Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

Article

ODK Scan: Digitizing Data Collection and Impacting Data Management Processes in the Tuberculosis Control Program of Pakistan

Syed Mustafa Ali [1,*], Rachel Powers [2], Naveed Anjum [1], Jeffrey Beorse [3], Muhammad Ishaq [1], Javariya Aamir [1] and Richard Anderson [3]

[1] Monitoring Evaluation and Learning Unit, Mercy Corps, Pak Palace, Murree Road, Rawal Chowk, Islamabad 44000, Pakistan; nbuzdar@mercycorps.org (N.A.); muishaq@mercycorps.org (M.I.), jaamir@mercycorps.org (J.A.)
[2] Information Systems, Village Reach, Seattle, WA 98102, USA; rachel.powers@villagereach.org
[3] Computer Science and Engineering, University of Washington, Seattle, WA 98105, USA; jbeorse@cs.washington.edu (J.B.), anderson@cs.washington.edu (R.A.)
* Correspondence: mustafa_30_84@yahoo.com; Tel.: +92-321-413-7073

Abstract: The present grievous situation of the tuberculosis disease can be improved by efficient case management and timely follow-up evaluations. With the advent of digital technology this can be achieved by quick summarization of the patient-centric data. The aim of our study was to assess the effectiveness of the ODK Scan paper-to-digital system during testing period of three months. A sequential, explanatory mixed-method research approach was employed to elucidate technology use. Training, smartphones, application and 3G enabled SIMs were provided to the four field workers. At the beginning, baseline measures of the data management aspects were recorded and compared with endline measures to see the impact of ODK Scan. Additionally, at the end, users' feedback was collected regarding app usability, user interface design and workflow changes. 122 patients' records were retrieved from the server and analysed for quality. It was found that ODK Scan recognized 99.2% of multiple choice bubble responses and 79.4% of numerical digit responses correctly. However, the overall quality of the digital data was decreased in comparison to manually entered data. Using ODK Scan, a significant time reduction is observed in data aggregation and data transfer activities, however, data verification and form filling activities took more time. Interviews revealed that field workers saw value in using ODK Scan, however, they were more concerned about the time consuming aspects of the use of ODK Scan. Therefore, it is concluded that minimal disturbance in the existing workflow, continuous feedback and value additions are the important considerations for the implementing organization to ensure technology adoption and workflow improvements.

Keywords: mHealth; ODK scan; mobile health application; digitizing data collection; data management processes; paper-to-digital system; technology-assisted data management; treatment adherence

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

1. Introduction

Tuberculosis (TB) is still a global and grave public health concern. It is the reason for ill-health of millions and cause of numerous deaths each year worldwide [1]. Despite the availability of effective treatment and strategies to improve access to treatment regimens, multidrug-resistant and extremely drug-resistant TB cases are emerging due to the lack of adherence to treatment protocols. This non-adherence is mainly because of the longer treatment course (6 to 8 months) and unpleasant side effects of the drug treatment [2].

Pakistan is among the TB high-burden countries, with the emergence of a large number of new TB cases (approximately 420,000 every year) and prevalence of multidrug-resistant TB. Pakistan ranks fifth for drug-susceptible TB and fourth for multidrug-resistant TB [3]. The burden of disease caused by TB and effectiveness of different programmatic interventions are better understood through an interpretation of the available epidemiological data [4]. The program management function, Monitoring and Evaluation (M&E), makes the best use of the available resources (including data) to guide and support programmatic decision-making. The M&E system plans and conducts data collection and analysis to inform program managers for keeping the performance measures on track [5]. However, unfortunately this immense resource is either underutilized or not used at all [4].

Health information system used in the disease-specific health programs are designed to collect patient-centric data to generate the actionable information to guide decision-making. Likewise, mostly paper-based patient recording and reporting systems are maintained and managed in the TB control programs [6].

Fidelity to treatment protocols is ensured through case management (DOTS strategy) and follow-up evaluations (sputum specimen for microscopic examination or culture at specific points of treatment period) [7]. Generally, in the TB control programs, paper-based patient records are used to prepare periodic reports of the patient enrolment, retention and treatment progress. These data summaries are also sent to different management levels for aggregation, analysis and reporting [6]. However, paper-based recording and reporting systems suffer from some inherent challenges, which include: checking data quality, data entry and aggregation, allotting treatment outcome correctly, filling gaps in data, analysing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

data and generating actionable information. These data management activities are inefficient, work-intensive and time consuming [8]. However, the advent of digital technology has positively impacted the organizational workflow [9].

Electronic Health Information Systems (eHIS) have generated better results on the three fronts of TB control programs: patient care, resource management and disease surveillance [8]. Increasingly, mobile health (mHealth) interventions are grabbing the attention of the public health professionals for their ability to improve the public health situation [10]. mHealth refers to the application of portable hand-held devices and communication capabilities in the health care programs [11]. During the past decade mHealth has been continuously evolving within the eHealth domain, where it has offered great potential to overcome many traditional barriers in delivering care and reducing the distance, time and cost [12]. Fortunately, most of the developing countries have already taken mHealth initiatives to unleash the power and reach of smartphones and considered them as an important resource for front line workers [10].

Among various reported uses of mHealth tools the most reported uses are: data collection, electronic decision support, provider's education and training, and communication [10]. However, a particular use of mHealth for data collection has more likelihood to improve the public health in low- and middle-income countries [13]. Noteworthy are those unpredictable implementation challenges and complicated situational factors that affect an introduction of mHealth application in primary healthcare settings of the developing world [14,15].

In the Public-Private mix model of the TB control program of Pakistan, chosen from freely available and open- sourced mHealth applications, we tested the ODK Scan paper-to-digital system for scanning paper forms and digitizing the data. Digitization of the data collection was experimented with an objective of reducing the time and resources spent on the data entry work. Additionally, the ODK Scan was chosen because it allows supervisors and managers to assess the data quality by viewing image snippet. The aim of this pilot study was to: (1) assess data quality aspects at both the app and user levels (2) assess the quantitative impact of ODK Scan on the data management processes (3) collect feedback

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

from the application user's experience to know app usability, user interface design and workflow changes.

2. Material and Methods

2.1. Description of Technology

Open Data Kit (ODK) Scan is an Android-based mobile application and a part of the ODK 2.0 suite of tools [9, 16]. It uses the device's built-in camera to capture an image of a scan-compatible paper form, which it feeds into an image processing algorithm. This algorithm segments the form into small image snippers that each contains a single data field from the original form [17]. The algorithm then converts hand-filled bubbles, structured numbered boxes, and QR codes into an editable record in the ODK 2.0 database. ODK Scan does not process free hand written text; instead it creates an empty text field in the record that can be manually filled in. Each digitized field includes a copy of its associated image snippet from the original form, which can be used to verify digitized fields or manually transcribe hand written text. Once the record is in the ODK 2.0 database, users can use the other tools, such as ODK Survey and ODK Tables to view, edit, verify, aggregate, and sync the data or create customized services, interactions and reports with the data [16].

2.2. ODK Scan Workflow

Before ODK Scan is to be used, a scan-compatible form must be designed using the form designer. Paper copies of the form image are printed out and the form's definition files are copied to the Android-based device. After the form is filled, the user takes a picture using the built-in camera of the smartphone. ODK Scan detects the data fields and auto-digitizes them. The user can then launch ODK Survey to view and verify data fields, and editing if necessary. After a user is sure about the data quality, the patient record is synced with an aggregate server. If no internet connection is available, the user can delay synchronization and continue collecting data until connectivity is restored. At any time, the user can view all synced records on the smartphone and run customized reports to view actionable information. For this trial we created a report that generates information about patients whose follow up appointments are either overdue or due in the coming days. Aside from

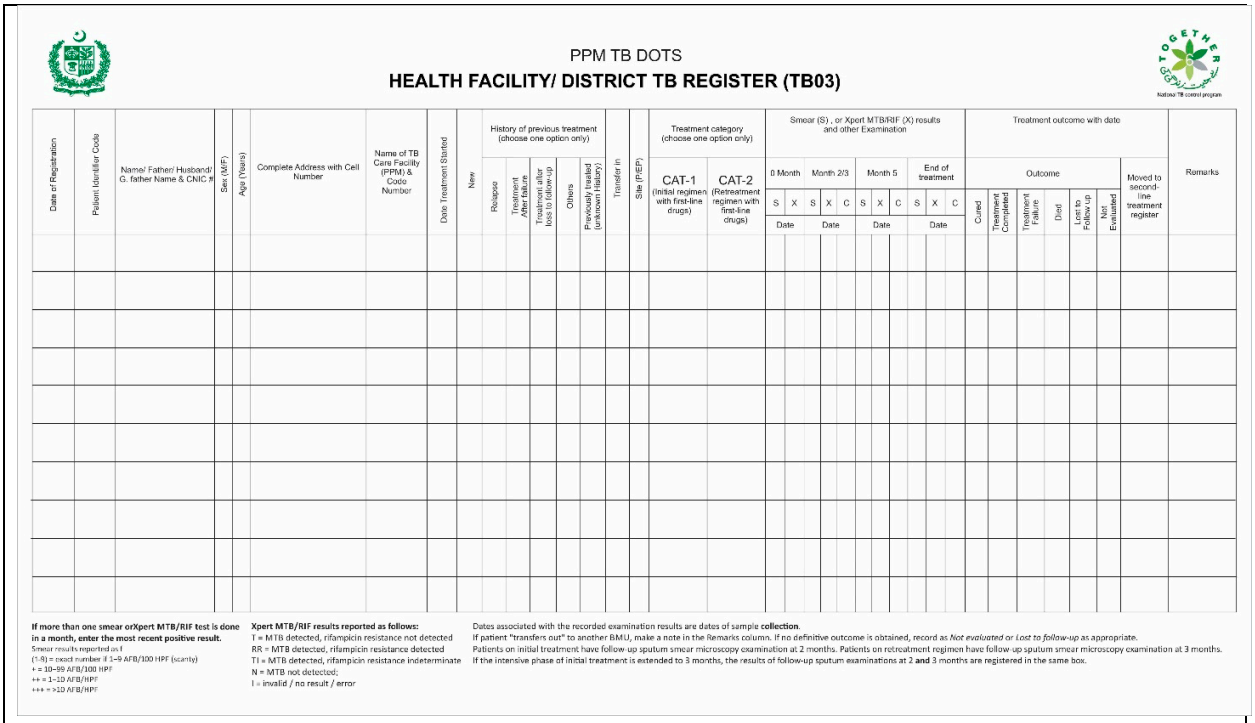
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

the syncing step of the ODK Scan workflow, all steps can be undertaken without internet connection.

2.3. Intervention

Among nine recording and reporting formats used in the TB control program of Pakistan, one recording format (TB03) was selected (Figure 1). Its scan-compatible form was designed using the ODK Scan Form Designer to create machine-readable fields for the app to recognize (Figure 2). The scan-compatible form consists of three types of machine-readable data fields: structured number boxes, bubbles and QR codes, one non-machine readable data field: hand written text fields (Figure 3). After the form design was prepared, it was pre-tested, necessary modifications in form design were made, and the forms were printed and bound into a register book. The intervention started at the beginning of May 2016 and lasted for three months in district-level field settings. Informed consent was sought from four of the field workers, belonging to two enrolled districts (Chakwal and Hafizabad) of the TB control program of Pakistan. The field workers (District Field Supervisors) were trained in using the Scan-compatible TB 03 form and the app workflow and were given smartphones (Samsung Galaxy J7 and Huawei Y6) and scan-compatible TB03 registers. Additionally, we provided them 3G-enabled mobile SIMs for internet connectivity and monthly top-up for sharing data via internet. We imposed no restriction on accessing or downloading any other mobile applications. Field workers were allowed to use smartphone, including making calls, text-messaging and internet browsing.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1



PPM TB DOTS

HEALTH FACILITY/ DISTRICT TB REGISTER (TB03)

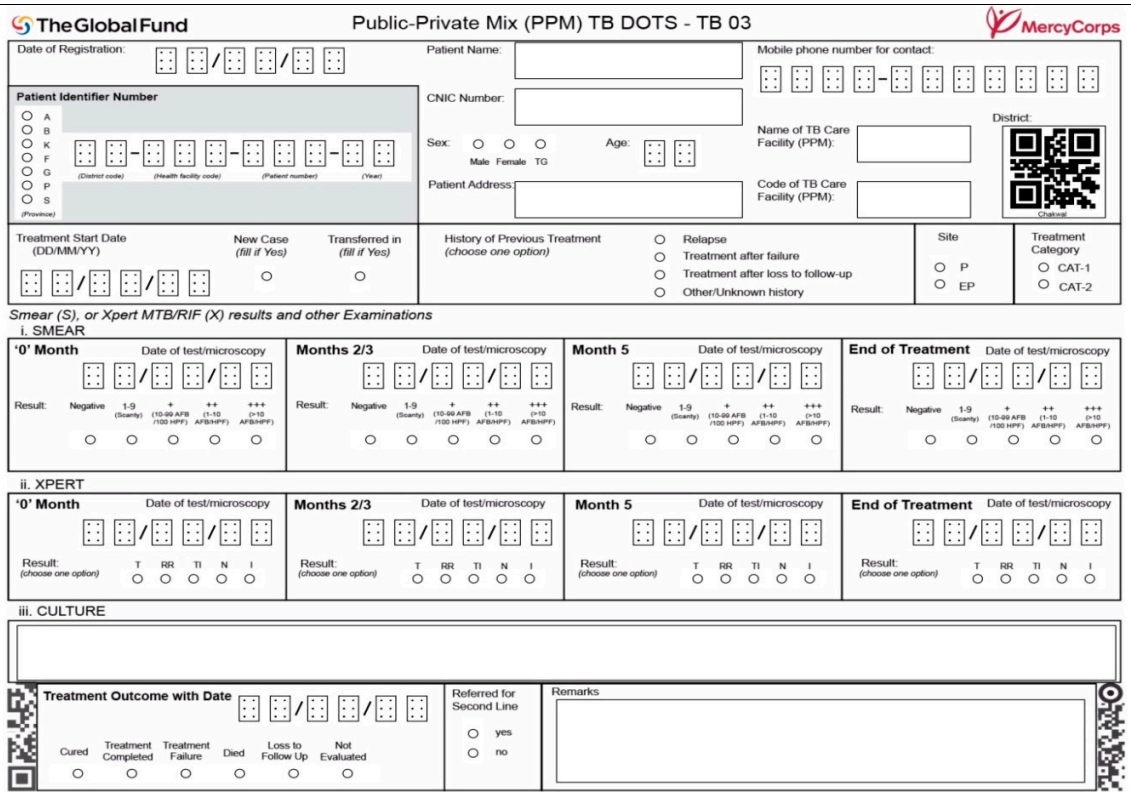
Date of Registration	Patient Identifier Code	Name/ Father/ Husband/ G. father Name & CNIC #	Sex (M/F)	Age (Years)	Complete Address with Cell Number	Name of TB Care Facility (PPM) & Code Number	Date Treatment Started	History of previous treatment (choose one option only): New	Relapse	Treatment After failure	Treatment after loss to follow-up	Others	Previously treated (unknown History)	Transfer in	Site (P/EP)	Treatment category (choose one option only): CAT-1 (Initial regimen with first-line drugs)	CAT-2 (Retreatment regimen with first-line drugs)	Smear (S) , or Xpert MTB/RIF (X) results and other Examination: 0 Month S	X	Date	Month 2/3 S	X	C	Date	Month 5 S	X	C	Date	End of treatment S	X	C	Date	Treatment outcome with date: Outcome: Cured	Treatment Completed	Treatment Failure	Died	Lost to Follow up	Not Evaluated	Moved to second-line treatment register	Remarks

If more than one smear orXpert MTB/RIF test is done in a month, enter the most recent positive result.
Smear results reported as f
(1-9) = exact number if 1–9 AFB/100 HPF (scanty)
+ = 10–99 AFB/100 HPF
++ = 1–10 AFB/HPF
+++ = >10 AFB/HPF

Xpert MTB/RIF results reported as follows:
T = MTB detected, rifampicin resistance not detected
RR = MTB detected, rifampicin resistance detected
TI = MTB detected, rifampicin resistance indeterminate
N = MTB not detected;
I = invalid / no result / error

Dates associated with the recorded examination results are dates of sample **collection**.
If patient "transfers out" to another BMU, make a note in the *Remarks* column. If no definitive outcome is obtained, record as *Not evaluated* or *Lost to follow-up* as appropriate.
Patients on initial treatment have follow-up sputum smear microscopy examination at 2 months. Patients on retreatment regimen have follow-up sputum smear microscopy examination at 3 months.
If the intensive phase of initial treatment is extended to 3 months, the results of follow-up sputum examinations at **2** and 3 months are registered in the same box.

Figure 1: Original TB03 Form used in Routine Data Management Process



The Global Fund — Public-Private Mix (PPM) TB DOTS - TB 03 — MercyCorps

Date of Registration: ☐☐/☐☐/☐☐

Patient Name:

Mobile phone number for contact:

Patient Identifier Number
A B K F G P S (Province)
(District code) – (Health facility code) – (Patient number) – (Year)

CNIC Number:

Sex: Male Female TG

Age:

Name of TB Care Facility (PPM):

District:

Patient Address:

Code of TB Care Facility (PPM):

Treatment Start Date (DD/MM/YY): ☐☐/☐☐/☐☐

New Case *(fill if Yes)*

Transferred in *(fill if Yes)*

History of Previous Treatment *(choose one option)*
- Relapse
- Treatment after failure
- Treatment after loss to follow-up
- Other/Unknown history

Site: P / EP

Treatment Category: CAT-1 / CAT-2

Smear (S), or Xpert MTB/RIF (X) results and other Examinations

i. SMEAR

'0' Month — Date of test/microscopy	Months 2/3 — Date of test/microscopy	Month 5 — Date of test/microscopy	End of Treatment — Date of test/microscopy
Result: Negative / 1-9 (Scanty) / + (10-99 AFB /100 HPF) / ++ (1-10 AFB/HPF) / +++ (>10 AFB/HPF)	Result: Negative / 1-9 (Scanty) / + (10-99 AFB /100 HPF) / ++ (1-10 AFB/HPF) / +++ (>10 AFB/HPF)	Result: Negative / 1-9 (Scanty) / + (10-99 AFB /100 HPF) / ++ (1-10 AFB/HPF) / +++ (>10 AFB/HPF)	Result: Negative / 1-9 (Scanty) / + (10-99 AFB /100 HPF) / ++ (1-10 AFB/HPF) / +++ (>10 AFB/HPF)

ii. XPERT

'0' Month — Date of test/microscopy	Months 2/3 — Date of test/microscopy	Month 5 — Date of test/microscopy	End of Treatment — Date of test/microscopy
Result: *(choose one option)* T / RR / TI / N / I	Result: *(choose one option)* T / RR / TI / N / I	Result: *(choose one option)* T / RR / TI / N / I	Result: *(choose one option)* T / RR / TI / N / I

iii. CULTURE

Treatment Outcome with Date ☐☐/☐☐/☐☐

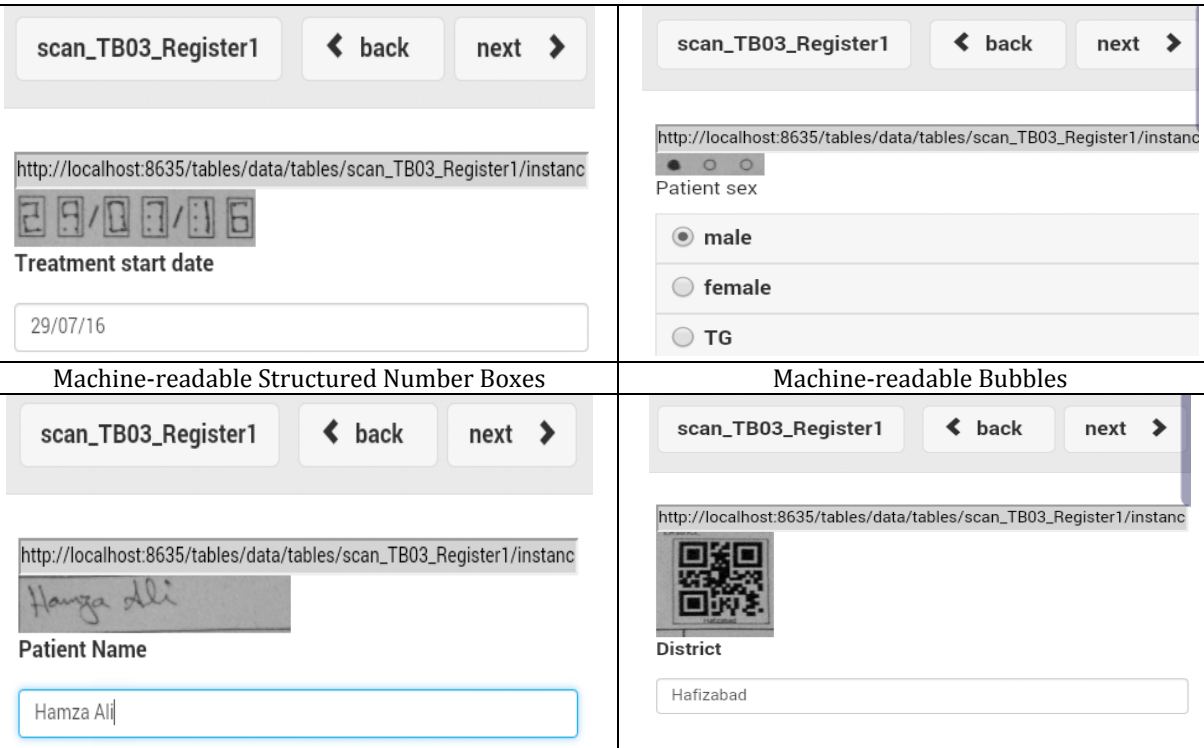
Cured / Treatment Completed / Treatment Failure / Died / Loss to Follow Up / Not Evaluated

Referred for Second Line: yes / no

Remarks

Figure 2: Scan-compatible TB03 form used in ODK Scan Paper-to-digital System displaying Bubbles, Number Boxes and Text Fields

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1



Machine-readable Structured Number Boxes	Machine-readable Bubbles
User transcribed Text Field	Machine-readable QR Code

Figure 3: Formats of Data Fields on ODK Scan-compatible TB03 form

2.4. Design

A sequential, explanatory mixed-method research approach was employed to elucidate the technology use for an effective implementation of the ODK Scan paper-to-digital system. This was done by applying the following three methods: (1) Assessment of the data quality was done at both app and user-level. At the application level, accuracy of ODK Scan's ability to recognize bubbles and number boxes was measured by comparing the raw Scan values (the digitized values which Scan's algorithm produced) with the data appearing in image snippets of the corresponding data fields. At the user-level, the workflow allowed for users to improve the quality of the data through validation and editing data fields manually. User-level data quality was assessed by comparing the manually and digitally entered records with hard-copies of the TB03 register (2) Time measures regarding form filling, data transfer, aggregation and verification were recorded in the routine data management process using direct observation. They were then compared with the endline measures for the same data management processes using the ODK Scan system. (3) Face-to-face interviews, using a semi-structured questionnaire, were carried out with each of the field

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

workers. Their app usability and workflow modification experiences were explored to assess the effectiveness of the ODK Scan paper-to-digital system from a usability perspective.

3. Findings

3.1. Data Quality

3.1.1. App-level quality measure: During an intervention period of three months, 122 patients' paper-based records were scanned, digitized and checked for quality. All of this data was synced and available on the cloud-based ODK aggregate server for access and remote analysis. The first task was to assess ODK Scan's ability to correctly parse the user filled machine-readable data fields, the structured number boxes and the bubbles. Accuracy rates for these fields were calculated by manually reviewing the 122 records, consisting of 10,600 digits and 8,000 bubbles, and comparing Scan's raw predicted value against the ground truth value of the image snippet. The recognition accuracy for the bubbles was found to be much higher than that of the number boxes. Including blank fields (fields that users did not fill in), the Scan app classified 99.2% bubbles and 79.4% digits correctly (Table 1).

Type of Field	All Fields	Blank Fields Excluded
Number	79.4%	77.6%
Bubble	99.2%	97.2%

Table 1: Measurement of Accuracy Rates for Recognition of Bubbles and Number Boxes

The misclassified number fields were further investigated by mapping accuracy measures of the number boxes on the TB form to specific fields on the physical paper. It was found that fields located on the right side of the page, the side farthest from the register book's binding suffered from lower accuracy measures than the number boxes which were located closer to the center or left hand side (Figure 4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

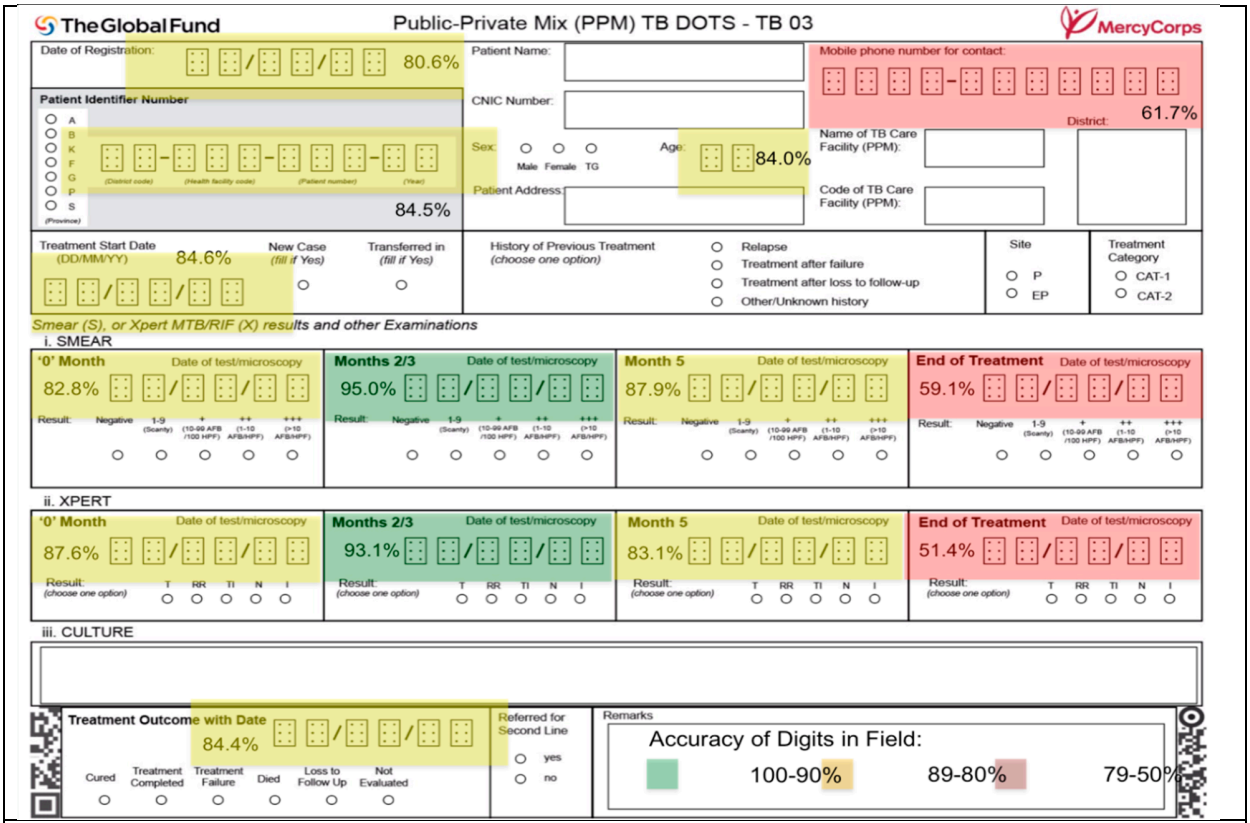


Figure 4: Mapping of the accuracy measures of the number boxes

3.1.2. *User-level quality measure:* The ODK Scan workflow includes viewing, editing and verifying data before it is synced with the server. Field workers were therefore given the complete responsibility for the quality assessment and improvement. However, in the existing manual data entry process, this responsibility is shared between a field worker and a data management officer (DMO). The DMO is responsible for data entry and quality at the next level of management. When synced records were checked and compared with the quality of manually entered records, it was found that for the manual data entry process a data accuracy rate of 94% was achieved, while the ODK Scan process only reached 86%. The accuracy of the data in the manual data entry process was calculated by comparing the scanned version of the hard-copies of TB03 forms to the excel spreadsheets they were translated into. Similarly, in the ODK scan process, data accuracy was calculated by comparing an exported excel sheet from the ODK aggregate server to the scanned version of the hard copies of TB03 forms.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

3.2. Impacting Data Management Processes

In the TB control program, the routine data management process begins with the collection of data from all enrolled General Practitioners (GPs). Once the consolidated paper-based report is prepared it is transferred to the next level of management (Sub-recipient office), which then passes through multiple desk reviews and is finally entered into an excel spreadsheet. The digitization of paper-based records takes a significant amount of time before it is shared with the next level of management (Principle recipient office) for finalization and progress review meetings (Figure 5). Time spent in form-filling, data collection, transfer, aggregation and verification were measured by the direct observation. At baseline, these measures were recorded by observing the routine data management process. However, at the endline the same measures were recorded by observing the ODK Scan workflow.

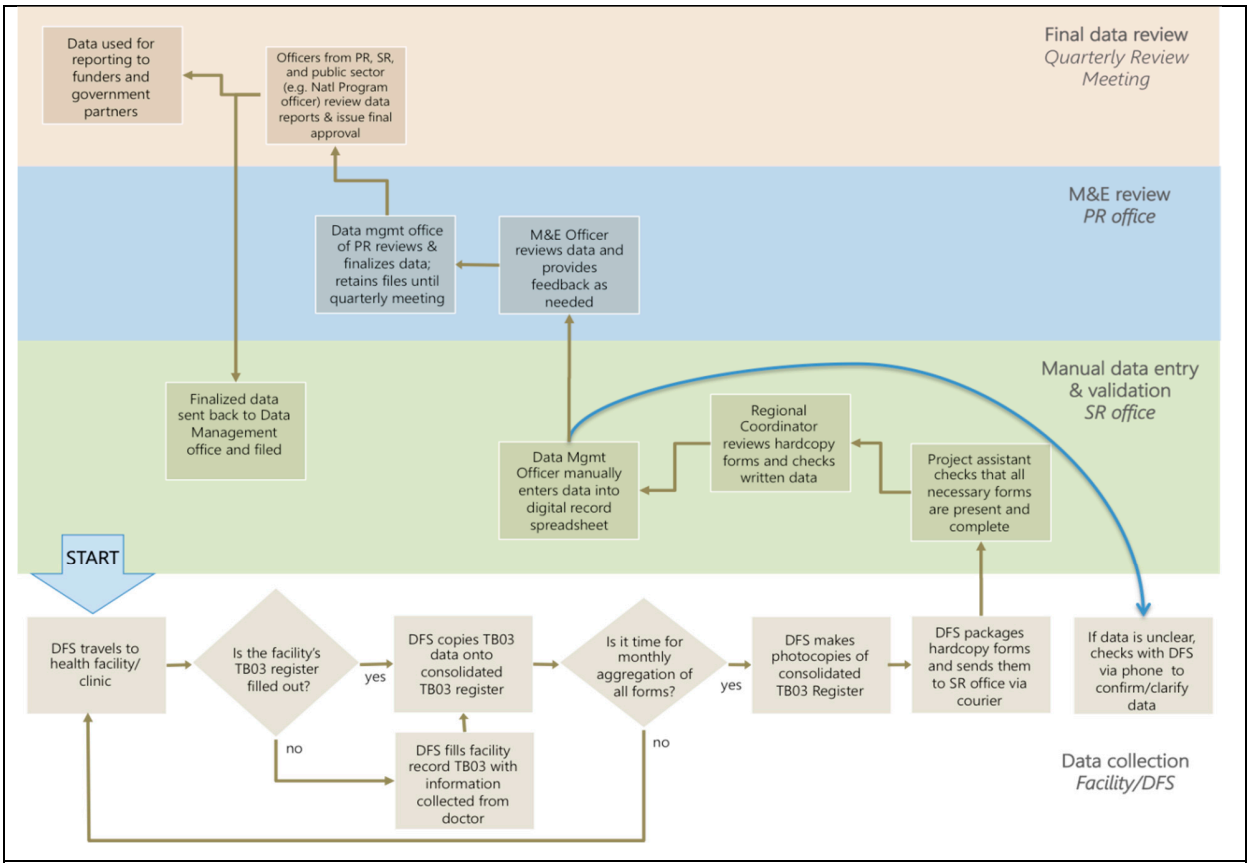


Figure 5: Routine Data Management Process in TB Control Program of Pakistan

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

Filling in the ODK scan-compatible TB form took longer than the old form. Extra attention was required for filling in bubbles and writing structured numbers on the guiding dots correctly. Each field worker (n=4) was observed while filling in three Scan-compatible forms. The average time taken to fill 12 Scan-compatible forms was 02 min 57 sec (max, 03:04 sec; min, 01:48 sec). However, on average it took 01 min and 27 secs for filling in one patient record using the existing TB03 register. The next aspect of the data management, data collection, was not impacted with the ODK Scan workflow because the process remained the same: visit all of the enrolled GPs of a district. Finally, with the routine data management process once data collection was completed, it took 1-2 days to transfer data via local courier service, in the form of hard-copies. However, with the ODK Scan process the same transfer of data only lasted a few minutes. It only required the user to get to an area of internet connectivity and send the data to the ODK Aggregate Server.

In the routine data management process, aggregation of records required manual data entry into a spreadsheet, and digitization of one record took 2 minutes. This digitization process was reduced to 10 seconds in the ODK Scan workflow. Moreover, the data verification process was greatly altered, as in the ODK Scan process the field worker took the sole responsibility of viewing and verifying data, while in the routine data management process this responsibility was shared between field worker and DMO. In terms of time, verification of paper based records took 20 seconds per record by the DMO, while in the ODK Scan process it took 3 minutes per record by the field worker. This verification process in ODK Scan workflow included checking of the data for completeness, consistency, and bubble and number recognition accuracy. ODK Scan process provided a significant time reduction for the data transfer and aggregation steps; however, time required for the verification and form filling were increased as compared to corresponding processes of the routine data management (Table 2, Figure 6).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

Data Management Processes	Routine Data Management Process	ODK Scan Process	Time Consumed in redesigned process
Form Filling (per form)	01 min 27 sec	02 min 57 sec	↑
Data Collection (complete)	3-4 days	3-4 days	=
Transfer (complete)	1-2 days	3 mins	↓
Aggregation (per record)	02 min	10 sec	↓
Verification (per record)	20 sec	3 min	↑

Table 2: Impact of ODK Scan on various processes of Data Management

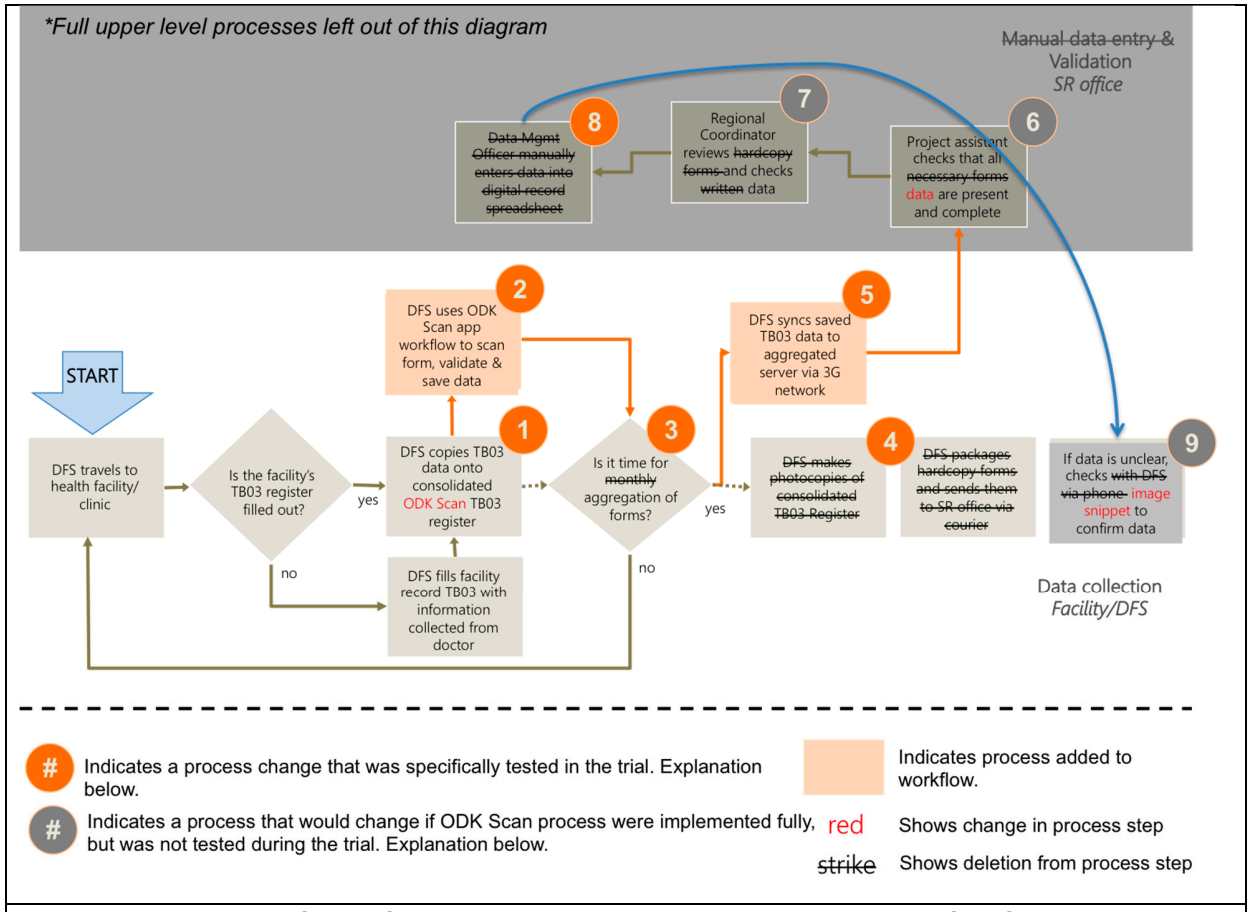


Figure 6: Redesigned Data Management Process in ODK Scan paper-to-digital system

3.3. Feedback on App Usability and Workflow Modifications

The field workers learned quickly how to use the smartphones and swiftly became comfortable with the application's workflow: most of them were already using smartphones in their personal lives. Nevertheless, the implementation of ODK Scan was

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

guided by the principles of *minimal disturbance in current workflow* and *generating value for users* to improve the adoption chances. Face-to-face interviews were conducted independently with each of the field workers at the end of the intervention period. These interviews were supported by a semi-structured questionnaire, which included hints to guide the interviewer to ensure the completeness of the feedback. The interviews were conducted by a research team member in the language of the enrolled field workers (Urdu). All interviews were digitally recorded, transcribed in Urdu and then translated into English. The interview recordings and the corresponding translations were then checked by another team member for consistency and completeness. Data is segregated based on the various steps of the app workflow for the ease of analysis and presentation.

3.3.1. Form Filling:

The ODK Scan process begins with filling in the ODK Scan-compatible form. This includes marking the correct bubbles for the bubble fields, linking guide dots for the structured number fields and writing in the text fields. Field test results showed that this step took more time as compared to the corresponding step of the routine data management process. Moreover, in terms of learning to fill in the new form, all of the field workers adapted to the form well and found it easy to locate the desired field. Nonetheless, almost all field workers were unhappy about the time-consuming aspect of the form filling for the ODK Scan process (Table 3), while at the same time they acknowledged the benefit of using number boxes and bubbles. One DFS said,

"This experience was very good. This is excellent. Especially dots and bubbles have eliminated any chance of the data management officer contacting us for clarification or to address any readability issue."

Interestingly, one DFS highlighted a different aspect of this, saying

"[I do not like] filling the dots on the form as they are time consuming. Also, filling the dots embarrasses me whenever I do that in the presence of others. It seems like I am a kid and practicing handwriting."

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

Field Workers (FW)	Learning to fill new form					Locating desired field					Ease of writing numbers					Ease of filling bubbles				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
FW1																				
FW2																				
FW3																				
FW4																				

Table 3: Self-administered questionnaire items
Note: Whereas, 1 corresponds to very easy and 5 corresponds to difficult.

3.3.2. Scanning form:

Scanning a Scan-compatible form requires the user to take a picture of the paper form using smartphone's built-in camera. This lets the application process the form in the background and detects machine-readable, hand-marked data fields. Once the processing of image is finished, a notification appears in the top toolbar. Field workers understood the importance of taking a good quality picture, and one described it as,

"A good picture is one which is not blurred, and properly focuses the form and covers all the edges of forms. No corner should be cut."

Field workers considered this step important and easy to learn. Knowing this step to be important for having accurate and better classification results, they preferred to take image at a relatively calmer place. One field work said,

"Because this is delicate task and to make its editing error-free, I conduct this exercise at home."

3.3.3. Verifying Data:

When an image is processed, the record is stored into the ODK 2.0 database. The user then launces the ODK Survey app by pressing the "Transcribe" button to verify the data. This allows the field worker to move field-by-field checking for any necessary edits. Most of the field workers considered the data verification process time-consuming, which may have contributed to the relatively lower data quality in comparison to the manual data management process. One field worker explained this process as,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

"I performed verification by pressing next option to move from one field to another. I checked all data points of one form and verify each field to confirm that data written on form is properly scanned and all mistakes are rectified. The information is easily viewable."

Moreover, one field worker explained the positive aspect of this verification step and said,

"Using this app feature, I feel more confident about the quality of data as I view and verify each data fields one-by-one. It feels like I am taking complete charge of the data quality"

3.3.4. Sending Data:

Sending data to the next management level was significantly easier with the ODK Scan process. In the Scan process, this simply involves getting to an area of the internet connectivity and sending data to the server by syncing it. Conversely, in the routine data management process, it is very time and energy consuming: copies of the TB register must be sent to the office through a reliable courier service. One of the field workers explained this process as,

"Using sync feature is easier because I do editing mostly at home then I reach an area of connectivity and sync forms. While, for sending paper forms via courier, first I have to carry big sized TB03 register to photocopier to make the copies of TB03, then I have to visit courier office to send the parcel. It will not reach the same day, at least it takes one day and then I have to call office to confirm if they have received the parcel/documents."

Another field worker explained the process of data sharing that he adopted over the intervention period,

"Once editing is finalized, then we connect device to Internet – mostly to WiFi – and open Core Services app for syncing. First I verify my email address then syncing is confirmed. Mostly I used to sync the data from public partner's office, and if I am far away then I go to local Internet café to sync data. To ensure my connectivity with Internet I check and confirm WiFi icon on top of the screen."

However, dependence on WiFi internet connection for syncing the data was challenging for two of the field workers, and one of them said,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

"It would be very great and easy for us if ODK can sync forms via 3G Internet connectivity. As it is, I scan and edit the forms at home and for syncing forms I have to travel to place where I can have access to WiFi which takes too much time."

3.3.5. Report Generation:

Regular sputum smear microscopy is an important medical aspect in ensuring adherence to the treatment protocol and declaring the treatment outcome of a TB patient. It is extremely difficult to review paper-based records and prepare a list of patients whose sputum smear microscopy is due in the next month. This time-consuming exercise is iterative and considered as a fundamental responsibility of the field workers for tracking treatment progress of the registered patients. With the ODK Scan process, generating this actionable information became easier by running a customized follow up visit report on the smartphone in an offline mode (Figure 7). This enabled field workers to keep a constant check on the treatment progress of the patients registered in his area of control. Moreover, all field workers talked very positively about the benefits of this report. One of them said,

"My existing practice is to review TB03 in the beginning of month to identify patients with follow up appointments. This task takes lot of time and mental energies as it is difficult to develop such list. With that view point, if I see this application, this is an excellent idea that I open app and all follow up visits will be appeared."

Others said;

"This is the best feature of this application, as now we know in advance when we have to call the patient for follow-up"

"Previously my practice was to visit all the GPs to enlist patients who have follow up visit due or overdue. Now with this report I do not have to visit the GPs to know patient's follow up visit date."

"I would say that, follow up visit report has revolutionized our work. Now, we will never miss out any patient and upon click we view all needed information required for follow up."

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

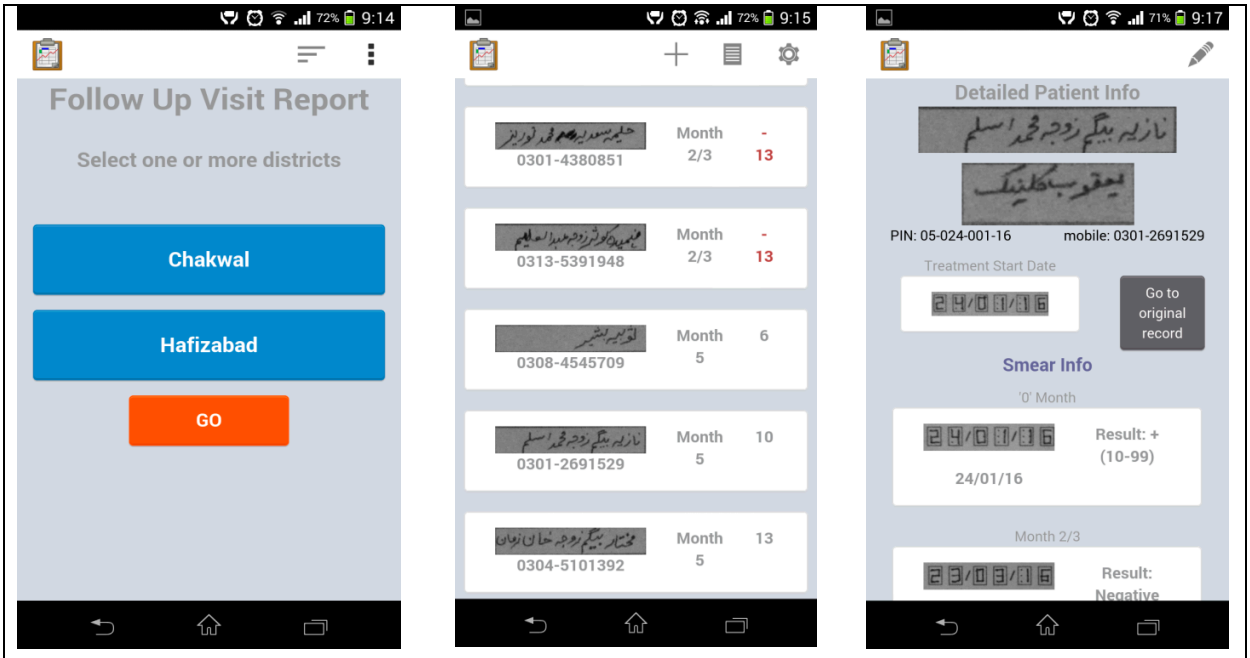


Figure 7: Process of Follow up Visit Report (left to right)

3.3.6. Suggestions for Improvement

After analyzing the content of interviews, it was evident that the field workers completely understood the app workflow and appreciated the benefits of using ODK Scan in improving their work efficiency and performance. Moreover, field workers were eager to continue application use. In addition to simple aggregation reports, field workers wanted to enhance the features of the application. For instance, they suggested including a bookmark feature to tag any patient record, if he/she is not accessible telephonically at the first attempt. With the number of the records available in a table on the user's device, the most popular requirement was patient look-up feature. As TB treatment involves multiple visits of the same patient for medical checkup, it is becoming increasingly difficult for field workers to retrieve a relevant patient's record from the ODK Tables default interface upon patient's revisit. Look-up feature will help field worker to avoid re-scanning and re-verifying form and follow up service will be updated in an existing patient's record.

Currently, synchronization involves the movement of data in either direction. It means, new records on the device are sent to the server and at the same time new records available at the server, synced by others, are downloaded onto the user's device. Deleting a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

record, either deliberately or mistakenly, is a potential threat to data loss. Therefore, defining user permission attributes to restrict data access to only the appropriate individuals would improve the Scan workflow. Lastly, one of the field workers suggested that instead of scanning, a direct-to-digital solution seems more feasible.

4. Discussion

This paper compares paper-based and technology-assisted data management processes and tries to extrapolate the value of mHealth solutions for capturing remote health data. In particular, the present study assesses the usability and technological aspects of the ODK Scan supported by the perspective of the field workers in relation to their work settings. However, it is also noted that in the developing country context, a prerequisite to the success of mHealth interventions is the consideration of developers and implementers about the field workers' expectations, needs and work settings [18, 19]. Moreover, a gap between the work practices and use of the new technologies is attributed by the poor implementation design of the technology-based interventions. [20]

In the Public-Private Mix model of the TB control program of Pakistan, despite available direct-to-digital data collection solutions, the ODK Scan paper-to-digital system was chosen because of the current healthcare system requirement. Largely, the healthcare system relies on the paper-based records for recording, reporting, verification, validation and other purposes. This setting informed the selection of ODK Scan paper-to-digital system. In other mHealth interventions where Android-based smartphones are used primarily for data collection, the ability of smartphones to transfer data in real-time from remote areas to centrally located offices is achieved by exploiting the internet connectivity [15, 21]. Similarly, in this field test, internet connectivity benefited field workers in transferring the data and labor involved in sharing the data in the form of hard copies was avoided. However, the ability of smartphones to transfer data by using 3G internet connectivity instead of WiFi would be highly beneficial. More than just digitization of patient's records, technology also supports report generation, case management and patient follow-up [22].

Data quality achieved through the ODK Scan process is not impressive if compared with its counter process. This drop in quality can be explained through the change in roles as

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

previously field workers were not solely responsible for ensuring data quality. In terms of time, the routine process of data review takes much longer than the new process of the data review and verification. Therefore, an alternative approach to improve the data quality would be to carry out supportive supervision, so that acceptable level of quality of the final set of data could be achieved [23]. Additionally, effective data verification approaches require determination of the data quality characteristics and the application of quality standards [24]. However, the benefits of the mHealth technologies are improved with increased validity of the digital data [25].

In the ODK Scan pilot test, unrestricted mobile and internet usage was authorized in good faith to generate ownership and responsibility among the field workers. It is also reported that free use policy has generated a sense of ownership and empowerment among the health workers by recognizing the significance and convenience of the smartphone usage in their practices [26].

The most challenging aspect of the use of ODK Scan was filling in Scan-compatible forms that included bubbles, number boxes and text fields. Additionally, this field test recorded significant time consumption on data verification and editing processes. On one hand, viewing and verifying all processed data values is beneficial for ensuring data quality, but on the other hand this requires the field worker to spend more time finalizing the scanned record after correcting Scan's errors. Therefore, more errors in the machine-readable data fields resulted in the frustration of field workers. Furthermore, lower recognition accuracy of the number boxes indicates a need for additional application engineering work to make ODK Scan more practical and workable solution. However, it is also an essential consideration for an implementing organization to include as many machine-readable data fields as required [27]. Additionally, the opinion of the field workers must be included during the form design process, so that the workload is anticipated and subsequently accepted by them and only important data fields are chosen for the ODK scan paper-to-digital system. Complexity of the mHealth interventions requires change management considerations, including changes in the behavior of healthcare providers and professionals, and systems and processes in delivering care [28, 29].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

Considering the short duration of the ODK Scan pilot phase, its impact on the TB case management aspects and treatment outcomes could not be measured. After feedback informed technology enhancements, its expansion plan in the TB control program of Pakistan will be prepared. In current mHealth domain, a key challenge is to make strong evidence available for advocating the proliferation of mHealth use in improving the health outcomes. Therefore in future studies, we aim to measure ODK Scan's effectiveness in ensuring fidelity to treatment protocols and improving treatment outcomes.

5. Conclusion

Although it is too early to demonstrate a direct link between digitization of data collection and improvement in health outcomes, mHealth technologies enable public health professionals and program managers to make data available for use in an easy and timely fashion. Prompt generation of actionable information from the digital records is major advantage of the technology. However, it is important to integrate application use within the routine workflow, providing minimal disturbance into existing practices, and continuous feedback are important factors for technology acceptance among the field workers. Efforts made to enhance technology features and improving users' experiences are important considerations for the successful implementation of mHealth interventions in the developing countries. However, a gradual adoption is necessitated by human resource factors and the pace of technological advancements. Hence, it is equally important to devise a national-level policy, or at least organizational policy, for the integration of the mHealth technologies into public health programs to maximize the opportunities to reach needy communities.

Acknowledgment

This paper was put together by the Monitoring, Evaluation and Learning Unit team of the Mercy Corps (MC) Pakistan. On behalf of the MC Pakistan, authors would like to thank the excellent support given by the technical teams of Village Reach and the University of Washington for collaborative mHealth work in the Public-Private Mix Model of the TB control program of Pakistan. Authors would also like to thank the participating field workers, Mudassar , Naveed, Shujja and Zuhaib, for their patience, support and feedback.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

Conflict of Interest

R.A. and J.B. are involved in the development and management of ODK, an open source platform for data collection.

Author Contributions

S.M.A prepared the final draft of the manuscript and contributed in development of the analysis plan. R.P. developed overall design of the study, conducted analysis of the quantitative data and developed graphical presentations. She also developed the data collection tool. N.B. prepared the analysis plan, implementation design and analyzed qualitative data. M.I. conducted interviews, translated them into English and managed field activities. M.I. and J.A. conducted literature search. J.A. managed the technical aspects of the application and also helped M.I. in translating interviews. Technical and development work was led by J.B. R.A. oversaw the development of technology and checked the final manuscript for consistency and scientific rigor. All authors read and approved the final manuscript.

Reference List

[1] World Health Organization (WHO). Global Tuberculosis Report- 2015.*WHO* 2015. Available online: http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf (Accessed on 27th June, 2016)

[2] Sibley, L.; Dennis, M.; Sarfas, C.; White, A.; Clark, S; Gleeson, F.; Mclntyre, A.; Rayner, E.; Pearson, G.; Williams, A.; Marsh, P.; Sharp, S. Route of delivery to the airway influences the distribution of pulmonary disease but not the outcome of Mycobacterium tuberculosis infection in rhesus macaque. *Tuberculosis Journal* 2016, Volume 96, Pages 141- 149. Available Online: http://www.sciencedirect.com/science/article/pii/S1472979215302110 (Accessed on 27th June, 2016)

[3] World Health Organization (WHO). Available Online: http://www.emro.who.int/pak/programmes/stop-tuberculosis.html (Accessed on 27th June, 2016)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

[4] World Health Organization (WHO). Understanding and Using Tuberculosis Data. *WHO* 2014. Available Online: http://apps.who.int/iris/bitstream/10665/129942/1/9789241548786_eng.pdf?ua=1&ua=1 (Accessed on 27th June, 2016)

[5] World Health Organization (WHO). A guide to Monitoring and Evaluation for collaborative TB/HIV activities. *WHO* 2015. Available Online: http://apps.who.int/iris/bitstream/10665/150627/1/9789241508278_eng.pdf?ua=1 (Accessed on 27th June, 2016)

[6] World Health Organization (WHO). Digital Health for End TB Strategy: An Agenda for Action. *WHO* 2015. Available Online: http://www.who.int/tb/areas-of-work/digital-health/Digital_health_EndTBstrategy.pdf (Accessed on 27th June, 2016)

[7] Division of Tuberculosis Elimination. "Treatment of Drug-Susceptible Tuberculosis Disease in Persons Not Infected with HIV. *Centers for Disease Control and Prevention (CDC)*, June 2012. Available Online: http://www.cdc.gov/tb/publications/factsheets/treatment/treatmenthivnegative.htm (Accessed on 27th June, 2016)

[8] Timimi, H.; Falzon, D.; Glaziou, P.; Sismanidis, C.; Floyd, K. WHO guidance on electronic systems to manage data for tuberculosis care and control. *Journal of the American Medical Informatics Association (JAMIA)* 2012, Volume 19. Available Online: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3534454/ (Accessed on 27th June, 2016)

[9] Dell, N.; Perrier, T.; Kumar, N.; Lee, M.; Powers, R,; Borriello G. Paper-digital workflows in global development organizations. CSCW'15 Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing, New York, USA, 2015, p 1659-1669. Available Online: http://dl.acm.org/citation.cfm?id=2675145 (Accessed on 27th June, 2016)

[10] Agarwal, S.; Rosenblum, L.; Goldschmidt, T.; Carras, M.; Goel, N.; Labrique, A.B. Mobile technology in support of frontline health workers: a comprehensive overview of the landscape, knowledge gaps and future directions. John Hopkins University Global mHealth Initiative 2016. Available Online:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

https://dl.dropboxusercontent.com/u/5243748/mFHW%20Landscape_2016%20Final.pdf (Accessed on 27th June, 2016)

[11] Sa, J.H.G.; Rebelo, M.S.; Brentani, A.; Grisi, S.J.F.E.; Iwaya, L.H.; Simplicio, M.A.; Carvalho, T.C.M.B.; Gutierrez M.A. Georeferenced and secure mobile health system for large scale data collection in primary care. *International Journal of Medical Informatics* 2016. Available online: http://www.ijmijournal.com/article/S1386-5056(16)30142-3/fulltext (Accessed on 27th June, 2016)

[12] Hamine, S.; Gerth-Guyette, E.; Faulx, D.; Green, B.B.; Ginsburg, A.S. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: A systematic review. *Journal of Medical Internet Research* 2015, 17(2):e52. Available online: http://www.jmir.org/2015/2/e52/ (Accessed on 27th June, 2016)

[13] van Velthoven, M.H.; Car, J.; Zhang, Y.; Marušić, A. mHealth series: New ideas for mHealth data collection implementation in low– and middle–income countries. Editorial. *Journal of Global Health*. 2013, 3(2), pp 1-3. Available online: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3868820/ (Accessed on 27th June, 2016)

[14] Kaplan, W.A. Can the ubiquitous power of smartphones be used to improve health outcomes in developing countries? *Globalization and Health*. 2006; 2:9. Available online: http://globalizationandhealth.biomedcentral.com/articles/10.1186/1744-8603-2-9 (Accessed on 27th June, 2016)

[15] Mosa, A.S.; Yoo, I.; Sheets, L. A systematic review of healthcare applications for smartphones. *BMC Medical Informatics and Decision Making*. 2012,12:67. Available online: http://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-12-67 (Accessed on 27th June, 2016)

[16] Brunette, W.; Sundt, M.; Dell, N.; Chaudhri, R.; Breit, N.; Borriello, G. Open data kit 2.0: expanding and refining information services for developing regions. HotMobile '13 Proceedings of the 14th Workshop on Mobile Computing Systems and Applications Article No.10.2013. Available online: http://dl.acm.org/citation.cfm?id=2444790 (Accessed on 28th August, 2016)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

[17] Dell, N.; Breit, N.; Wobbrock, J.O.; Borriello, G. Improving form-based data entry with image snippets. *GI' 13 Proceedings of Graphic Interface*. 2013. Pages 157- 164. Available Online: http://dl.acm.org/citation.cfm?id=2532156 (Accessed on 27th June, 2016)

[18] Paul, R.J.; Ezz, I.; Kuljis, J. Health information systems: a patient-user perspective. *Health Systems*. 2012, 1(2):85-95. Available online: http://link.springer.com/article/10.1057%2Fhs.2012.17 (Accessed on 14th July, 2016)

[19] Shachak, A.; Montgomery, C.; Dow, R.; Barnsley, J.; Tu, K.; Jadad, A.R.; Lemieux-Charles, L. End-user support for primary care electronic medical records: a qualitiative case study of users' needs, expectation, and realities. *Health Systems*. 2013,2(3):198-212. Available online: http://link.springer.com/article/10.1057%2Fhs.2013.6 (Accessed on 14th July, 2016)

[20] Ali, S.M.; Giordano, R.; Lakhani, S.; Walker, D.M. A review of randomized controlled trials of medical record powered clinical decision support system to improve quality of diabetes care. *International Journal of Medical Informatics*. 2016, 87:91-100. Available online: http://www.ijmijournal.com/article/S1386-5056(15)30085-X/abstract (Accessed on 14th July, 2016)

[21] Earth Institute. Barriers and gaps affecting mHealth in low and middle income countries: a policy while paper. *Washington DC: mHealth Alliance*. 2010. Available online: http://cgsd.columbia.edu/files/2012/11/mHealthBarriersWhitePaperFINAL.pdf (Accessed on 15th July, 2016)

[22] Dwolatzky, B.; Trengove, E.; Struthers, H.; et al. Linking the global positioning system (GPS) to a personal digital assistant (PDA) to support tuberculosis control in South Africa: a pilot study. *Int J Health Geogr.* 2006;5:34. Available online: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1563457/ (Accessed on 14th July, 2016)

[23] Luk, R.; Bancroft, E. Lesson Learned: ICTs for supply chain management in Low-Resource Settings. *Village Reach and Dimagi*. 2013. Available online: https://www.k4health.org/toolkits/mhealth-planning-guide/lessons-learned-icts-supply-chain-management-low-resource-settings (Accessed on 15th July, 2016)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2016 doi:10.20944/preprints201608.0232.v1

[24] Divorski, S.; Scheirer, M.A.; Improving data quality for performance measures: results from a GAO study of verification and validation. *Evaluation and Program Planning*. 2001, 24:83-94. Available online: http://www.sciencedirect.com/science/article/pii/S0149718900000495 (Accessed on 15th July, 2016)

[25] Poku, A.M.; Katz, A.T. Development of mobile health capabilities for remote data collection in resource-limited settings. 2014, *Johns Hopkins APL technical digest*, 32(4):679-684. Available online: http://www.jhuapl.edu/techdigest/TD/td3204/32_04-Poku.pdf (Accessed on 19th July, 2016)

[26] Little, A.; Medhanyie, A.; Yebyo, H.; Spigt, M.; Dinant, G., Blanco, R. Meeting community health worker needs for maternal health care service delivery using appropriate mobile technologies in Ethiopia. 2013. *PLoS one* 8(1):e77563. Available online: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0077563 (Accessed on 19th July, 2016)

[27] Dell, N.; Crawford, J.; Breit, N.; Chaluco, T.; Coelho, A.; McCord, J.; Borriello, G. Integrating ODK Scan into the community health worker supply chain in Mozambique. ICTD'13 Proceedings of the sixth international conference on information and communication technologies and development. 1:228-237. Available online: http://dl.acm.org/citation.cfm?id=2516611 (Accessed on 19th July, 2016)

[28] The PLOS Medicine Editors (2013). A reality checkpoint for mobile health: three challenges to overcome. *PLoS Med*. 10(2):e1001395. Available online: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001395 (Accessed on 19th July, 2016)

[29] Tomlinson, M.; Rotheram-Borus, M.J.; Swartz, L.; Tsai, A.C. Scaling up mHealth: where is the evidence? *PLoS Med.* 2013. 10(2):e1001382. Available online: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001382 (Accessed on 19th July, 2016)

© 2016 by the authors; licensee *Preprints*, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).