponds to considering a clinic as flagged if it was highlighted in red anytime 15 to 45 days prior to an unannounced visit).

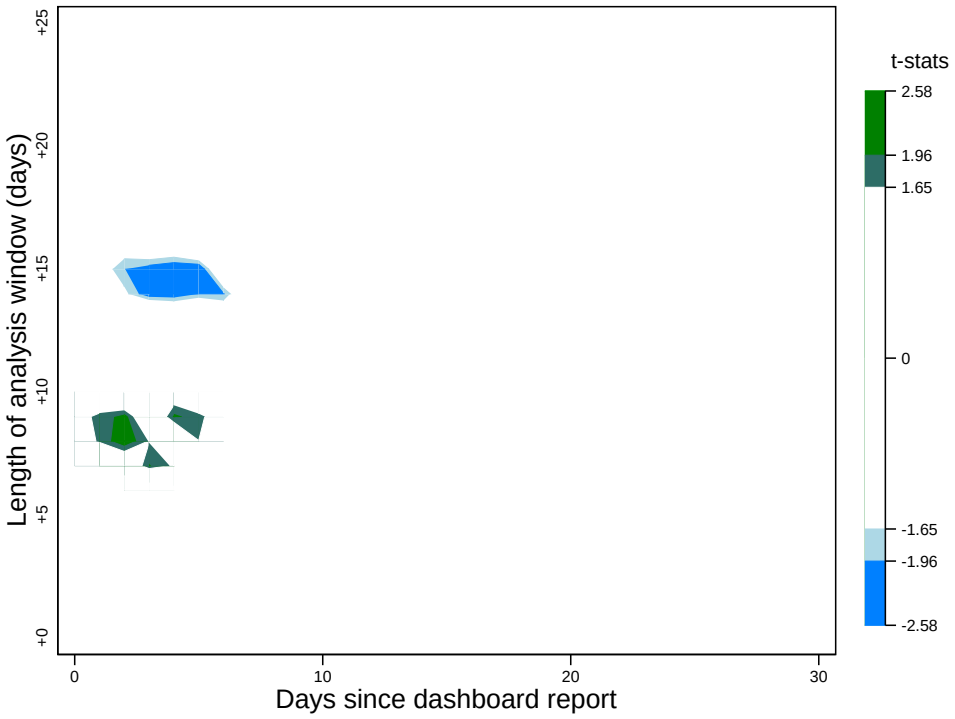


Figure A3: Survey Before Flagging Placebo

Notes: This figure presents placebo tests for the flagging results in Table 3, reversing the flagging window to be after our independent survey. Each figure reports t-stats from 750 hypothesis tests analogous to that conducted in Table 3 column (3) that Flagged = 0, with these placebo definitions of flagging, varying the window for which we define a clinic as flagged *after* a primary unannounced visit to a clinic along two dimensions—we vary the length of the window being used along the x-axis and the delay from when a clinic is highlighted in red to when the window begins along the y-axis (so for example, a length of 30 and delay of 15 corresponds to considering a clinic as flagged if it was highlighted in red anytime 15 to 45 days after an unannounced visit).

B Duties of Health Inspectors

Health inspectors' primary duty is to directly monitor health service delivery across their sub-district. This monitoring takes three official forms: (i) inspection of health facilities, (ii) management of facilities and staff, and (iii) supervising immunization campaigns (such as Polio vaccination campaigns).

Health inspectors are nominally required to inspect all health facilities in their sub-district once per month. These facilities include Basic Health Units (2,496 clinics in this study spread across 123 health inspectors), Rural Health Centers (293 RHCs), Tehsil Headquarter Hospitals (88 THQs), and District Headquarter Hospitals (34 DHQs). The average health inspector is thus responsible for roughly 24 inspections per month. We have provided the inspection form for BHUs in Appendix Section E. The inspection forms for these other facilities are similar in nature. In control districts, health inspectors report spending an average of 97 minutes per day on inspections, roughly split between BHUs and other facilities (see Appendix Table A3).

Health inspectors report spending an average of 78 minutes per day on management tasks, which typically take place at the office (health inspectors' offices are in their sub-district's THQ). These management tasks are mainly planning in nature—monitoring medicine and supply usage and potential stock-outs, conducting other needs assessments, and planning for special interventions (such as special programs for treating Tuberculosis). Management tasks also include preparing paperwork related to health inspectors' monitoring activities.

Health inspectors report spending an average of 97 minutes per day supervising immunization campaigns. This requires the supervision and monitoring of frontline health workers such as Lady Health Workers as they rove around the sub-district conducting door-to-door vaccinations. These campaigns are not constant – they tend to run for one to two weeks every few months – but when they are running they require the near full attention of health inspectors.

Besides these three official streams of duties, health inspectors are expected to perform

any task assigned to them by more senior health officials as they are the individual most responsible for health at the sub-district level. The most common task in this category is attending "official meetings" (an average of 67 minutes per week), which generally entails health inspectors being called into their DHQ to provide comments at a meeting between a senior health official and someone visiting the district.

C Hiring and Assignment Process for Medical Officers (Doctors)

There are two different hiring processes for Medical Officers. They are either hired centrally through through the Punjab Provincial Service Commission (PPSC) or locally at the district level. If hired through the PPSC, a Medical Officer becomes part of the provincial bureaucracy. The PPSC is tasked with the hiring of human resources for several arms of the provincial government. The commission floats an advertisement with details of the hiring process. Individuals who have passed the doctor certifications (M.B.B.S.) and are registered with Pakistan Medical and Dental Council are eligible to apply to these positions. The top candidates are called in for a test and further shortlisted candidates are interviewed by a selection committee. The committee consists of senior officials from the PPSC, the Health Department, the Director General Health Services office, and a senior medical expert. Merit lists generated based on performance in the interview are then communicated to the Health Department by the PPSC.

The provincial Health Department decides on assignment of these Medical Officers to facilities across the province. They can be assigned to any district and the senior health official of that district can assign them to any facility within that district. That means, at least theoretically, these Medical Officers can be moved across district lines any time. However, in practice it is rare for a Medical Officer to be moved across district lines after their first assignment, unless they specifically request it.

The second route of hiring Medical Officers runs through the Executive District Officers (EDO), or senior health officials, at the district level. The Office of the EDO Health advertises vacant positions locally, and shortlisted applicants are interviewed by the EDO personally. Recommendations of the EDO are conveyed to the Health Department, which then issues offer letters to the successful applicants for assignment in that particular district. However, these Medical Officers are hired for a fixed term and are not considered permanent employees. This route exists so that the EDOs can swiftly respond to any surge in demand for health services at the local level. The time required for hiring through this route is much shorter than the hiring process through the PPSC.

Medical Officers that are hired by the EDO of a district can only be assigned to a health facility within that district. This means such officers cannot be moved to a facility outside their own district.

D Details on data collection

Our survey protocol was as follows: First, our survey teams were trained by senior enumerators and our team members at four regional hubs. Following these trainings, the teams made visits to health clinics in their assigned districts and remained in regular contact with their team leaders and our research team. Data collection and entry then followed back-checks and other validation processes consistent with academic best practice. Surveys took three weeks to field in each wave.

When verifying whether health inspections had occurred, enumerators' process was as follows: First, they checked a facility's physical register for the last recorded inspection. Second, they verified what they found in the register with whichever health staff was currently in charge of the facility. In some cases, enumerators were unable to confidently verify whether or not an inspection had occurred in the prior month. We treat such cases as missing data for analysis and verify in Appendix Table A1 that such cases do not correlate with treatment

assignment at baseline and in Appendix Table A2 that we do not have treatment-related attrition across waves.

During time use surveys, inspectors listed the time they spent on a variety of tasks during the two working days prior to our survey. Inspectors picked up to three out of 10 possible categories of work to account for each hour between 8am and 6pm. In addition, they were asked to identify when they arrived at and left from work. We also collected additional information from health inspectors during these interviews, and we also interviewed doctors and senior health officials across our sample through similar face-to-face interviews. We briefly discuss some of the other data collected through these interviews in Section 4.5.

E Paper Inspection Form

INSPECTION FORM OF BASIC HEALTH UNIT
HEALTH DEPARTMENT (GOVERNMENT OF THE PUNJAB)
A - BASIC HEALTH UNIT INFORMATION

HMIS Code: ☐☐☐☐☐☐☐ Name of BHU: _____

Managed by: i) PRSP ☐ ii) Dist. Govt. ☐ (please tick only one option)

Mauza: _____ UC Name: _____ UC No. _____

NA No. _____ PP. No. _____ District: _____ Tehsil: _____

Name of incharge of the facility: _____ Designation: _____

Mobile No.: _____ BHU's Phone (with code): _____

Name of DDOH: _____ Reference No. _____

Date & Time of arrival for inspection: ____/____/____ Time: Hours _____ Minutes _____ am / pm

B - CLEANLINESS AND GENERAL OUTLOOK OF THE FACILITY
(tick relevant column)

Sr. No	Location	Good	Average	Poor
1	Boundary Wall			
2	Lawns			
3	Waiting Area			
4	Building			
5	Labour Room			
6	Wards			
7	Toilets			

C - DISPLAYS
(tick relevant column)

ITEMS		Yes	No
Signboards/Direction Board displayed			
Display in the MO/ Incharge office:	1) Organogram		
	2) Map of Union council showing all localities		
	3) Statistics of the Union Council and the BHU		
	4) Tour Programme of 'outreach team'		

D - AVAILBILITY OF UTILITIES
(tick relevant column)

Sr. No	Name of Utility	Not Available	Available	
			Functional	Non Functional
1	Electricity			
2	Telephone			
3	Water supply System			
4	Sui Gas			
5	Sewerage System			
6	Other			

E - DISPOSAL OF HOSPITAL WASTE
(tick relevant column)

Sr. No	Mode		Yes	No
1	Hospital Waste Segregated as per guidelines			
2	Hospital Waste Lying Open			
3	Burnt by:	(a) Incinerator		
		(b) Other means		
4	Buried			
5	Carried away by municipality			
6	Any other (Please state)			

F - PURCHI FEES (give amount)

Fee Deposited during the current financial year till the last calendar month: ***(Rs.)*** _____

Purchi fee being charged @ ***Rs.*** _____ *per patient.*

G - PATIENTS TREATED IN LAST CALENDAR MONTH (give numbers)

Sr. No	*Cases*	*Numbers*	Sr. No.	*Cases*	*Numbers*
1	OPD Cases		7	Children vaccinated outside BHU	
2	Percentage of pervious day OPD Cases registered with NIC No.		8	TB Patients under Treatment	
3	Deliveries at BHU		9	Hepatitis "B" Vaccination Done	
4	No. of PCD slides prepared		10	Antenatal Cases Checked	
5	No. of referrals to other hospitals		11	Family Planning Visits	
6	Children vaccinated at BHU		12	No. of referrals by LHW	

H- SCHOOL HEALTH PROGRAM

S. No.	Indicator	Numbers
1	Total Number of Students referred by SH&NS during previous month	
2	Total Numbers of student treated at BHU referred by SH&NS	
3.	Total Numbers of School visited during previous month	
4.	Tour program approved and displayed in DHQ	Yes / No

I - DOCTORS (give numbers)

Sanctioned Posts	Filled Posts	Present

This doctor is posted in this BHU but is not present during the visit

Details regarding absence of doctors. (Do not write anything if a doctor is present.)

Sr. No	Designation	Name of Doctors	Type of Absence on the monitoring day. (tick only one box)					Days of absence during last three calendar months	
			UA	SL	OD	St. L	LC	UA	Other Types
1	MO WMO								

Unauthorized absence (UA), Sanctioned leave (SL), On official duty outside the BHY (OD), Short leave (St.L), Late Comer (LC).

J- PARAMEDICS (OTHER THAN DOCTORS)

PARAMEDICS INCLUDES:	Medical Assistant, Medical Technician, Health Technician, Dispenser, Other

Sr. No	Staff Category	Sanctioned	Filled Posts	Present
1	Paramedical Staff			

Details regarding absence of Paramedics. (Do not write anything if staff is present.)

Sr. No	Designation	Name of Paramedic	Type of Absence on the monitoring day. (tick only one box)					Days of absence during last three calendar months	
			UA	SL	OD	St. L	LC	UA	Other Types
1									
2									
3									

K - PREVENTIVE / OUTREACH STAFF

PREVENTIVE / OUTREACH STAFF INCLUDES.	LHV, RHI, Midwife, Dai, Vaccinator CDC Supervisor, Sanitary Inspector Sanitary Worker, School Health & Nutrition Supervisor

Sanctioned Posts	Filled Posts	Present

Details regarding absence of Preventive / outreach staff. (Do not write anything if staff is present.)

Sr. No	Designation	Name of Staff	Type of Absence on the monitoring day. (tick only one box)					Days of absence during last three calendar months	
			UA	SL	OD	St. L	LC	UA	Other Types
1									
2									
3									

L - ADMIN / SUPPORT STAFF

ADMIN / SUPPORT STAFF INCLUDES:	Computer Operator, Naib Qasid, Chowkidar, Mali, Sweeper

Sanctioned Posts	Filled Posts	Present

Details regarding absence of Admin/Support staff. (Do not write anything if staff is present.)

Sr. No	Designation	Name of Staff	Type of Absence on the monitoring day. (tick only one box)					Days of absence during last three calendar months	
			UA	SL	OD	St. L	LC	UA	Other Types
1									
2									
3									

M- VACANT POSTS (please write full name of posts)

Sr. No	Name of Post	Number of Vacant Post	Sr. No.	Name of Post	Number of Vacant Posts

U - MONTHLY PERFORMANCE

Sr. No		Monthly Target	Performance
1	No of children given full immunization coverage		
2	Daily OPD attendance		
3	Delivery coverage at facility		

V– GENERAL REMARKS

Time of Departure from the facility: Hours _____ Minutes _____ am / pm

Certified that this basic Health Unit was inspected today by the undersigned and the information stated above is as per facts and record.

_____ **Signature of DDOH/MEA**

_____ **Signatures & Stamp of MO/Incharge**

_____ **Signatures of DMO/EDOH**

F Training Manual For Smartphone Application Use

MANUAL
FOR
HEALTH FACILITY INFORMATION AGGREGATION SYSTEM

DIRECTORATE GENERAL HEALTH SERVICES
SUPPORTED BY
PUNJAB HEALTH SECTOR REFORMS PROGRAM

CONTENTS

1. INTRODUCTION

The Health Department of Government of the Punjab is committed to adopting state-of-the-art technology to strengthen governance and improve service delivery for all citizens.

For this purpose, the Punjab Health Sector Reforms Program (PHSRP), with technical assistance from International Growth Centre (IGC) Team, is supporting DGHS and district health managers in strengthening the internal monitoring system of the Health Department. This is being done by introducing a mobile phone based information management system that is being rolled out across different districts of the province.

This initiative will improve the internal information transmission within the Health Department and will ensure that timely, authentic and actionable information is sent quickly from individual facilities to district and provincial health managers on such crucially important issues as absenteeism, medicine stock outs, availability and functionality of equipment etc.

Android-based smartphones have been provided to those district supervisory officers, such as Executive District Health Officers (EDOs), District Health Officers (DOs), and Deputy District Health Officers (DDOs), who have been tasked with the collection of performance related data from Basic Health Units (BHUs), Rural Health Centers (RHCs) and Tehsil and District Headquarters (THQs and DHQs).

The report submitted by these officers through the phone will be recorded on a website and automatically analyzed for use by managers at various levels. It is expected that this information will become a powerful tool for management both for district and central level officials. This is expected to bring about marked improvement in health service delivery management, particularly at primary and secondary levels of healthcare, leading to better health outcomes for the poor and disadvantaged in the province.

At Directorate General Health Services, Director, District Health Information System (DHIS), supported by the PHSRP and IGC team, is the focal person for implementation of the program at the provincial level. Overall responsibility for the program at the district level lies with EDOs, and Statistical Officers (SOs) are the designated focal persons for managing the system at the district level.

This manual contains basic information about the program and the phone, as well

as details of how to submit data and deal with some problems that may arise.

2. ABOUT THE PHONE

The HTC Explorer runs on Android 2.3.5 with HTC's latest custom interface - Sense 3.0, and is equipped with a 3.2 inch capacitive touch screen.

The phone has 4 capacitive touch buttons on the front- HOME, MENU, BACK and SEARCH.

With a 600 MHz processor based on the latest mobile technology, 512 MB of RAM and a 2 GB SD card, the phone is well equipped to deal with advanced tasks associated with smart-phones today.

The phone can be used for browsing the internet using either GPRS or WIFI. It is also equipped with a GPS device and a 3 MP camera which can capture high-resolution images and videos.

For detailed instructions regarding how to undertake different tasks on the phone and a comprehensive guide to unlocking the full potential of the device, please visit the following website:

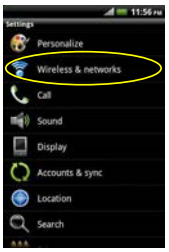
http://www.htc.com/uk/help/htc-explorer/#overview

If you encounter any further problems while using the phone, please contact the helpline given at the end of this document.

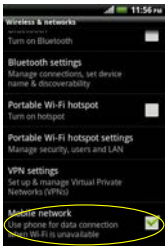
3. ABOUT THE APPLICATION

The Android application is very intuitive and simple to use. Before running the application, you must ensure that you are connected to the internet and the GPS is switched on. To confirm that you are connected to the internet, tap the 'Internet' icon on the home screen to launch the phone browser and try opening any webpage (e.g. yahoo.com); if the webpage opens up, it means you are connected to the internet. In this case, tap the phone's 'HOME' capacitive touch button to return to the home screen. To confirm if GPRS (internet) is enabled or not, tap the phone's 'MENU' capacitive touch button while on the home screen and select 'Settings' tab that pops on the bottom right of the screen, as shown below:

Choose 'Wireless & networks' from the list of settings that appear on the screen.



Then scroll down the page to check whether the option of 'Mobile network' is selected or not.



If it is selected, as shown, then the GPRS is switched on. If not, switch it on by checking this option. Confirm again by returning to the home screen by tapping "HOME" and opening any webpage using the phone's browser. If it still does not open, report the issue on the helpline given at the bottom of this document. If the website opens, go back to the home screen.

To check if the GPS is on or off, check the power control widget on the main screen (the dark grey bar at the top with five large symbols); if the GPS symbol is highlighted, as shown below, the GPS is on. If not, tap the GPS symbol to toggle it on, before starting the application.

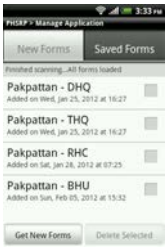
Once it is confirmed that the phone is connected to the internet and the GPS is switched on, tap the PHSRP icon on the home screen to start the application.

The application main screen has three buttons- 'Start New Form', 'Send Finished Forms' and 'Manage Application'- as shown below:

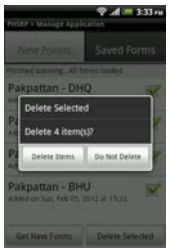
In order to start making entries, the application needs to first download the relevant forms. There are four forms for each district; one for each type of facility- BHU, RHC, THQ and DHQ. For the case of the phones handed out, the relevant forms have already been downloaded. However, in case there are any revisions made, all concerned officials will be notified that the forms will have to be updated. **Do not** delete the forms unless you are formally notified to do so.

3.1. How to update forms if notified

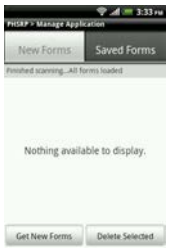
To update the forms on the application **if you are notified to do so**, tap the 'Manage Application' button. Considering Pakpattan as an example, the following screen will be displayed:



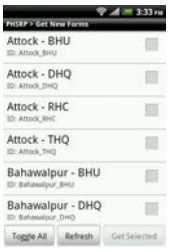
Select all the forms, or just the ones that you need to update as notified, by tapping on the checkboxes on the right, and tap the 'Delete Selected' button at the bottom right. A confirmation will be displayed as follows:



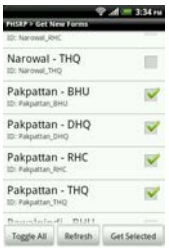
Tap 'Delete Items' to confirm and the selected forms will be deleted. If all the forms are deleted, the following screen will be displayed:



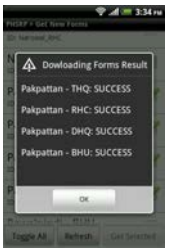
Now, tap 'Get New Forms', to retrieve the updated forms. The application will use the internet to list the updated forms of all districts for download as follows:



If you encounter an error at this point, it means you are not connected to the internet. Ensure that you are connected to the internet by following the instructions given previously and try again. If you encounter an error again, report the issue immediately on the helpline given at the end of this document. If there is no error and the above screen is displayed, scroll vertically to find the forms of your district and select them all by tapping the checkboxes on their right as shown:

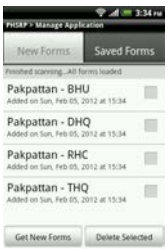


Then, tap 'Get Selected' to download the updated forms of your district. Once the forms are successfully downloaded, the following screen will be displayed:



If there is some sort of error at this point, try downloading the forms again. If you are still unsuccessful, report the issue on the helpline to get an immediate solution.

If all forms are successfully downloaded and the above screen is displayed, tap 'OK' and the following screen will be displayed:



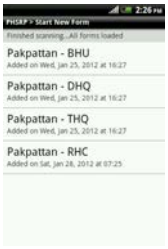
Tap the phone's 'BACK' capacitive touch button at the bottom of the screen to get to the main screen of the application again. You are all set to continue to making and submitting entries now.

3.2. How to fill a form

At this point, it is important to note that completing a form and submitting a form are two different tasks that are performed separately. Filling a form does not require an internet connection, so you can enter data from your inspection visits and save the completed forms regardless of whether the internet is working or not. However, submitting the forms requires an internet connection.

To start filling a form, tap the 'Start New Form' button from the main screen of the application.

The following screen will be displayed, prompting you to choose the type of facility:



Before moving on, it is important to note that if you want to close or discard the entry at any point before saving and exiting, tap the BACK capacitive button on the phone and choose 'discard entry'. If you tap BACK by mistake, simply tap 'Cancel' on the dialogue box that pops up.

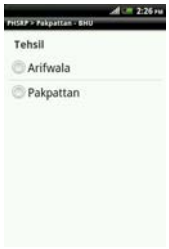
Furthermore, if you accidentally tap the phone's 'HOME' capacitive touch button and end up at the home screen while filling in the form, simply tap the PHSRP application icon again to load the application again and it will return you to the screen you were previously at in the form with all previous entries made on the form intact.

3.2.1. How to fill a BHU form

To fill a BHU form, choose the BHU form from the list shown above and the following screen will be displayed, instructing how to navigate through the form:

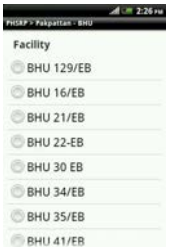
It is important to note here that you will be able to scroll back and forth within the form to check or change your entries before you complete the form, by scrolling laterally in one direction or the other, but whenever you scroll to a screen that requires numerical input from the keypad that pops up (as explained later), all numerical entries will be cleared and you will have to re-enter them.

Scroll laterally, as instructed, to start filling in the form. The next screen will allow you to choose the Tehsil in which the BHU is located, as shown below:

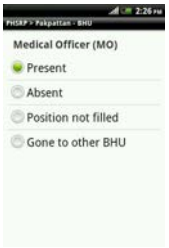


It is important to note at this time that some screens require at least one entry by the user, and you will not be able to move forward in the form unless it is made. To demonstrate, if you attempt to move forward in the form by scrolling laterally when it prompts you to enter the Tehsil in which the facility is located, the following message will appear on the screen:

You will have to select one of the options and then scroll laterally to move to the next screen. The next screen will require you to choose the BHU you are visiting from a list of all the BHUs present in that Tehsil. For demonstration, we select the Tehsil of Arifwala and scroll to the next screen. The following list is displayed:



Scroll vertically to find and choose the specific facility you are visiting, and scroll laterally to move to the next screen:



This screen relates to the availability status of the Medical Officer at the facility. An important thing to note here is that for all non-PRSP districts, the last option will not be shown on this screen as it does not apply to them. As Pakpattan is a PRSP district, the 'Gone to other BHU' option is available on the form.

Another important thing to note here is that all officers are required to make these entries from the perspective of a citizen visiting the facility- so even if the MO is on official leave or out on some official business at the time of the inspection visit, he/she would be marked absent. However, officers would also be required to take a note regarding the reason for absence of the MO in their diaries for such exceptional cases.

For demonstration, we choose Present and scroll laterally to the next screen:

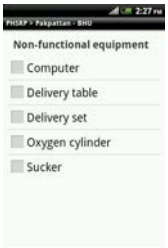
This screen requires you to check all the people not present at the BHU. As mentioned for the case of the MO, the officer will mark people absent based on the perspective of a visiting citizen- if someone is out on official business or on official leave or even if the position is not filled etc., the position holder will be marked as absent, and a note will be made in the officer's diary about the reason for absence for these exceptional cases.

If all the staff is present, you can scroll laterally to move to the next screen without marking any checkbox on this screen. The next screen requires you to mark tablets not available at the facility, as shown:

Scroll vertically and mark all the tablets that are out of stock at the BHU. If all tablets are present, scroll laterally to the next screen without marking any checkbox.

Repeat the same procedure for 'Injections', 'Syrups' and 'Other Medicines' in the subsequent screens as shown:

The next screen will require you to mark all equipment that is not functional. Unavailable equipment will also be marked as non-functional:



Leave the screen unmarked if all equipment is available and functional, and scroll laterally to the next screen.

The next screen will require you to tap in numerical values for the number of OPD cases last month, number of deliveries last month and number of Antenatal cases last month. A keypad will pop up at the bottom automatically so that you can enter the numbers. Tap on the entry bar of the next field to enter its number after you are done with the first one, and then move on to the third one after you are done with the second one. All **three** fields must be filled in order to move to the next screen. To get to the third field, you will have to scroll vertically lower down the page. While scrolling, ensure that you are avoiding the keypad, as scrolling over the keypad will not work.

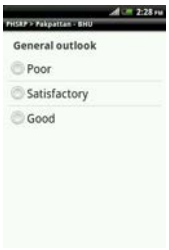
Once all three entries are filled, scroll laterally to move to the next screen. Once again, ensure that you avoid the keypad as scrolling laterally over the keypad will not work.

The next screen will require you to enter the **mobile numbers** of any **two** randomly selected delivery patients from the BHU records from **last month**. The entry fields are designed to detect invalid numbers, and the application will not let you move to the next screen unless you enter two **valid** mobile numbers.

Once the two mobile numbers are entered, scroll laterally to move to the next screen.

The next screen will require you to enter **mobile numbers** of any **two** randomly selected ANC patients from **last month**. The entry fields on this page are also designed to detect invalid numbers, and the application will not let you move to the next screen unless you enter two **valid** mobile numbers.

Once the numbers are entered, scroll laterally to move to the next screen:



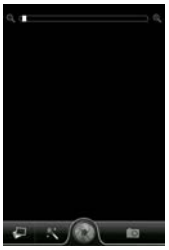
Choose the most appropriate option and scroll laterally to move to the next screen.

This screen will require you to mark which information was displayed in the BHU. Leave the screen unmarked and scroll laterally to the next screen if none of these were displayed at the facility.

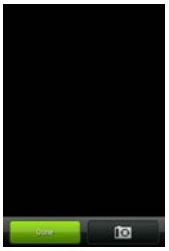
Mark the options appropriately and scroll laterally to move to the next screen.

The next screen will require you to take a clear picture of yourself with the essential staff present at the BHU, as shown below:

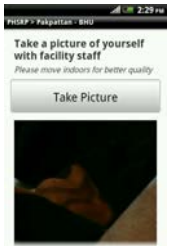
Tap the 'Take Picture' button to load the camera. For better picture quality, it is advisable to take the picture indoors and have someone take it for you. To take the picture, have that person tap the silver button in the centre-bottom of the screen, as shown below:



When the picture is taken, you will be given the option of retaking it if you are not satisfied with it. Tap the camera icon on the right to load the camera again and take a better picture, as shown below:



Once you are satisfied with the picture, tap the 'Done' button on the left, and you will be taken to the following screen:

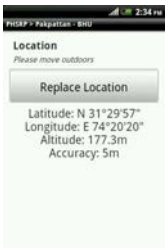


Scroll laterally to move forward. If, instead, you want to view the picture in full screen again, tap the picture preview box at the bottom, and you will be able to view it in full screen:

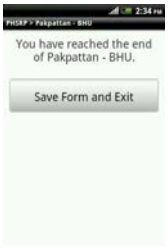
Tap the phone's 'BACK' capacitive touch button at the bottom to return to the previous screen. Once there, scroll laterally again to move to the next screen:

Tap 'Record Location' and the phone will record its location using GPS, network information and GPRS. It is advisable to move outdoors to record location as GPS signals are stronger outdoors. While you wait for the location to be recorded, you might see the accuracy radius values decreasing gradually:

When accuracy radius falls to 5 m, the following screen will be displayed:



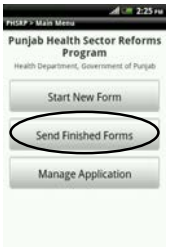
GPS satellites are not always in range hence it might take some time for the phone to narrow down its location. If, even after waiting for five to ten minutes, the phone is unable to record its location, ensure that the GPS is toggled on and try again. If, still, the phone is unable to record its location, contact the helpline immediately for quick resolution. Once the location is recorded, the above screen will be displayed. To move forward, scroll laterally again to get to the following screen:



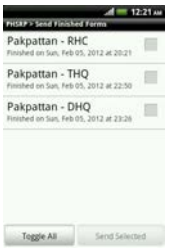
Tap 'Save Form and Exit' to complete the entry. A message will be displayed notifying you that the form was saved successfully and you will be taken back to the main screen of the application.

3.3. How to submit completed forms

Once you have completed the form (after pressing the 'Save Form and Exit' button), it needs to be submitted. After completing the form, tap the 'Send Finished Forms' button on the application main screen:



This will take you to a screen where all your completed and un-submitted forms are listed. Select the one you would like to submit or select all if you want to submit all, and tap the 'Send Selected' button on the bottom right of the screen.



If the submission was successful, a message will appear saying so, and the respective completed forms will vanish from this list. If all were selected and successfully sent, all will disappear. Tap the phone's 'BACK' capacitive button to return to the application's main screen.

If there is any error in submission, it can be because of the internet not working. In that case, confirm if the internet is working and try submitting the form/s again. If you are still unsuccessful, report your issue on the helpline given at the end of this document.

Tap the phone's 'HOME' capacitive touch button to exit the application and return to the home screen of the phone once you have successfully submitted the forms.

Helpline: 0308 4091080

G Training Manuals For Dashboard Use

MANUAL
FOR
HEALTH FACILITY INFORMATION MANAGEMENT SYSTEM
(ONLINE DASHBOARD)

Directorate General Health Services
Supported By

Punjab Health Sector Reforms Program (PHSRP)
International Growth Centre (IGC)
Lahore University of Management Sciences (LUMS)

CONTENTS

1. Introduction

The Health Department of Government of the Punjab is committed to adopting state-of-the-art technology to strengthen governance and improve service delivery for all citizens.

For this purpose, the Punjab Health Sector Reforms Program (PHSRP), with technical assistance from International Growth Centre (IGC) Team, is supporting DGHS and district health managers in strengthening the internal monitoring system of the Health Department. This is being done by introducing a mobile phone based, online information management system.

This initiative will improve the internal information transmission within the Health Department and will ensure that timely, authentic and actionable information is sent quickly from individual facilities to district and provincial health managers on such crucially important issues as absenteeism, medicine stock outs, availability and functionality of equipment etc.

Android-based smartphones have been provided to those district supervisory officers, such as Executive District Health Officers (EDOs), District Health Officers (DOs), and Deputy District Health Officers (DDOs), who have been tasked with the collection of performance related data from Basic Health Units (BHUs), Rural Health Centers (RHCs) and Tehsil and District Headquarters (THQs and DHQs).

The report submitted by these officers through the phone will be recorded on a website, known as the 'Dashboard', and automatically analyzed for use by managers at various levels. It is expected that this information will become a powerful tool for management both for district and central level officials. This is expected to bring about marked improvement in health service delivery management, particularly at primary and secondary levels of healthcare, leading to better health outcomes for the poor and disadvantaged in the province.

At Directorate General Health Services, Director, District Health Information System (DHIS), supported by the PHSRP and IGC team, is the focal person for implementation of the program at the provincial level. Overall responsibility for the program at the district level lies with EDOs, and Statistical Officers (SOs) are the designated focal persons for managing the system at the district level.

This manual explains what information is available on the online dashboard and how it is displayed, to help managers at different levels to utilize this powerful tool to its full potential in order to improve health care in the province.

2. The Dashboard

The online dashboard can be accessed any time over the internet through the following link:

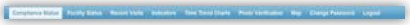
punjabmodel.gov.pk/phsrp/dashboard

When you open the link, the following page will be displayed, prompting you to enter your username and password, and giving you the option of saving these credentials for automatic login the next time you open the link, as shown in Figure 1.

Figure 1

To access the dashboard, you have to enter the unique username and password already communicated to you and click on 'Login'. Once successfully logged in, you can also change your password for the dashboard by accessing the Change Password section in the blue bar. When you are done using the dashboard, you can click on 'Logout' to end the session.

Figure 2



As shown in Figure 2, the blue bar near the top of the page contains all the major sections of the dashboard, allowing you to effortlessly navigate from one part of the online tool to another.

One major feature of this tool is the 'Print' button/icon which is located to the right, just below the blue bar. Clicking this allows you to take a snapshot of whatever is currently being displayed on the dashboard and print it out.

It is important to note that there are two levels of access for the dashboard- the district level and the provincial level. All DCOs, EDOs, DOs and DDOs have access to the district level but not the provincial level, ergo when they log in, they are shown the district level by default. The relevant higher up senior officers, however, have access to the district level as well as the provincial level, so when they log in, their default view is the provincial level, but they can also choose to access the district level by choosing from a drop down list of districts near the top of the webpage.

2.1. The District Level

2.1.1. Compliance Status

The first page that is displayed when you log in the dashboard is the Compliance Status section. Officers can use this section to track their compliance performance for the current month as well the months before. They can also gauge their current standing compared to fellow health officers in the district with respect to compliance.

The most prominent characteristics of this page are the 2 bar charts and the table below them.

The first bar chart represents the percentage compliance of all the health officers in the district for the **last calendar month**, disaggregated by facility type. This is calculated as follows:

Percentage compliance= (total visits performed last month / visits assigned last month) x 100

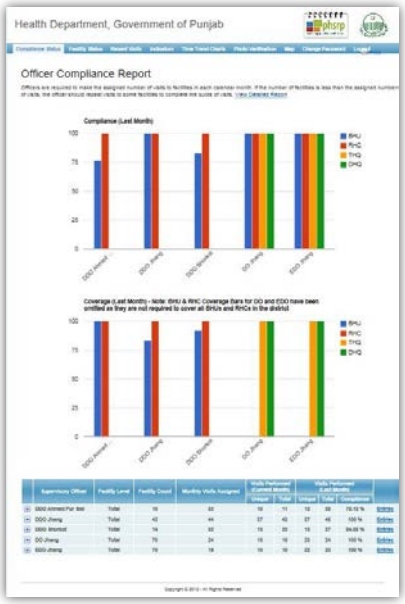
The bars are color coded by facility type, as explained by the legend displayed on the page. Compliance is 100% if the officer performed 100% of the visits assigned to him or more.

The second bar chart represents the percentage coverage of all health officers in the district for the **last calendar month**, disaggregated by facility type. This is calculated as follows:

Percentage coverage= (1 – (no. of assigned facilities not visited by **any** officer last month/ facility count)) x 100

Once again, the bars are color coded by facility type, as explained by the legend displayed on the page, as shown in Figure 3.

Figure 3

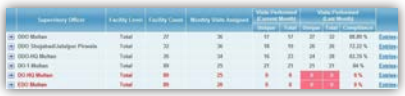


The distinction between what the two charts convey is important and is easily explainable using an example. Suppose there are 10 facilities in an officer's jurisdiction and he is assigned a total of 10 visits. If he visits every single facility once, his compliance as well as his coverage will be 100%. If he visits only 1 facility 10 times during the month, his compliance will still be 100% but his coverage will be 10%. Similarly, suppose if the assigned visits are 20 and the facility count is still 10; if he visits each facility once (leading to a total of 10 visits), his compliance will be 50% but his coverage will be 100%. Officers should strive for 100% compliance as well as the maximum possible coverage (which can be less than 100% only in cases where facility count exceeds the number of assigned visits).

The table below the charts gives detailed information regarding compliance figures. The '+' icon before every officer's designation in the 'Supervisory Officer' column can be clicked to expand the table to show information disaggregated by facility type. The information displayed in the table includes the facility count, monthly assigned visits, unique and total visits performed during the current month, unique and total visits performed last month, and the percentage compliance for last month, for every officer in the district, disaggregated by facility type as well as in total.

For cases in which compliance in the last calendar month is low, the table is highlighted red, as shown in Figure 4.

Figure 4



The last column provides hyper links, allowing you to jump directly to the relevant entries in the 'Recent Visits' section. The Recent Visits section will be explained in detail later on.

If you are interested to see compliance figures for months before the last calendar month, you can click on the 'View Detailed Report' hyperlinked text located near the top of the page.

Note: Should you find that a visit to a particular facility is not being displayed on the dashboard despite being successfully submitted from the Android smart phone allotted to you, please convey it immediately at the helpline given at the end of this document.

2.1.2. Facility Status

The Facility Status section gives you a list of all the facilities in the district, arranged by the **date of last visit** with the oldest visited at the top. It is designed to enable you to keep track of facilities that are being neglected. The facilities are color coded, according to the legend displayed on the page, as shown in Figure 5.

Figure 5

The page has different tabs for the different facility types. Each tab displays a table which displays the facility name, the Tehsil/Town it is located in, the designation of the officer who last visited the facility, the date of the last visit and the number of days since the last visit. The corresponding columns also have filters in-built that allow you to view selective information if you choose to.

The table also contains a column for Summary Report. Clicking the icon in this column for any row will take you to a page displaying details regarding the last visit to the facility as well as the second last visit, in addition to Tehsil variable averages (from 30 days from the last visit). Figure 6 shows a cropped screenshot of the page.

Figure 6

Clicking on the icon in the Recent Visits column for any facility, instead, will take you to the Recent Visits section showing you a list of all entries made for that facility, as shown in Figure 7.

Figure 7

Officers should ensure that all the facilities listed in the Facility Status section are green- some can be blue for cases in which the facility count is more than the assigned visits. Orange or red rows represent neglected facilities and they should be visited as soon as possible.

2.1.3. Recent Visits

The Recent Visits section lists all entries as they come in, with the latest submitted on top. There are different tabs for different facility types. Each facility type tab contains a date filter, which allows you to view entries submitted during a particular time period, and a table consisting of entries, as shown:

Figure 8

To view entries submitted between certain dates, choose the start and end dates from the drop down calendars displayed by clicking on the two white text boxes immediately below the facility type tabs respectively, and click the 'Filter by Period' button.

Some of the entries in the table might be highlighted red, as shown in the above screenshot. These represent facilities where significant staff absence was reported. The table also allows you to display only the highlighted entries or the non-highlighted entries separately, in addition to displaying them all together. The drop down filter for the column labeled 'Absence' can be used to toggle between the selections.

The table also contains information that includes the facility name, the Tehsil/Town it is located in, the visiting officer, the date of visit, the availability status of the MO and the availability status of other staff. It also provides filters for all these categories for selective searches.

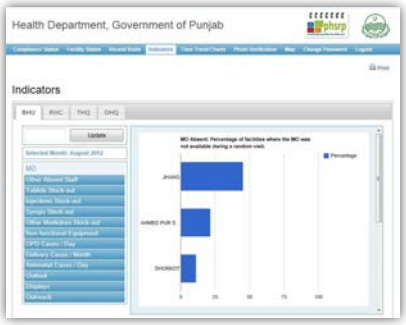
The Summary Report icon at the end of every entry in the table can lead you to a page displaying details regarding the last visit to the facility as well as the second last visit, in addition to Tehsil/Town variable averages (from 30 days from the last visit) as already depicted in Figure 6.

As already mentioned, should you find that a visit to a particular facility is not being displayed on the dashboard despite being successfully submitted from the Android smart phone allotted to you, please convey it immediately at the helpline given at the end of this document.

2.1.4. Indicators

The Indicators section displays charts comparing performance of the different Tehsils/Towns based on the various indicators reported during facility visits. Once again, there are different tabs for different facility types, and different indicators, in some cases, for different tabs. The following screenshot should give you an idea of what the page looks like:

Figure 9



It is important to note that while there are multiple BHUs, RHCs and THQs in each district, the number of DHQs is one or zero. Hence, instead of a comparison across Tehsils/Towns as for the case of BHUs, RHCs and THQs, the DHQ section compares DHQs **across districts**. Furthermore, all indicator charts that display data expressed in percentages in the DHQ section have an additional **red** bar which reflects percentage compliance in every district. The compliance bars are intended to be a gauge of how many visits' data is used to derive the charts- ergo, the higher the compliance, the more reflective is the value of the variable of the actual situation in the corresponding district.

For all tabs, there is a text box allowing you to choose which month you want to see the data for. The page displays charts for the last calendar month by default. If you want to access charts for some previous month, you need to click on the white text box, select the month and year from the drop down

menu, click 'Done', and then click the 'Update' button located to the immediate right.

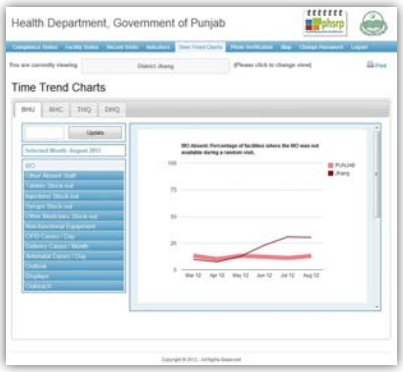
Most indicators in the list have multiple charts that are displayed when you click on any one of them. All charts have descriptive labels that clearly indicate what they represent. Tables 1 through 4 in the appendix show how the charts are arranged for each facility type.

These charts can prove to be a very powerful tool for Tehsil-wise comparison based on the different performance related indicators. However, if taken in isolation, interpretations derived from them may be misleading. For example, if Tehsil 'A' shows 0% MO absence while Tehsil 'B' shows 20% MO absence, it doesn't necessarily imply that Tehsil 'A' is better in MO attendance than Tehsil 'B'. It is possible that only a single visit was performed in Tehsil 'A' in the entire month- during which the MO was present- while, out of the 10 visits performed in Tehsil 'B', the MO was absent in only 2. Ergo, the information displayed in the charts should always be interpreted while considering compliance figures.

2.1.5. Time Trend Charts

The Time Trend Charts section contains line graphs representing the change over time in all the indicators of the different facility types present in the Indicators section as shown in Tables 1, 2, 3 and 4 in the appendix. The general layout of this section is very similar to that of the Indicators section, with the same indicator tabs and option to select a different month for all facility types. However, there is one key difference; the charts contain two lines- a thin one representing the district average and a thick one representing the provincial average- allowing you to compare the average **district** performance on each indicator to the **provincial** average, over time, instead of comparing across Tehsils/Towns of the same district. Figure 10 shows how the webpage might look.

Figure 10



You can also compare the performance of any **Tehsil/Town** compared to the **district** average over time. This can be done by clicking the drop down button near the top of the page and selecting the Tehsil/Town you want to compare with the district average. In the charts that will be displayed as a result, the thick line would represent the district's average and the thin line would represent the Tehsil/Town average.

These charts can prove to be very useful in observing and comparing trends in different indicators over time, at the provincial, district, as well as the Tehsil/Town level.

2.1.6. Photo Verification

To verify staff presence, the smart-phone Performa requires officers to take pictures of the essential staff present at the facility they are visiting. The Photo Verification section displays all these, sorted by the most recent visit, by officer designation. Figure 11 shows the layout of the page.

Figure 11

You can view the full size version of any picture by clicking on it. Health officers responsible for supervision of BHUs, RHCs, THQs and DHQs are advised that the pictures submitted should not be blurry or unclear in any way for the convenience and effectiveness of photo verification.

2.1.7. Map

When you click on the tab for the Map section, a separate window (or tab, depending on your browser) will open, displaying a map of Pakistan and its surrounding areas as shown in Figure 12.

Figure 12

For completing an entry for a facility visit, the smart-phone Performa requires the supervisory officer to record the location of the facility using the phone's GPS. All successfully submitted entries show up on this map when you zoom down to individual district.

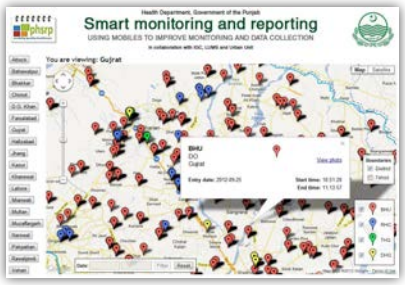
In order to view entries for any district, you need to click on the relevant district tab from the list on the left. Once you zoom in, all the relevant entries will show up as place-marks color-coded with respect to the facility type, as shown in Figure 13.

Figure 13

You can zoom further in or out using the zooming tool in the upper left corner of the map. The map also allows you to show or hide District and Tehsil boundaries, and even switch between Map and Satellite view. Furthermore, the date filter allows you to see only those entries submitted during a certain time period.

Clicking on any place-mark reveals a few details regarding the entry that include the supervisory officer's designation, the date the entry was made, the start and end time of the visit and a link to the picture taken for the entry, as shown in Figure 14.

Figure 14



The map allows for spatial review of the coverage and compliance in the District or Tehsil/Town, which can prove to be very useful for circumstances in which information regarding the location and spread of the facilities is crucial.

2.2. The Provincial Level

As already mentioned, when you log in to the dashboard with an account that has provincial level access as well as district level access, your default view of the dashboard is the provincial level view. However you can access the district level view for any district by choosing it from the drop down list that appears when you click the 'Punjab' button, which is right below the blue bar near the top of the page.

The Recent Visits and Photo Verification sections in the provincial level view are blank as the usefulness of a combined list of entries or verification pictures coming in from all districts is very limited.

Apart from that, the Map section for both the levels is exactly the same.

2.2.1. Compliance Status

Once again, the first page displayed after a successful login is the Compliance Status section. This is just like the Compliance Status section in the district level view except that instead of a comparison across Tehsils/Towns in a district, you have a comparison of compliance **across districts**.

The bars in the two charts are color-coded in the same way as in the district level view, and the table below the charts gives detailed information regarding compliance figures for districts, rather than supervisory officer. Again, the '+' icon can be clicked to expand the table to show information disaggregated by facility type. The information displayed in the table includes the facility count, monthly assigned visits, unique and total visits performed during the current month, unique and total visits performed last month, and the percentage compliance for last month, for every district, disaggregated by facility type as well as in total.

Figure 15 shows how the page might look like.

Figure 15

Districts with low compliance in the last calendar month will be highlighted in red. The last column provides hyperlinks, allowing you to jump directly to the relevant entries in the 'Recent Visits' section, as in the district level view.

Moreover, if you are interested to see compliance figures for months before the last calendar month, you can click on the 'View Detailed Report' hyperlinked text located near the top of the page, in same way.

This section is very useful for senior officials to track the compliance and coverage status of all districts and compare them if need be.

2.2.2. Facility Status

The Facility Status section in the provincial level view is radically different from that in the district level view, as apparent from Figure 16.

Figure 16

The page displays a single bar chart representing the percentage of facilities that are being neglected in each district. The bars are color-coded based on the facility type.

The criterion for a facility to be considered neglected is that it is not visited by any supervisory officer in the current month as well as the last two calendar months. Senior officials can easily identify which district has the highest percentage and take appropriate measures to rectify the situation.

2.2.3. Indicators

The Indicators section in the province level view is very similar to that in the district level view in terms of layout and structure. The variables are exactly the same as those in the district level view, as detailed in Tables 1, 2, 3 and 4 in the appendix.

One major difference between the two views, however, is that instead of a comparison across **Tehsils /Towns** in a district, the provincial level charts compare performance across **districts** for all the indicators.

Also, indicator charts in the province level view contain extra red bars representing compliance for the BHUs, RHCs and THQs as well as the DHQs, whereas this is only true for DHQs in the district level view of the Indicators section. As previously explained, the compliance bars serve as a gauge of how many visits' data is used to derive the charts- meaning that the higher the compliance, the more the value of the variable is reflective of the actual situation in the corresponding district

Figure 17 depicts a screenshot of the section.

Figure 17

This section can prove very useful to track performance of and across districts in terms of various indicators.

2.2.4. Time Trend Charts

The Time Trend Charts section in the province level view is exactly the same as that in the district level view, except that there isn't an extra line for any district on any of the charts; just a thick line representing the trend of provincial averages for the same indicators over time, as depicted in Figure 18.

Figure 18

As already mentioned, you can move to the district level view if you want a comparison of the provincial average with a district's average, or even to the Tehsil/Town level view if you want a comparison of the district average with a Tehsil/Town's average, over time.