Evaluating the Accuracy of Data Collection on Mobile Phones: A Study of Forms, SMS, and Voice

Somani Patnaik, Emma Brunskill and William Thies

***Abstract*—While mobile phones have found broad application in reporting health, financial, and environmental data, there has been little study of the possible errors incurred during mobile data collection. This paper provides the first (to our knowledge) quantitative evaluation of data entry accuracy on mobile phones in a resource-poor setting. Via a study of 13 users in Gujarat, India, we evaluated three user interfaces: 1) electronic forms, containing numeric fields and multiple-choice menus, 2) SMS, where users enter delimited text messages according to printed cue cards, and 3) voice, where users call an operator and dictate the data in real-time.**

Our results indicate error rates (per datum entered) of 4.2% for electronic forms, 4.5% for SMS, and 0.45% for voice. These results caused us to migrate our own initiative (a tuberculosis treatment program in rural India) from electronic forms to voice, in order to avoid errors on critical health data. While our study has some limitations, including varied backgrounds and training of participants, it suggests that some care is needed in deploying electronic interfaces in resource-poor settings. Further, it raises the possibility of using voice as a low-tech, high-accuracy, and cost-effective interface for mobile data collection.

I. Introduction

Mobile devices have shown great promise for improving the efficiency and effectiveness of data collection in resource-poor environments. Compared to a traditional process that relies on paper-and-pencil forms with subsequent transcription to a computer system, mobile devices offer immediate digitization of collected data at the point of survey. This allows for fast and automated data aggregation. It also improves adherence to complex or context-dependent questionnaires, as the device determines which questions should be answered or skipped.

The benefits of mobile data collection have been demonstrated mostly in the context of personal digital assistants (or PDAs) [31], [10], [8], [2], [32], [12], [24], [9], [16], [4], [3], [15]. Given the recent explosion of mobile phones around the world, there is growing excitement in extending the successes achieved on PDAs to a phone-based platform. While high-end phones provide the same capabilities as PDAs, low-end phones lack features such as high-resolution displays and touch-screen capabilities. To empower the full population of nearly 4 billion mobile phone subscribers [26] with the capabilities of mobile data reporting, it will be important to establish usable interfaces that are portable to inexpensive phones, and there have been a number of recent efforts in this space (see for example [13], [1], [22], [25], [7]).

Manuscript received September 22, 2008.

Somani Patnaik and Emma Brunskill are with the Massachusetts Institute of Technology (email: somanip@mit.edu, emma@csail.mit.edu). During this research, William Thies was affiliated with both the Massachusetts Institute of Technology and Microsoft Research India (email: thies@microsoft.com).

In migrating mobile data collection from PDAs to cell phones, a critical issue is ensuring the accuracy of data entry. In the context of healthcare, an errant entry may prevent life-saving treatments from reaching patients, or may cause the prescription of unnecessary treatment that is costly and dangerous. In financial applications, entry errors may jeopardize the economic standing of communities that are already very poor. Due to the importance of this issue, several researchers have studied the error rates incurred as PDAs are deployed in developing regions. As detailed later (in Table II), the error rates are generally less than 2% (i.e., 2 errors per 100 entries) in programs where users received at least an hour of training [12], [24], [4]. However, in the context of mobile phones, studies of data accuracy are distinctly lacking. The closest work is by Parikh et al., where a hybrid system of paper forms and camera-equipped mobile phones has demonstrated error rates of less than 1% [28]. For standalone data collection on low-end phones, we are unaware of any previous study with a rigorous evaluation of data accuracy. This research opportunity is highlighted in Table I.

In this paper, we provide a quantitative evaluation of data entry accuracy using low-cost mobile phones in a resource-constrained environment. We evaluate three practical user interfaces for entering data on a mobile phone: electronic forms, SMS, and voice. Electronic forms consist of numeric fields and multiple-choice menus, and can be implemented in Java or a native phone platform. The SMS interface requires users to send a structured SMS messages to a server, with logical fields separated by delimiters in the message. The voice interface represents a normal telephone call, with a live human operator that enters the data into a centralized spreadsheet.

We evaluated these interfaces in a study of 13 health workers and paramedical staff over a month-long period in Gujarat, India. Each participant was trained and evaluated on all of the interfaces. We focus on the collection of health data relevant to tuberculosis (TB), as we anticipate deploying an electronic system in a real TB treatment program. The data in this paper represent only simulated patient interactions.

Our results indicate an error rate of 4.2% for electonic forms, 4.5% for SMS, and 0.45% for voice. These represent the fraction of questions that were answered incorrectly; as each patient interaction consisted of eleven questions, the probability of error somewhere in a patient report is much higher. For both electronic forms and SMS, 10 out of 26 reports (38%) contained an error; for voice, only 1 out of 20 reports (5%) contained an error (which was due to operator transcription). As detailed in Section VI, error rates

	PDAs	Cell Phones
Published error rates	Malaria monitoring in Gambia [12] Clinical study in Gabon [24] Tuberculosis records in Peru [4] Sexual behavior surveys in Peru [3]	*None?*
Other programs	SATELLIFE [15] DataDyne EpiSurveyor [31] EpiHandy [10] Infant health in Tanzania [32] e-IMCI project in Tanzania [8] Respiratory health in Kenya [9] Tobacco survey in India [16] Ca:sh project in India [2]	Cell-Life in South Africa [13] Jiva TeleDoc in India [1, p.42] Pesinet in Mali [22] Malaria monitoring in Kenya [25] Voxiva Cell-PREVEN in Peru [7]

TABLE I
PREVIOUS WORK IN EVALUATING THE ACCURACY OF MOBILE DATA COLLECTION IN THE DEVELOPING WORLD.

are distinctly higher for health workers than for hospital staff, though this difference may also be influenced by variations in our training environment.

We were surprised and alarmed by these results. In our own treatment program, our original intent was to utilize electronic forms. However, we consider it to be an unacceptable risk that 38% of submitted forms – containing critical health information – may contain errors. For this reason, we have overhauled our plans and will implement a treatment program using voice rather than forms or SMS. While the cost of a live operator may be prohibitive in many countries, in India it proves to be very cost-effective. The increased cost of a human operator is more than compensated by the decreased cost of voice-only handsets, voice-only cellular plans, decreased training time, and decreased literacy requirements for health workers. We offer a more detailed analysis in Section VII.

While the results of this study have changed our own approach to implementing mobile data collection, we caution the reader in extending the results of the study beyond its original context. In particular, we are focused on the scenario in which users have limited cell phone familiarity and there is limited time to perform training. If either of these variables changes, it may be possible to implement high-accuracy mobile data collection with electronic forms or SMS. Also, while the error rates that we report on mobile phones are 3-8x higher than those previously reported for PDAs, our data are unable to distinguish whether this difference is due to the devices, or due to other aspects of the study demographics, training, and evaluation. A future study could address this question directly by evaluating both phones and PDAs in the same context.

Despite these limitations, our study is the first (to our knowledge) that evaluates data entry accuracy on mobile phones. Based on our results, we submit only that electronic forms and SMS may need further validation before gaining widespread deployment in accuracy-critical applications, and that voice may deserve more attention as a high-accuracy and low-cost means of data collection.

The rest of this paper is organized as follows. We start by reviewing related work on mobile data collection (Section II). Then we consider the tradeoffs between electronic forms, SMS, and voice (Section III) and detail our implementation of each interface (Section IV). We describe the setup of our user study (Section V) and the results obtained (Section VI), and we discuss the implications (Section VII). We conclude in Section VIII.

II. Related Work

As summarized in Table I, there have been several initiatives to apply PDAs and cell phones for mobile data collection in the developing world. While a fraction of the studies on PDAs includes an experimental analysis of the error rate incurred, we are unaware of any study which systematically measures the accuracy of data entry on a cell phone. This is the principal novelty of our work.

Lane et al. provides a review of nine randomized controlled trials that compare the effectiveness of PDAs and paper forms for data collection [21]. Six of the trials reported entry accuracy, with varying results: two studies found PDAs to be more accurate than paper [20], [29], three studies found the accuracy to be similar with both methods [17], [23], [36], and one study found that paper was more accurate [35]. None of the trials were in the context of the developing world (they took place in North America and Europe).

Previous studies of PDA entry accuracy in the developing world are summarized in Table II. In cases where workers received at least an hour of training, error rates are under 2% (i.e., 2 errors per 100 questions). As early as 1991, Forster et. al evaluated the use of PDAs for a malaria morbidity study in the Gambia [12]. Employing secondary-educated workers who received five days of training, they report error rates between 0.1-0.6% and argue that the PDAs offer improved accuracy and efficiency over paper forms. Missinou et al. employed PDAs in a clinical study in Gabon, employing four clinicians who had no prior PDA experience and received 8 hours of training [24]. They report a 1.7% rate of discrepancy between PDAs and paper forms, and note that clinicians preferred the PDAs. Blaya et al. found that error rates improved from 1.3% (with paper forms) to 0.37% (with PDAs) in reporting tuberculosis bacteriology data in Peru[1] [4]. The authors also argue that PDAs are cost-effective [5].

[1]Blaya et al. reports errors per form, rather than errors per entry [4]. Via personal communication with the author, we determined that there were an average of 7.5 entries per form, yielding the error rates quoted here.

Application	Location	PDA	Education Level	Training	Error Rate
Malaria morbidity [12]	Gambia	Psion Organizer II XP	Secondary	5 days	0.1%-0.6%
Clinical study [24]	Gabon	Palm m500	3 M.D.s, one clinical officer	8 hours	1.7%
Bacteriology data [4]	Peru	Palm Zire	Post-secondary (2-3 years)	16 hours	0.37%
Sexual behavior [3]	Peru	Palm Zire	Secondary or less	2-3 mins	14%

TABLE II
ERROR RATES MEASURED BY PREVIOUS RESEARCHERS IN APPLYING PDAS FOR MOBILE DATA COLLECTION IN THE DEVELOPING WORLD.

Higher error rates have been reported in the case of self-administered surveys, when limited training is possible. Bernabe-Ortiz et al. evaluate the use of PDAs for surveys of sexual behavior in Peru [3]. To protect patient privacy, the PDAs were intended for use by actual subjects, rather than by health workers. As only some subjects had finished secondary education, and subjects received only 2-3 minutes of training, the authors observed a 14% discrepancy between electronic and paper forms. However, the error rate was substantially lower for subjects who had finished secondary schooling.

Additional programs have applied PDAs for data collection in the developing world, but have not provided a rigorous analysis of entry accuracy. SATELLIFE uses PDAs for disseminating and collecting medical information in numerous countries [15]. There are anecdotal reports that the PDAs improved data quality [19], and the benefits of decreased error rates were estimated on a five-point scale [6]. Users of the system have also rated its usability [11]. However, we are unaware of a quantitative assessment of the error rates incurred. DataDyne EpiSurveyor [31] has been widely deployed for data collection in Sub-Saharan Africa; while it has been argued that the system is more accurate than paper forms [30], we are unaware of a controlled study. EpiHandy also provides tools for deploying electronic forms on PDAs and has been deployed in South Africa, Uganda, and elsewhere [10]. PDAs have also found application for gathering infant mortality data in Tanzania [32], for pediatric care (as part of the e-IMCI project) in Tanzania [8], for assessing respiratory health in Kenya [9], for surveying tobacco use in India [16] and for maternal and child health (as part of the Ca:sh project) in India [2]. These studies lack formal evaluations of entry accuracy.

Cell phones have also found broad application for mobile data collection in the developing world. Cell-Life employs electronic forms on mobile phones to improve TB and HIV treatment in South Africa [13], [33]. Electronic forms are also used by Jiva TeleDoc for improving rural healthcare in India [1, p.42], and by Pesinet for monitoring infant health in Mali [22]. Mobile phones with forms are also being used to monitor malaria in Kenya [25]; while PDAs were also piloted, the authors note that phones are more intuitive due to worker familiarity. Voxiva's Cell-PREVEN uses interactive voice response and voice recording to monitor adverse events amongst sex workers in Peru [7]. We are unaware of any quantitative evaluation of entry accuracy in these projects.

To avoid the complexities of navigating electronic forms, the CAM framework offers a hybrid system in which paper forms are used for organization while phones are used for data entry [27]. Each field on the paper form is annotated with a barcode, which is recognized by a camera on the phone prior to data entry. Users that lacked prior camera or computer experience were trained to a level of comfort within 5 to 15 minutes. A separate study measures error rates of 1% or below using the CAM system [28]. This represents an interesting and useful design point, especially in cases where paper forms are already ingrained into the workflow. We focus on solutions that are independent of any paper workflow, and which do not necessarily require a camera-phone (while Java-phones often have cameras, our SMS and voice solutions are suitable to the most inexpensive phones).

While electronic forms have been widely deployed, there are fewer solutions that rely on user-constructed SMS messages for mobile data collection. One example is a system from Dimagi, Inc. which monitors water treatment plants in India [34]. We are unaware of other systems which rely on a cue card (as we do in our evaluation) for submitting a structured SMS message to a server.

Others have considered broader issues in the contextual design of user interfaces for data collection in the developing world. Examples include interface design for Auxiliary Nurse Midwives in India [14] and a methodological framework for evaluating health devices [18]. Our focus is on assessing the entry accuracy for a range of standard interfaces.

III. USER INTERFACES

Three of the central modes on a cell phone that can be used to perform data collection are voice, SMS and an electronic forms application. Data collection performed by voice can be further split into systems that link the data collector with a live operator, those that connect to an automated interactive voice response system, and those that allow the user to record a message. We focus our discussion around live voice operators, SMS and electronic form based systems, and examine some of the strengths and weaknesses of these various approaches. We use SMS to refer to data collection systems that involve information entered by a structured text message: in particular we assume that the information is entered by following a small cue sheet with a flowchart that directs the collector how to enter the data. To our knowledge, using cue cards to guide data entry by text message has not been done previously. In contrast, electronic forms (particularly on personal digital assistants) have been widely used. In this paper, we use the term "electronic forms" to denote any external application that can be placed on a phone, and that automatically guides the user how to enter data, through the use of text, menus or other tools. In a voice operator interface, the user simply calls a live operator, who asks the user a series of questions to elicit the information needed. Figure 1 illustrates each interface as used in our particular experiment.

Electronic Forms Interface

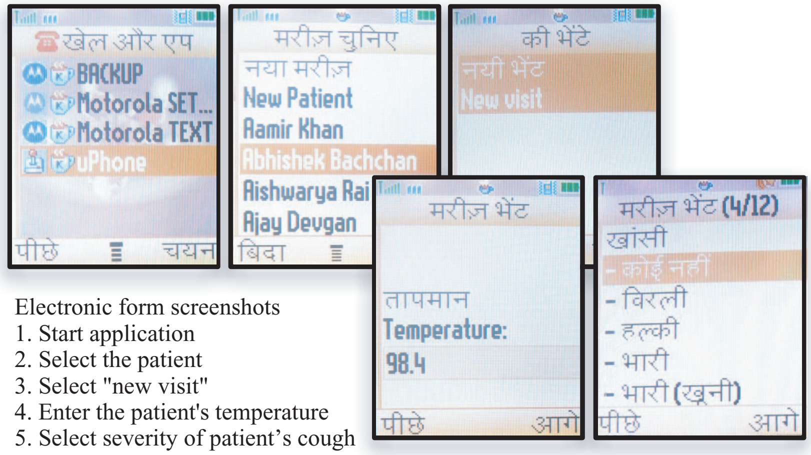
a)

General Strengths
Easy patient identification
Ongoing cost is low (SMS or data plan)
Can store visits when connectivity is poor

General Weaknesses
Requires programmable phones
Requires basic literacy skills
Hard to alter survey questions
Hard to enter in free-form notes
Application can be deleted by user

Our Results: Accuracy & Efficiency
We measured 4.2 errors per 100 entries
The average interaction was 99 seconds

1 2 3



Electronic form screenshots
1. Start application
2. Select the patient
3. Select "new visit"
4. Enter the patient's temperature
5. Select severity of patient's cough

4 5

SMS + Cue Card Interface

b)

General Strengths
Can be used with any phone
Ongoing cost is low (SMS)
Many workers familiar with SMS

General Weaknesses
Requires basic literacy skills
Changing survey requires new cue card
Hard to enter in free-form notes
No confirmed receipt of data delivery
Worker can forget or lose cue card
Quite easy to fake visits (copy old SMS)

Our Results: Accuracy & Efficiency
We measured 4.5 errors per 100 entries
The average interaction was 97 seconds

SMS Cue Card

1. Create a new SMS Message
Press Center Button
Select "Messages"
Select "Create Message"
Select "New Short Message"

2. Switch to Numeric Input Mode
Press Menu Button
Select "Entry Mode"
Select "Numeric"

3. Enter the ID of the Current Patient
Aamir Khan - **Press** 1
Abhishek Bachchan - **Press** 2
Aishwarya Rai - **Press** 3
…

4. Enter a Space
Press *

11. Enter the Patient's Cough
No Cough - **Press** 1
Rare Cough - **Press** 2
Mild Cough - **Press** 3
Heavy Cough - **Press** 4
Severe Cough (with blood) - **Press** 5

21. Check Yourself
Your finished message should be formatted similarly to the following:
10 372 62 68 4 1030007

Voice Interface

c)

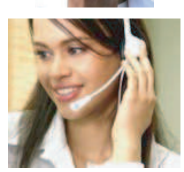
General Strengths
Can be used with any phone
No literacy required of workers
Easy to change survey questions
Easy to add in free-form notes
Hard to fake a visit: operator can ask new questions

General Weaknesses
Ongoing cost of operator salary
Voice plans often higher cost than SMS
Awkward 3-way social interaction

Our Results: Accuracy & efficiency
We measured 0.45 errors per 100 entries
The average interaction was 140 seconds

Patient

Worker



Operator

Sample Voice Interaction

Operator: *Hello. What is your name?*
Worker (to operator): *My name is Lipika. I am calling to record a patient visit.*
Operator: *What patient are you visiting?*
Worker (to patient): *What is your name?*
Patient: *Pavathi (reading from note note sheet)*
Worker (to operator): *Pavathi.*
Operator: *That's Pavathi, right?*
Worker (to Operator): *Yes*
(operator records name)
Operator: *What is her temperature?*
Worker (to patient): *What is your temperature?*
Patient: *97.1 (reading from note sheet)*
Worker (to operator): *97.1 degrees.*
Operator: *97.1 deg.*
(operator records temperature) …

Fig. 1. The three user interfaces evaluated in this paper: a) electronic forms, b) SMS + cue card, and c) voice.

In general, there are a variety of factors that affect the choice of a data collection interface. These may be loosely categorized into operation, effectiveness and cost. Figure 1 supplements the below discussion by summarizing some of the strengths and weaknesses of each interface.

A. Operation

We use "operation" to refer to factors involved with the general infrastructure of the data collection system. Initially there is the investment of time to set up the system, and then train the workers who will be performing data collection on the system. For voice or SMS interfaces, the set up time for workers is minimal: each worker must simply be provided with a phone, if he or she does not currently own one. However, electronic forms require that the application be downloaded onto the phone, which requires either an Internet-enabled phone in an area of good connectivity, or specialized development tools and an external computer.

Training time for each application is an open issue, and is one of the factors we investigate further in this study. Worker education and worker cell phone familiarity are likely to affect how easy it is to set up each user with an interface, and train them how to use it. We expect that a voice interface requires the least amount of education and background to get users equipped to start performing data collection. In particular, a voice interface does not require that its users be literate.

System coverage and reliability are also critical factors to ensure good data collection. Voice calls have priority over SMS, and there is the possibility of lost SMS messages. The delivery mechanism with electronic forms can vary: both GPRS and SMS can be used. GPRS has the advantage that there is an acknowledgment of whether the data was sent; however not all locations have coverage. From the user side, voice appears to be the most reliable and has the most far reaching coverage; however, this also requires that there exists a sufficient number of operators so that users can always reach a person when they call. If this is not always possible then there may be a reliability penalty as users may have to call back later (or wait for the operator to return their call).

In addition to reliability, a good system should enable some degree of flexibility. Despite good initial prototyping, it may sometimes be important to be able to modify the data collection interface, fix an error, improve usability, or add or remove information to be collected. If users have an Internet enabled phone and are always working in areas of high connectivity, then updating an electronic form system is quite feasible. However, if this is not the case, then users must reprogram their phone using the same specialized tools needed for initial set up. SMS is similarly challenging to update since a new cue card must be distributed to direct the user to enter the data. In contrast, voice is trivial to update, as the operator can simply ask a new set of questions.

B. Effectiveness

In any data collection effort, one of the key considerations is the effectiveness of the program at obtaining high quality data. High quality data can perhaps be characterized by two simple criteria: whether or not the data is intentionally faked by the user, and the accuracy of data that is not intentionally faked (which is the focus of this paper).

Intentionally faked data can lead to incorrect conclusions and potentially lead to significant misallocation of resources when interventions are based on false data. There may be an incentive to fake data when users are busy and collecting real data is time consuming, due to the data recording itself or transportation time to reach the source of the data (such as visiting remote patients). Unfortunately in SMS systems it is quite easy to fake data, particularly for cell phone savvy users that can copy and paste prior SMS messages. Faking electronic forms is slightly harder as it requires the user to sequentially fabricate data across an entire form. It requires the most effort for users to fake data while speaking on the phone, as the operator can always ask a new question to try to ascertain if the user is fabricating the data.

Voice also has the benefit that it is easy for users to convey additional information (not included in the original survey), whereas it is more challenging to spell out text using the keypad, particularly in other languages which may or may not be supported on a given phone. Voice is also likely to have fewer operational risks: users may accidentally delete the form application, or forget their SMS cue card, but since an operator can always call a worker directly, the voice system is fairly robust. Voice also makes it easy for users to correct previous visits, by simply calling back the operator. This is also easy to do by modifying and resubmitting a saved electronic form.

However, it is also important to consider the speed of data entry, how much the user likes the interface, and the accuracy of data entry. To our knowledge there are no prior studies comparing the accuracy and speed of data entry using SMS, electronic forms and voice. Since we regard these as some of the most critical factors in choosing an interface, this is a large motivation for our current study.

C. Cost

One of the other important considerations is cost: the most beautiful, user-friendly, accurate interface may still not be practical if the cost overhead is too high for the particular problem. Costs consist of fixed one time costs as well as ongoing marginal costs.

For all three interfaces users must have a cell phone. An electronic form requires a programmable phone (such as a Java-enabled phone or Windows phone) but both SMS and voice applications can be used with any phone. The ongoing cost for an SMS phone depends on the rate per message which is typically quite low. An electronic form can send data using SMS or through a data plan; typically SMS is cheaper depending on the amount of data that is being collected. Voice minutes are frequently more expensive than SMS. But most importantly, voice has the ongoing cost of the salary of the operator, which is an additional overhead not shared by electronic forms or SMS.

IV. User Interface Instantiation

The prior section discussed some of the general factors important to consider when designing and selecting a data collection interface. We now discuss the context for our data collection effort and the interfaces we evaluated.

A. *Domain context*

Soon the authors, along with other collaborators, intend to conduct a trial that examines whether increased information and monitoring can improve heath outcomes and adherence during tuberculosis treatment in Bihar, India. Treatment will be conducted by having tuberculosis patients regularly visit health workers and receive drugs as part of a directly observed therapy (DOT) strategy. During these visits, health workers will collect data about their patients and report this information by mobile phone back to a central office. This information will be aggregated and analyzed to inform doctors and the trial manager about which patients may need to be visited, for example, if a patient is not improving or is experiencing adverse side effects. To support this effort we need a user interface that enables fast and accurate data collection.

The data collected during a patient visit will include both identification and health status information. The worker will enter in information to identify both the worker name (done only once at the start of treatment, in the case of forms and SMS) as well as the patient name. In addition, the health worker will record the patient's current temperature, weight and pulse, as well as the presence or absence of seven symptoms: night sweats, chest pain, loss of appetite, nausea, coughing with blood, yellow eyes and fatigue. These symptoms were chosen based on advice gathered from tuberculosis health experts. The worker will also record whether the patient's current cough is absent, rare, mild, heavy or severe with blood.

The trial intervention is centered around the hypothesis that better, more frequent data collected about patients can improve tuberculosis health outcomes and therefore high quality data collection is critical. However, even if an interface encourages high quality data, it is still essential that such a data collection method also be easy to use and affordable in order for such an intervention to have widespread applicability. Originally we were planing to use electronic forms for data collection. However, since there appeared to be a dearth of literature in evaluating mobile data collection accuracy, we decided to evaluate the accuracy, speed and usability of three mobile phone interfaces. The results of this evaluation influenced our choice of an interface for use in the treatment program.

B. *Electronic forms implementation*

We created a Java application which provides a sequence of electronic forms that guide the worker to request information from the patient. The worker identification number is encoded once into the phone and is included with each recorded visit. The worker has to either enter numeric data or make a selection from a multiple-choice menu to encode symptoms.

The electronic forms underwent several design iterations, including gathering feedback from a 3-day session with 22 health workers in Bihar, India, prior to this study. Based on feedback from the workers in Bihar, we choose to employ hybrid English/Hindi menus for some of the forms, since some medical terms are easier to understand in English, but others are easier to understand in Hindi. We also changed from using multi-select lists (with a checkbox per symptom) to using individual yes/no questions.

Figure 1a shows a series of screenshots of the form interface used for the present study. The Java application can be set up to either relay this information via SMS or GPRS. This distinction is important for cost considerations but does not affect the interface testing considered here.

C. *SMS implementation*

For the SMS interface we designed a cue card that instructs the worker how to record information about the patient into a text message; Figure 1b displays a subset of the cue card used. All information is coded numerically; this is done to reduce the amount of cell phone familiarity necessary, as well as to increase the speed of data entry. Participants enter in data as prompted by the cue card and then send the text message at the end of the interaction. The final part of the cue card as displayed in Figure 1b shows a sample text message.

D. *Voice implementation*

For the voice interface the worker calls a live operator. The operator asks the worker a series of questions about the patient's health, which prompts the worker to ask the patient that question. This means that workers interact simultaneously with an operator and a patient; we are unaware of previous programs that have taken a similar approach. Figure 1c displays a sample interaction. The live operator confirms answers with the worker; this adds to the length of each call but is done to increase accuracy. This can be particularly important when the phone connection is poor or there is background noise.

V. Study Methodology

The user study took place in the Surat and Bharuch districts of the Indian state of Gujarat during July and August of 2008.

A. *Participants*

As detailed in Table III, the study participants consisted of six community health workers and seven hospital paramedical staff. The community health workers were associated with the Dahej public health center; five of the paramedical staff were at the Reliance Tuberculosis hospital; and the remaining two paramedical staff were at the dispensary of the Sardar Vallabhbhai National Institute of Technology. The study participants were recruited through contacts of the first author.

Initially, we had hoped to perform the study entirely with community health workers, as they are often the primary agents of remote data collection (including in our upcoming tuberculosis treatment program). However, this turned out to be infeasible because some community health workers were unable to travel to the Dahej public health center for training and testing, and it was not feasible for us to travel to each worker's home. This prompted us to recruit participants from two other centers. There were also some logistical challenges in performing the studies due to adverse weather conditions and the bomb blasts occurring in July 2008 in the Surat area.

The education level of the health workers ranged from 10 to 12 years, while the education of the hospital staff ranged from 10 years to a B.A. degree. The average age of the study participants was 26.4 years (range 19-35). Seven participants owned a cell phone, four participants had used but did not own a cell phone, and two participants had never used a cell phone previously. Eleven of the participants were native Gujarati speakers and all spoke Hindi.

B. Training

Participants were trained by at least two trainers in small groups of at least two. Initially, examples were presented on a whiteboard and participants were instructed to practice entering in the data on either electronic forms or as an SMS using the cue card. After this stage, a paper with a set of example patients was handed out, and participants were instructed to practice entering in this data. In the final stage, participants were instructed to practice role playing patient–worker interactions with each other.

Participants received variable amounts of training, ranging from 45 minutes to 8 hours, depending on their experience and availability. The longer training sessions were not necessarily more effective, as they were performed in larger groups. While it would have been desirable to achieve more uniform training, this was difficult given the logistics of transportation and worker schedules. Prior to the completion of training, all participants had completed at least two perfect interactions on both electronic forms and SMS, and at least one perfect interaction on the live operator mode.

Throughout the user study, we employed Motorola L6i cell phones for training and testing. This is the cheapest Java-enabled phone from Motorola (the source of our current development tools) that is available in India; see Appendix A-1 for a cost analysis. All interfaces and related tools (cue cards, etc.) were presented in Hindi, and the mobile phones used had dual Hindi menus.

C. Testing

Participants were tested in pairs, alternating who was being tested on data entry, and who was playing the fake patient for that data point. The order of the interfaces was randomized: for a given participant pairing, the order of voice, SMS, and electronic forms was alternated. For the voice interface, the first author acted as the operator and was located outside of the room testing was being conducted in; however, there was always an additional person associated with the experiment inside the room at all times with the participants.

During testing, each participant performed two complete patient–worker interactions (in the role of the worker) for each of the forms and SMS interfaces. For the voice interface, the six community health workers completed only one interaction, while others completed two interactions (we did not anticipate that voice would become a focal point of this study until halfway through our experiments).

The lag time between training and testing was exactly one day for seven of the participants, and ranged between half a day and two days for the remaining participants. All participants received a brief refresher and supervised entry session immediately prior to testing.

VI. Results

The results of the user study are detailed in Table III. We present both the accuracy of data entry, as well as the time needed to interview patients and report the data.

On average, electronic forms and SMS offered comparable error rates of 4.2% and 4.5% per entry, respectively. The voice interface proved to be approximately 10x more accurate, with an error rate of 0.45% per entry. While only one out of thirteen participants performed perfectly on both the forms and SMS interfaces, twelve out of thirteen participants performed perfectly on voice. A Student's two-tailed, unpaired t-test revealed that voice had a significantly lower error rate than electronic forms ($p < 0.01$) and SMS ($p < 0.01$); no significant difference was found between the error rates of electronic forms and SMS ($p = 0.84$).

It is important to note that our results indicate a bimodal distribution of error rates: participants 7-13 performed notably better than participants 1-6. While there are many compounding differences between these participants, including the manner in which we conducted training, we refer to them by their occupation in order to simplify the discussion; participants 1-6 are health workers while participants 7-13 are hospital staff. As summarized in Table III, health workers exhibited an error rate of 7.6% for forms and 6.1% for SMS, while hospital staff exhibited an error rate of 1.3% for forms and 3.2% for SMS. In addition, the only voice error occurred with health workers.

Unfortunately, our data are insufficient to explain the differences observed between these two groups of participants. On average, the hospital staff were older, more educated, and more likely to own a cell phone than the health workers. It is plausible to suspect that these factors contributed to the higher accuracy achieved by hospital staff. However, due to logistical reasons, our training procedure also differed between the two groups: health workers were trained in a large group for 6-8 hours, while hospital staff were trained in small groups for 1-2 hours. Our trainers were also somewhat more experienced when working with hospital staff, as health workers were trained first. We re-iterate, however, that training continued until all participants were able to complete two perfect trials on forms and SMS, and one perfect trial on voice.

To better understand the error rates observed using each interface, we tabulate the exact sources of error in Appendix A-2. We classify errors by their entry type (numeric, multiple-choice, yes/no). We also inspect whether each error could be detected, by a trained eye, using the submitted data only; in the future, such errors could potentially be flagged or

ID	Occupation	Education Level	Age	Owns Cell Phone?	Used Cell Phone?	Total Training (Hours)	Accuracy of Entries (Wrong / Total) Forms	SMS	Voice	Time per Interaction (Average) Forms	SMS	Voice
1	Health worker	pre-secondary (class 10)	25		X	8	1 / 22	3 / 22	0 / 11	2:00	1:45	3:07
2	Health worker	pre-secondary (class 10)	25		X	6	2 / 22	1 / 22	0 / 11	1:55	1:12	2:29
3	Health worker	pre-secondary (class 10)	30		X	6	1 / 22	1 / 22	0 / 11	2:15	2:05	2:50
4	Health worker	secondary (class 12)	19			8	2 / 22	1 / 22	0 / 11	1:33	1:27	2:34
5	Health worker	secondary (class 12)	19		X	6	2 / 22	0 / 22	1 / 11	1:45	1:27	2:12
6	Health worker	secondary (class 12)	20	X	X	6	2 / 22	2 / 22	0 / 11	1:35	2:10	2:00
7	Hospital staff	pre-secondary (class 10)	30			2.5	0 / 22	2 / 22	0 / 22	2:25	1:40	2:05
8	Hospital staff	secondary (class 12)	32	X	X	2	0 / 22	1 / 22	0 / 22	1:42	1:17	2:35
9	Hospital staff	secondary (class 12)	28	X	X	0.75	0 / 22	1 / 22	0 / 22	1:30	1:17	1:55
10	Hospital staff	post-secondary (B.A.)	35	X	X	1.5	1 / 22	0 / 22	0 / 22	1:25	3:15	2:00
11	Hospital staff	post-secondary (D. Pharm.)	26	X	X	1	0 / 22	0 / 22	0 / 22	1:05	0:55	2:10
12	Hospital staff	post-secondary (D. Pharm.)	24	X	X	1	0 / 22	1 / 22	0 / 22	1:07	1:25	1:52
13	Hospital staff	post-secondary (M.S.W.)	30	X	X	0.75	1 / 22	0 / 22	0 / 22	1:10	1:15	3:15
						Average (health workers only)	7.6%	6.1%	1.5%	1:50	1:41	2:32
						Average (hospital staff only)	1.3%	3.2%	0%	1:29	1:35	2:16
						Average (across all interactions)	**4.2%**	**4.5%**	**0.45%**	**1:39**	**1:37**	**2:20**
						Std. Dev. (across all interactions)	5.9%	6.4%	2.0%	0:28	0:45	0:28

TABLE III

RESULTS OF THE USER STUDY. ALL PARTICIPANTS WERE EVALUATED ON TWO INTERACTIONS WITH THE FORMS INTERFACE AND TWO INTERACTIONS WITH THE SMS INTERFACE. THE COMMUNITY HEALTH WORKERS (1-6) WERE TESTED ON ONE INTERACTION WITH THE VOICE INTERFACE, WHILE THE PARAMEDIC HOSPITAL STAFF (7-13) WERE TESTED ON TWO INTERACTIONS. AVERAGES AND STANDARD DEVIATIONS ARE SHOWN AT BOTTOM.

automatically fixed using self-correcting forms. Finally, we tabulate whether each error is potentially dangerous (e.g., a severe cough reported as a mild cough would prevent a physician from delivering needed care).

Electronic forms witnessed errors in each entry type; only three of the twelve errors are evident from the values submitted, while five errors may be dangerous. Surprisingly, eight of the errors were due to numeric entry problems on the electronic forms. Two errors were due to a mis-placed decimal point in the temperature entry; while our interface automatically places the decimal point if needed, the user failed to enter the right number of digits in the temperature.

The SMS interface also witnessed errors in each entry type; out of thirteen errors, eight are detectable and seven may be serious. Three of the errors could perhaps be averted with a revision of the SMS cue card: to indicate the absence of a patient cough, many participants entered the code “0” rather than the desired (though perhaps less intuitive) value of “1”. Unlike the forms interface, workers sometimes entered the wrong patient identity when using SMS.

The voice interface witnessed only a single error for the entire duration of the trial. We consulted a videotaped record of the interaction in question (we taped one interaction for each participant), and found that the error was incurred by the operator in translating the participant’s report into a spreadsheet. While such transcription errors could indeed occur in practice, it is encouraging that the participants were not responsible for any errors on the voice interface.

While the voice interface offered the lowest error rates, it also led to the longest entry times. Electronic forms and SMS averaged 1:39 and 1:37 per interaction, respectively, while the voice interface required 2:20 on average (1.43x higher than forms and SMS). One factor that contributed to the slower entry rates using voice was the cellular coverage in our study area; the connection between participants and the operator was highly unreliable. The audio quality was frequently degraded beyond recognition, and calls were occasionally dropped and re-started. While many resource-poor environments have excellent cellular coverage (including the area of Bihar that we are planning to target with our treatment program), the weak coverage in our study area nonetheless reflects a realistic hazard of voice in some environments.

In addition to quantitative results, we also solicited qualitative feedback from each participant, asking them to rank the interfaces by their order of personal preference. The forms and SMS interfaces were most popular amongst the participants, with each receiving six votes as the most popular interface. Only one participant preferred the voice interface to the others. This feedback is indicative of the poor phone connections experienced during the trial; many found voice to be frustrating due to the bad call quality. We were surprised that any participants preferred the SMS interface, given the relatively cryptic message that is produced in the end; however, participants noted that fewer keys are required under SMS than under electronic forms (which requires scrolling and selection). We also note that 8 of the 13 participants preferred the interface on which they demonstrated the fastest entry time.

VII. DISCUSSION

In addition to the factors examined in our experiment, cost is a critical variable for selecting a data collection interface. For the purposes of our own decision making with regards to selecting an interface for our tuberculosis treatment program, we performed a simple cost analysis. Details are provided in

Appendix A-1, but in summary, the expected cost for data collection for each patient during his/her treatment is US $7.89 using electronic forms, US $4.59 using voice, and US $2.99 using SMS[2]. These results show the cost of voice is competitive with the cost of the other two interfaces. Though SMS is slightly cheaper, in order for tracking patient symptom status to be helpful, it is essential that the reported data be close to error-free. This data will be used to guide doctor intervention, and faulty data may lead to unnecessary visits or worse, missed visits when a patient is sick. The voice interface had close to perfect accuracy and was significantly more accurate than SMS or electronic forms. Voice also allows for additional, unscripted information to be easily collected, and provides a social dimension to the health worker's job. We anticipate that this social dimension could potentially lead to higher performance and a lower turnover rate amongst workers, since talking to an operator is likely to increase the worker's feeling of being supported and integrated in a larger project. Voice also allows for verification to be performed easily: operators can simply request the worker to verify the data entry just given, which can be particularly useful for unusual entries. In addition, a voice interface can be replicated very easily in other contexts– no special software or cue cards need to be developed, and any cell phones can be used. While voice requires longer entry times for workers, this represents a very small fraction of their overall working day. For all these reasons, we have now decided to use a voice interface for our upcoming tuberculosis treatment program.

Despite the many advantages of voice, there are still several challenges that must be addressed in practice. In our upcoming treatment program, workers will be actively examining and collecting data from patients and must report this information back to an operator. Calling the operator and keeping him on the line as the worker examines the patient may lead to a slightly awkward social interaction. Another more general challenge for voice interfaces is how to handle scenarios in which a user calls and the operator line is busy. One potential solution for these two challenges is to have the worker write down the data on paper and then call the operator. This introduces an additional opportunity for transcription errors but has the side benefit of creating a paper trail that may be used for later verification. To handle missing calls the operator could be responsible for calling back workers, or workers could leave a message that would be transcribed by the operator.

An alternative solution to these challenges would be to use an interactive voice recognition (IVR) system. IVR could also be useful when there is very frequent data collection or when each survey questionnaire is very long. Hybrid live-operator-IVR systems are also possible, such as directing the worker initially to an IVR system, but automatically transferring the worker to a live operator if the patient symptoms entered are worrisome. We look forward to exploring solutions for handling these different tradeoffs, and considering IVR solutions, as part of our future work.

[2]We use Motorola phones for the electronic forms due to our current set of development tools. Moving to the cheapest available Java-enabled phone would decrease the forms cost to $5.39. However, in practice the cost of voice phones can also be reduced by leveraging existing phones in the community. The cost of voice remains competitive with forms in most practical scenarios.

VIII. Conclusion

Given the widespread excitement in using mobile phones for collecting and analyzing data in the developing world, it is important to establish that the data entered on these devices meets the strict accuracy requirements of health, finance, and other applications. In this study, we provide a quantitative evaluation of data entry accuracy on mobile phones using electronic forms, SMS, and voice interfaces in a resource-poor setting.

Our results indicate that, within the context of our study, the error rates for electronic forms (4.2% of entries wrong) and SMS (4.5% of entries wrong) may be too high to deploy these solutions in a critical application. In contrast, the accuracy of the voice interface was an order of magnitude better (0.45% of entries wrong), with only a single error observed across all trials. This result has influenced us to overhaul our plans for an upcoming tuberculosis program in Bihar, India, to switch to a voice-only interface. Employing a voice interface requires the employment of an operator, and may not be cost-effective in all countries. However, in India, the cost of this operator is more than compensated by the lower cost of voice-only handsets, voice-only cellular plans, decreased training time, and decreased literacy requirements on health workers.

While this study provides an initial data point for the accuracy of data collection on mobile phones, further research is needed to distinguish the factors that are responsible for the errors observed. In the case of electronic forms, we observed error rates that are 3-8x higher than previously measured on PDAs. Our data are insufficient to diagnose whether this difference is due to the devices themselves (screen resolution, touch screen vs. keypad, etc.) or due to other aspects of the evaluation (worker education, training duration, etc.). A future study could address this question directly by evaluating PDAs and mobile phones in the same focus group. However, it is not our goal in this paper to prescribe the optimum device for mobile data collection. Rather, we aim only to highlight that there exists at least one context in which electronic forms and SMS may be too error-prone for large-scale deployment in an accuracy-critical application. In this same context, there is evidence that a low-tech alternative (voice) provides an accurate and cost-effective solution.

IX. Acknowledgments

We are very grateful to the many people who graciously facilitated and participated in our user study. We thank the hospital staff at the Sardar Vallabhbhai National Institute of Technology in Surat, and also at the Reliance TB Hospital in Hazira. We thank Joshnaben Godia of the Taluka Development office in Vagra, Gujarat, and especially Suprava Patnaik for all her help with our study. This work was supported in part by the MIT Public Service Center.

REFERENCES

[1] *Achieving the Health Millennium Development Goals in Asia and the Pacific: Policies and Actions Within Health Systems and Beyond*. United Nations Publications, Dec. 2007.

[2] V. Anantraman, T. Mikkelsen, R. Khilnani, V. S. Kumar, N. R. Machiraju, A. Pentland, and L. Ohno-Machado. Handheld computers for rural healthcare: Experiences from research concept to global operations. In *Workshop on Development by Design*, 2002.

[3] A. Bernabe-Ortiz, W. H. Curioso, M. A. Gonzales, W. Evangelista, J. M. Castagnetto, C. P. Carcamo, J. P. Hughes, P. J. Garcia, G. P. Garnett, and K. K. Holmes. Handheld computers for self-administered sensitive data collection: A comparative study in Peru. *BMC Medical Informatics and Decision Making*, 8:11, 2008.

[4] J. Blaya and H. S. F. Fraser. Development, implementation and preliminary study of a PDA-based bacteriology collection system. *Proceedings of the AMIA Annual Symposium*, 2006:4145, 2006.

[5] J. A. Blaya, W. Gomez, P. Rodriguez, and H. Fraser. Cost and implementation analysis of a personal digital assistant system for laboratory data collection. *The International Journal of Tuberculosis and Lung Disease*, 12(8):921–7, Aug. 2008.

[6] Cost effectiveness study report for the PDA data capture and transmission. SATELLIFE, Sept. 2004. http://www.healthnet.org/coststudy.php.

[7] W. H. Curioso, B. T. Karras, P. E. Campos, C. Buenda, K. K. Holmes, and A. M. Kimball. Design and implementation of Cell-PREVEN: A real-time surveillance system for adverse events using cell phones in Peru. *Proceedings of the AMIA Annual Symposium*, 2005:176180, 2005.

[8] B. DeRenzi, N. Lesh, T. Parikh, C. Sims, W. Maokla, M. Chemba, Y. Hamisi, D. S. Hellenberg, M. Mitchell, and G. Borriello. E-IMCI: Improving pediatric health care in low-income countries. In *CHI*, 2008.

[9] L. Diero, J. Rotich, J. Bii, B. Mamlin, R. Einterz, I. Kalamai, and W. Tierney. A computer-based medical record system and personal digital assistants to assess and follow patients with respiratory tract infections visiting a rural Kenyan health centre. *BMC Medical Informatics and Decision Making*, 6(1):21, 2006.

[10] T. Engebretsen. Acceptance of information technology by health research projects in low-income countries: Intention to use and acceptance of using EpiHandy (IUAUE). Master's thesis, Agder University College, Dec. 2005.

[11] Evaluation of the SATELLIFE PDA project, Feb. 2003. Bridges.org.

[12] D. Forster, R. H. Behrens, H. Campbell, and P. Byass. Evaluation of a computerized field data collection system for health surveys. *Bulletin of the World Health Organization*, 69(1):107–11, 1991.

[13] R. Fynn. Remote HIV/AIDS patient monitoring tool using 3G/GPRS packet-switched mobile technology. In *Appropriate Healthcare Technologies for Developing Countries*, pages 129–138, 2006.

[14] S. Grisedale, M. Graves, and A. Grünsteidl. Designing a graphical user interface for healthcare workers in rural India. In *CHI*, 1997.

[15] T. Groves. SatelLife: Getting relevant information to the developing world. *British Medical Journal*, 313(7072):1606–9, 1996.

[16] P. C. Gupta. Survey of sociodemographic characteristics of tobacco use among 99,598 individuals in Bombay, India using handheld computers. *Tobacco Control*, 5(2):114–120, June 1996.

[17] R. N. Jamison, R. H. Gracely, S. A. Raymond, J. G. Levine, B. Marino, T. J. Herrmann, M. Daly, D. Fram, and N. P. Katz. Comparative study of electronic vs. paper VAS ratings: A randomized, crossover trial using healthy volunteers. *Pain*, 99(1-2):341–7, Sept. 2002.

[18] D. R. Kaufman and J. B. Starren. A methodological framework for evaluating mobile health devices. *Proceedings of the AMIA Annual Symposium*, 2006:978, 2006.

[19] H. Ladd and N. Sewankambo. Uganda health information network: Technical report (submitted to the International Development Research Centre). SATELLIFE & Uganda Chartered HealthNet, Dec. 2004. http://www.healthnet.org/idrcreport.html.

[20] S. O. Lal, F. W. Smith, J. P. Davis, H. Y. Castro, D. W. Smith, D. L. Chinkes, and R. E. Barrow. Palm computer demonstrates a fast and accurate means of burn data collection. *The Journal of Burn Care & Rehabilitation*, 21(6):559–61; discussion 558, 2000.

[21] S. J. Lane, N. M. Heddle, E. Arnold, and I. Walker. A review of randomized controlled trials comparing the effectiveness of hand held computers with paper methods for data collection. *BMC Medical Informatics and Decision Making*, 6:23, 2006.

[22] Mali: Mobile service helps bring down infant mortality, 2007. Balancing Act News, Issue No. 364, http://www.balancingact-africa.com/news/back/balancing-act_364.html.

[23] J. S. McBride, R. T. Anderson, and J. L. Bahnson. Using a hand-held computer to collect data in an orthopedic outpatient clinic: A randomized trial of two survey methods. *Medical Care*, 37(7):647–51, July 1999.

[24] M. A. Missinou, C. H. O. Olola, S. Issifou, P.-B. Matsiegui, A. A. Adegnika, S. Borrmann, D. Wypij, T. E. Taylor, and P. G. Kremsner. Short report: Piloting paperless data entry for clinical research in Africa. *Amer. Journ. of Trop. Medicine and Hygiene*, 72(3):301–303, Mar. 2005.

[25] Mobilizing against malaria. *Expanding Horizons (by Nokia)*, pages 12–13, Feb. 2007.

[26] Nokia CEO message, 2008. http://www.nokia.com/A4942317.

[27] T. S. Parikh. Using mobile phones for secure, distributed document processing in the developing world. *Pervasive Computing*, 4(2), 2005.

[28] T. S. Parikh, P. Javid, S. K, K. Ghosh, and K. Toyama. Mobile phones and paper documents: Evaluating a new approach for capturing microfinance data in rural India. In *CHI*, 2006.

[29] P. Quinn, J. Goka, and H. Richardson. Assessment of an electronic daily diary in patients with overactive bladder. *BJU International*, 91(7):647–52, May 2003.

[30] J. Selanikio. Advantages of using handheld computers for data collection. DataDyne. http://www.datadyne.org/files/Brief-HandheldAdvantages.pdf.

[31] J. Selanikio and R. Donna. Datadyne brief. DataDyne. http://www.datadyne.org/files/DataDyne_brief.pdf.

[32] K. Shirima, O. Mukasa, J. Schellenberg, F. Manzi, D. John, A. Mushi, M. Mrisho, M. Tanner, H. Mshinda, and D. Schellenberg. The use of personal digital assistants for data entry at the point of collection in a large household survey in southern Tanzania. *Emerging Themes in Epidemiology*, 4(1):5, 2007.

[33] D. Skinner, U. Rivette, and C. Bloomberg. Evaluation of use of cellphones to aid compliance with drug therapy for HIV patients. *AIDS Care*, 19(5):605–607, 2007.

[34] SMS data management. Dimagi, Inc., 2008. http://www.dimagi.com/content/sms-data-management.html.

[35] B. Tiplady, G. Crompton, M. Dewar, F. Boellert, S. Matusiewicz, L. Campbell, and D. Brackenridge. The use of electronic diaries in respiratory studies. *Drug Information Journal*, 31(3):759–764, 1997.

[36] I. Walker, C. Sigouin, J. Sek, T. Almonte, J. Carruthers, A. Chan, M. Pai, and N. Heddle. Comparing hand-held computers and paper diaries for haemophilia home therapy: A randomized trial. *Haemophilia*, 10(6):698–704, Nov. 2004.

APPENDIX A-1. COST ANALYSIS

In our basic cost analysis we first assume that the treatment pool is 1000 patients. In our treatment program each worker is responsible for 10 patients, so there is a total of 100 workers. Each worker must be equipped with a cell phone. Our current development tools for electronic forms are tied to Motorola, and require a Java-enabled phone. The cheapest such phone in India is the Motorola L6i which is 75 US dollars. In contrast, both the SMS interface and voice interface can be used on any cell phone, one of the cheapest of which is the Motorola Motofone F3 ($26). Voice calls are slightly more expensive than text messages: a 3-minute voice call, which is longer than the average time in our experimental study, is about 3 rupees (0.065 US dollars, Airtel carrier). SMS messages using Airtel are 1.5 rupees per message (0.0327 US dollars). The average call length in our user study is 2 minutes and 20 seconds; therefore conducting 100 calls would require slightly under 4 hours. We therefore anticipate that a 100 call load would be reasonable for an operator working 8–9 hours per day, in order to include a liberal number of breaks. Our program design involves each worker visiting each patient to record symptom information every two weeks. At this rate a single operator working five days per week could handle the 1000 calls over the two week period. Based on our experience in hiring a qualified operator in Bihar for $100 per month, we choose a conservative estimate of an operator salary of $200 per month. The length of treatment is six months.

Given the assumptions above, we calculate the total cost per patient over the course of the treatment for each interface. Note that we are only focusing here on the aspects of the interfaces that lead to different costs and we are not considering the salary of the workers or additional program overhead. The cost of the phone per patient is simply the cost per phone multiplied by the number of workers divided by the number of patients. The cost of an operator per patient is the salary of the operator per month ($200), multiplied by the 6 month treatment length, divided by the number of patients, yielding a cost of $1.20 per patient. Workers will upload health information approximately 12 times per patient (once every two weeks). Therefore the cost of communication per patient is equal to the cost for each data entry (either SMS or a voice call) multiplied by 12. Table IV displays the cost breakdown per patient.

Due to the high cost of phones that can support external applications, such as Java-enabled phones, voice is cheaper than electronic forms over a single 1000-patient program, even given the ongoing cost of an operator salary. SMS is the cheapest since it requires no operator and can be used with any phone. Perhaps most important is that the cost for each interfaces is less than $10, a small sum compared to the total cost of approximately $90–100 needed to treat a tuberculosis patient in India.

Interface	Fixed Cost	Marginal (Ongoing) Cost	Total cost
Forms	$7.50	$0.39	$7.89
Voice	$2.60	$1.99	$4.59
SMS	$2.60	$0.39	$2.99

TABLE IV

APPROXIMATE COST PER PATIENT INCURRED BY EACH USER INTERFACE AS PART OF A 6-MONTH TUBERCULOSIS TREATMENT PROGRAM IN INDIA. FIXED COSTS COVER THE PHONE, WHILE MARGINAL COSTS COVER TRANSMISSION VIA VOICE OR SMS, AND, WHERE APPLICABLE, THE CALL OPERATOR SALARY. HEALTH WORKER SALARIES DO NOT DEPEND ON THE INTERFACE AND ARE EXCLUDED.

This cost analysis assumes that we continue to use the Motorola L6i Java-enabled phone for the electronic forms interface. There are some cheaper Java-enabled phones that we may be able to use in the future, such as the $50 Nokia 2626, but this would require us to obtain new development tools. This would change the cost per patient for electronic forms to be $5.39. This still means that voice is less expensive than forms in terms of cost per patient. Also, the cost of voice could be further reduced by leveraging existing phones belonging to the health workers.

While the above analysis is conducted for a specific program in India, informal data suggests that in some other countries voice may also be worth considering. For example, the average salary of call center operators in Peru is approximately 150 US dollars per month. The biggest cost considerations when comparing interfaces in new locations are likely to be the operator salary, the cost of voice calls compared to SMS, and the expected frequency and duration of conversations between workers and the operator.

APPENDIX A-2. DETAILED LOG OF ALL DATA ENTRY ERRORS

Error Number	Interface Mode	Entry Type	Entry Name	Correct Entry	Actual Entry	Error Detectable?	Error Dangerous?
1	Forms	Multiple-choice	Cough	"mild"	"none"		X
2	Forms	Multiple-choice	Cough	"heavy"	"mild"		X
3	Forms	Numeric	Temperature	100.3	103.0		X
7	Forms	Numeric	Temperature	100.8	108.0	X	
4	Forms	Numeric	Temperature	98.5	98		
5	Forms	Numeric	Temperature	98.7	98.687		
6	Forms	Numeric	Temperature	100.2	100.0		
8	Forms	Numeric	Weight	62	*empty*	X	
9	Forms	Numeric	Weight	68	67		
10	Forms	Numeric	Weight	68	93	X	
11	Forms	Yes/No	Fatigue	Yes	No		X
12	Forms	Yes/No	Nausea	No	Yes		X
13	SMS	Multiple-choice	Cough	"1" (none)	"0" (disallowed)	X	
14	SMS	Multiple-choice	Cough	"1" (none)	"0" (disallowed)	X	
15	SMS	Multiple-choice	Cough	"1" (none)	"0" (disallowed)	X	
16	SMS	Multiple-choice	Cough	"3" (mild)	"0" (disallowed)	X	
17	SMS	Multiple-choice	Cough	"5" (severe)	*missing*	X	X
18	SMS	Multiple-choice	Patient ID	"6" (Akshaye Khanna)	"5" (Akshay Kumar)		X
19	SMS	Multiple-choice	Patient ID	"7" (Anil Kapoor)	"1" (Aamir Khan)		X
20	SMS	Numeric	Temperature	1003	103		X
21	SMS	Numeric	Weight	54	45		X
22	SMS	Numeric	Weight	62	826	X	
23	SMS	Numeric	Weight	69	59		X
24	SMS	Yes/No	Yellow eyes	"6"	"2"	X	
25	SMS	Yes/No	Fatigue	"0000007"	"000007"	X	
26	Voice	Numeric	Weight	69	59		X