Field Evaluation of a Camera-Based Mobile Health System in Low-Resource Settings

Nicola Dell[1] **Ian Francis**[2] **Haynes Sheppard**[2] **Raiva Simbi**[3] **Gaetano Borriello**[1]

[1]Computer Science and Engineering
University of Washington
{nixdell;gaetano}@cs.washington.edu

[2]Global Solutions for
Infectious Diseases
ifrancis@gsid.org; haynes.sheppard@gmail.com

[3]Zimbabwe Ministry of
Health and Child Care
raivasimbi@gmail.com

ABSTRACT
The worldwide adoption of mobile devices presents an opportunity to build mobile systems to support health workers in low-resource settings. This paper presents an in-depth field evaluation of a mobile system that uses a smartphone's built-in camera and computer vision to capture and analyze diagnostic tests for infectious diseases. We describe how health workers integrate the system into their daily clinical workflow and detail important differences in system usage between small clinics and large hospitals that could inform the design of future mobile health systems. We also describe a variety of strategies that health workers developed to overcome poor network connectivity and transmit data to a central database. Finally, we show strong agreement between our system's computed diagnoses and trained health workers' visual diagnoses, which suggests that our system could aid disease diagnosis in a variety of scenarios. Our findings will help to guide ministries of health and other stakeholders working to deploy mobile health systems in similar environments.

Author Keywords
Mobile phone; smartphone; mHealth; point-of-care diagnostics; camera; sensors; vision; ICTD; HCI4D.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
The rapid increase in mobile device penetration throughout the world is providing an opportunity for mobile systems to play an important role in the delivery of health and information services to people in developing countries. Mobile devices are portable, battery-powered, relatively low-cost and can connect to the Internet via cellular networks. These properties make them more suitable for deployment in low-resource environments than traditional desktop computing platforms. As a result, the emerging research area of mobile health has evolved to study the design and use of mobile systems for healthcare delivery in low-resource settings. Categories of research include informing people about health issues [5], improving adherence to medical protocols [6], providing remote consultation [13] and collecting data in survey format [11].

MobileHCI '14, September 23 - 26 2014, Toronto, ON, Canada
Copyright 2014 ACM 978-1-4503-3004-6/14/09…$15.00
http://dx.doi.org/10.1145/2628363.2628366

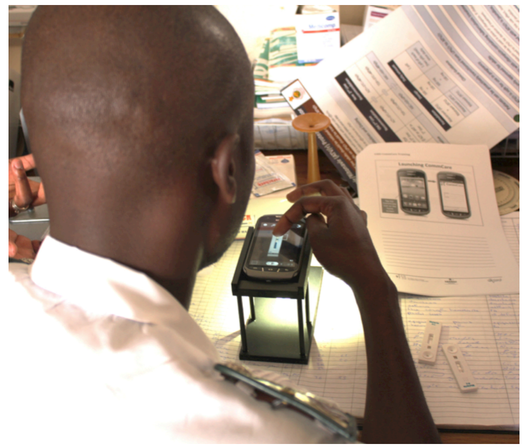

Figure 1: A health worker uses our system in a rural clinic in Zimbabwe to analyze diagnostic tests for malaria.

In addition, a variety of emerging technologies aim to transform a mobile device into a portable medical platform using sensors built-in or attached to the device. SpiroSmart [9] uses the built-in microphone to measure lung function, CellScope [20] uses a lens attachment to turn the camera into a microscope, while NETRA [16] attaches an optical probe to the camera to detect eye anomalies. Although these sensor-based systems may have the potential to transform healthcare in low-resource settings, the majority of them are yet to be evaluated with health workers in the field. As a result, little is known about the challenges and issues that might arise when health workers attempt to integrate mobile sensor-based systems into their patient care routines.

Our work extends the state-of-the-art by exploring the user interaction and deployment challenges that arise when health workers in low-resource settings integrate a mobile camera-based system into their clinical workflow. We designed a system that uses a smartphone's built-in camera and computer vision to capture, analyze and report diagnostic tests for infectious diseases (see Figure 1). Previous research [3] describes an initial implementation of the system's image analysis algorithms. In this paper, we provide an extensive field evaluation of the system with sixty health workers at five hospitals and clinics in Zimbabwe. This is the first time the system has been evaluated with real users in low-resource settings. We explore how health workers are able to integrate the system into their daily patient care routines, both at small, rural clinics and at large, urban hospitals, and we expose a variety of issues that are generally applicable to a range of mobile sensor-based systems (such as data collection issues, user errors, infrastructure challenges, etc.).

Our field evaluation focuses on three important research questions: (i) the impact of the mobile system on health workers' patient care routines, (ii) the impact of poor infrastructure on system usage and data collection, and (iii) the quality of the mobile system's automatically computed diagnoses. Our findings show that health workers were able to integrate the system into their clinical workflow and successfully collect extensive amounts of test data after only sixty minutes of training. Furthermore, health workers used our system to capture and analyze diagnostic tests consistently over an eight-week period. In addition, health workers developed a variety of strategies to overcome poor network connectivity and transmit data to a centralized database. Finally, we show strong agreement between the system's computed diagnoses and the visual diagnoses provided by well-trained health workers, which suggests that the system could assist with disease diagnosis in a variety of scenarios. Taken together, our findings will help ministries of health and other stakeholders to assess the viability and acceptability of deploying mobile sensor-based systems to assist health workers in the field. In addition, our insights will guide future researchers working to deploy mobile health systems in similar environments.

BACKGROUND AND RELATED WORK

Rapid Diagnostic Tests for Low-Resource Settings

Health workers in low-resource settings often lack access to convenient, affordable and usable diagnostic technologies that could help them to quickly diagnose and treat infectious diseases. To address this challenge, low-cost, commercially available rapid diagnostic tests (see Figure 2) have been developed and are now routinely used throughout the world to diagnose a variety of common diseases [12]. For example, the 2013 World Malaria Report [21] showed that a total of 205 million rapid diagnostic tests for malaria were manufactured and sold globally in 2012.

Although the potential benefits of these new diagnostic technologies are immense, little attention has been paid to the challenges faced by the health workers responsible for administering tests and interpreting their results. To run a test, a health worker places a drop of blood on the test, waits a recommended amount of time (usually 10-20 minutes) and then interprets a series of colored control and test lines (see Figure 2) to determine the result. Although interpreting tests may appear to be simple, there are many challenges involved in reading the tests by eye, particularly for health workers in low-resource settings. Human interpretation is subjective and health workers often lack confidence in their ability to read test results correctly [18]. The variety of tests available is also expanding rapidly [12] and the numbers and positions of result lines vary across diseases and test brands, which increases the potential for human error. Researchers are also in the process of developing more complex tests whose results will require quantification [1] or time-sensitive analysis [4] that will be extremely difficult or impossible to analyze by eye. Finally, health workers in low-resource settings may have poor eyesight that could impair their ability to interpret the test.

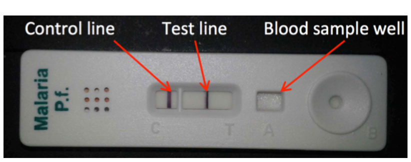

Figure 2: A low-cost rapid diagnostic test for malaria.

In our work, we address these challenges by building a mobile system to objectively interpret the tests. Analysis of tests can become a standardized, auditable and adjustable process that can be changed without retraining users. Commodity smartphones can analyze a variety of tests, negating the need for specialized reader devices, and new tests can be added to the system as they become available. Furthermore, a mobile system can also collect and transmit relevant patient and disease data to a central database. Providing governments and epidemiologists with timely, accurate statistics regarding the diseases detected could aid outbreak detection, supply chain management, evidence-based decision-making, health system evaluation and global disease surveillance and control efforts.

Mobile Health Systems for Low-Resource Settings

A range of prior research highlights the potential for mobile systems to improve healthcare in low-resource settings. Research areas include informing people about health issues [5] and improving adherence to medical protocols [6]. Several mobile data collection systems have also been widely deployed [5, 8, 11]. However, these deployments focus on manually collecting textual data in survey format rather than processing rich data collected from sensors. Telemedicine systems allow users to capture images in the field and transmit them for remote analysis, either by a human expert [13], or by complex image processing algorithms running in a data center [17]. However, many health facilities in low-resource settings lack the reliable network connection required to transmit high-resolution images for off-site analysis. In addition, the expert analyzing the images is likely to be a busy medical professional and it may take days to send a result back to the health worker. These challenges suggest that it may be beneficial to perform any necessary computation locally on the device rather than transmitting it for off-site analysis.

As discussed in the previous section, innovative mobile health systems have also recently been developed that aim to process data collected from sensors built-in or attached to the device [9, 18, 22]. Although these systems have the potential to transform healthcare delivery in low-resource settings, they are yet to be evaluated with real users in the field, and little is known about the issues that might arise when these systems are integrated into the clinical workflow. A notable exception by Chaudri et al. [2] uses an external temperature sensor attached to a smartphone to monitor the pasteurization of human breast milk. The system was evaluated in two South African hospitals with promising preliminary results, although the study did not consider how the system may be used in small, rural clinics.

Mobile Systems for Analyzing Diagnostic Tests

In addition to our work, several recent mobile systems also target the capture and analysis of rapid diagnostic tests. Matthews et al. [14] describe a mobile system that uses image processing to automatically interpret a paper-based test for dengue fever. However, the algorithm has not yet been fully evaluated and only targets a single, special-purpose test for dengue rather than a range of tests for a variety of diseases. Mudanyali et al. [15] developed a smartphone-based diagnostic test reader that controls the environment in which images of tests using lighting. By contrast, our system uses only the device's built-in flash to light the image, which is advantageous since requiring additional battery-powered parts increases the likelihood that one of the parts will run out of charge and render the system non-functional.

The FIO Corporation [7] markets a smartphone-based diagnostic test reader capable of imaging a variety of tests. However, the device is enclosed in a specialized container, which limits it to reading diagnostic tests and prevents it from becoming a general-purpose medical platform. In addition, users must purchase the company's proprietary devices and cloud-based services. By contrast, our system is open-source, allows users to add their own tests to the system, and works with a variety of mobile devices and server solutions. Lastly, to the best of our knowledge, FIO's system is the only smartphone-based diagnostic test reader that has been evaluated in the field [19]. However, the study focused solely on the accuracy of the system's medical diagnoses and did not consider any data collection, usability or infrastructure challenges faced by users in the field.

SYSTEM DESIGN

Our work extends the state-of-the-art by exploring the user interaction and deployment challenges that arise when health workers in low-resource settings integrate a mobile camera-based system into their clinical workflow. Specifically, we designed and built a smartphone system that uses the device's built-in camera to capture and analyze rapid diagnostic tests for infectious diseases. The rest of this section describes our design principles, user interaction and system workflow, and implementation details.

Usage Scenario and Design Principles

We target the following usage scenario: a health worker in a rural clinic has been issued a mobile device to assist with clinical tasks. The health worker enters a patient exam room to assess a patient. After spending time interacting with the patient and using the device to collect patient data, the health worker decides to perform a diagnostic test for malaria. (S)he obtains a blood sample from the patient and administers a rapid diagnostic test for malaria. Then, (s)he uses the device's built-in camera to capture an image of the test. The system processes the image, displays the test result on the screen, and transmits the data to a database if and when there is sufficient connectivity. Finally, based on the test result, the health worker recommends the appropriate treatment or selects another test to run.

In addition to being appropriate for our target problem of interpreting diagnostic tests, many of our design principles are also applicable to a range of other mobile sensor-based systems. For example, similar usage scenarios might apply to capturing images for cell-phone microscopy [16] or sound to analyze lung function [9].

Commercially available devices

Requiring users to import, configure and maintain custom hardware systems may present a significant barrier for many users in low-resource settings. Therefore, our system runs on commercially available Android devices and only uses sensors that come built into these devices. The system does not use any additional hardware or specialized reader device and analyzes rapid diagnostic tests using only images captured from the built-in camera. We have tested the system using a variety of smartphones with 5 mega-pixel cameras and above. The variety of Android devices available in low-resource settings will allow users with varying needs and budgets to choose devices that best fit their requirements. In addition, smartphones are multi-purpose devices that can also provide utility beyond diagnostics (such as voice and text communications, patient management or data reporting).

Local computation

Prior research explores the idea of transmitting images for remote analysis [13, 19]. However, many low-resource settings lack sufficient Internet connectivity to reliably transmit high-resolution images. Thus, our system performs all of the computation and image processing on the device, which allows the system to be fully functional in the absence of a network connection.

Asynchronous transmission

In many low-resource settings, the quality of the network connection might vary with network traffic, time of day and power outages. Thus, our system supports asynchronous data transmission in which collected images and data are stored locally on the device until a network connection or sufficient bandwidth becomes available for transmission.

Data collection

In addition to capturing and processing images, our system allows users to collect and report data (including text, barcodes, GPS data, *etc*.) about the patients seen and diseases detected. Rather than building a new data collection platform, we integrated the image processing components of the system with CommCare [5], a widely used data collection platform based on Open Data Kit [11].

Easily configurable

We wanted to make it easy for test manufacturers and clinical experts to add new rapid diagnostic tests to the system and to control the sensitivity of the image-processing algorithm for their specific test. To add a new test, a clinical expert creates a simple text file, called a *test description file*, which specifies the size and location of the regions that contain the test results. The test description file is uploaded to the device and used to configure the mobile system for use by health workers in the field. The health workers will not adjust the system parameters. Instead, health workers will simply use the smartphone's built-in camera to capture images of tests that have previously been added to the system by clinical experts. A typical test description file consists of about ten lines of text and may be created using any text editor or using a graphical web tool that we developed. Thus far, clinical experts in the laboratory have created and tested description files a variety of rapid diagnostic tests for different diseases and test brands, including tests for malaria, HIV and syphilis [3].

Interaction Design and System Workflow

A prior paper describes the system architecture and image processing algorithms in detail [3]. Here, we provide an overview of the steps required for a health worker to interact with the system to capture and process a diagnostic test. A health worker begins by launching the system and using the mobile application to record any relevant patient data (such as age, gender, *etc*.) They then select the type of diagnostic test from a menu on the screen that contains a list of all the tests that have been added to the system. Selecting a test automatically launches the Android camera application, which allows users to take (and, if necessary, re-take) and image of the relevant diagnostic test. When the user is happy with the image that they captured, they press a button to save and process the image.

The next step in the workflow is to align the captured image, which involves separating the part of the image that contains the diagnostic test cartridge from the background. After aligning the image, the system displays the test on the screen, and provides users with an opportunity to re-take the image should the test appear to be misaligned. If the test has aligned correctly, the user presses a button to compute the diagnosis. The system uses test description file that corresponds to the chosen diagnostic test to locate and process each region on the test that the clinical expert specified as showing a result. For each region, the system uses a thresholding algorithm (described in detail in [3]) to determine if a line is present in that region. Finally, the system uses the combination of detected lines to determine the final test result. After the result has been computed, it is displayed on the screen with an image that shows the user the results of processing so that they can verify the diagnosis. Figure 3 shows the system's user interface for a processed malaria test with a positive result. After checking the result, the user saves and exits the application. The system then stores all of the data associated with the captured test locally on the smartphone, and transmits this data, along with an image of the captured test, to a central database if and when a network connection is available.

Figure 3: The mobile system's user interface, showing a processed diagnostic test for malaria with a positive result.

FIELD EVALUATION

We conducted an eight-week field study with sixty users at five hospitals and clinics in Zimbabwe to explore the impact that using our mobile camera-based system had on health workers' patient care routines. Although we have extensively tested the system with rapid diagnostic tests for multiple diseases in the laboratory, including tests for HIV, syphilis and flu [3], our field evaluation in Zimbabwe targeted only diagnostic tests for malaria. The practical constraints of integrating a new system into government public health facilities limited our deployment to diagnostic tests that are already routinely purchased and distributed by the Ministry of Health and Child Care. Rapid diagnostic tests for malaria are available at all hospitals and clinics in Zimbabwe and are used to test and treat patients in a variety of clinical departments, including the maternal and child health, outpatient, opportunistic infections and maternity departments. Health workers in Zimbabwe have been trained to administer rapid diagnostic tests for malaria and use them on a daily basis, and the Ministry of Health allowed us to analyze these tests. We acknowledge that only evaluating the mobile system with malaria tests is a limitation of our study. However, since the process of administering rapid diagnostic tests for other diseases is the same as administering tests for malaria, the workflow to analyze tests for other diseases would be identical.

Research Questions

This study was the first time that the system has been evaluated with health workers in low-resource settings. We formalize our study into three research questions:

Q1: To what extent are health workers in low-resource settings willing and able to integrate a mobile camera-based system into their daily patient care routine?

Although many health workers in Zimbabwe own basic mobile phones, the majority have never used a touchscreen device. We wanted to assess how easily health workers learned to interact with these new devices, how easily groups of health workers within a clinic were able to use a single device as a shared resource, and how system usage might vary across health facilities of different sizes. Finally, we wanted to explore the range of user errors that occurred when health workers used the system in the field.

Q2: What infrastructure challenges are faced by health facilities using the system and what strategies might be used to overcome these challenges?

The rapid growth of mobile networks in developing regions has led to the widespread deployment of 2G/3G networks. However, network data speeds are usually erratic and availability can be intermittent. We wanted to examine how network connectivity varied at different sites and see if the sites possessed sufficient bandwidth to transmit collected data and images. In addition, we were interested in what strategies were used to keep the devices safe and charged.

Q3: To what extent do the automatically processed diagnoses computed by the mobile system agree with the visual diagnoses made by trained health workers?

One of our long-term goals is for the system to help health workers who may not have much medical background to correctly diagnose patients. However, since this study is the first field evaluation of the system, we did not want to affect the standard of patient care at this stage. Thus, health workers who participated in our study were trained nurses whose interpretation of tests is currently used to guide patient care. All patient treatment was based on their visual diagnoses collected prior to image processing. This enabled a blinded comparison of the visual diagnoses and the system diagnoses to assess if the system might be used to direct patient care in the future.

Study Sites

Following approval of the study protocols by the IRB, we identified five clinical study sites in the Manicaland province: the Mutare provincial hospital, the Hauna and Nyanga district hospitals, and the Tombo and Zindi rural health centers. We selected sites that typically have a high prevalence of malaria testing and deliberately targeted facilities that ranged in size from large urban hospitals to small rural clinics to explore how system usage varied at different levels of the healthcare hierarchy. Mutare hospital, located in Zimbabwe's third largest city, is the biggest hospital in Manicaland province. Nyanga district hospital is in a town roughly 100km from Mutare. Hauna hospital is located in a substantially more rural and difficult to access valley roughly 100km from Nyanga. Finally, the Tombo and Zindi clinics are in small, rural villages roughly 80km from Nyanga and 60km from Hauna respectively.

Table 1: Summary of study sites, participants and devices.

Study site	No. of Participants	No. of Devices
Mutare hospital	27	4
Nyanga hospital	17	4
Hauna hospital	11	4
Zindi clinic	3	2
Tombo clinic	2	1
All sites	**60**	**15**

Participants

We recruited 60 health workers (47 female) across the five study sites (see Table 1). Participation was limited to trained health workers who conduct malaria diagnostic tests regularly as part of their job. Participants ranged in age from 24 to 64 years (M = 35) and had between 1 and 42 years (M = 6) of experience employed as health workers. Most participants owned basic mobile phones, but 32 had never used a touchscreen device and another 15 had less than six months experience using a touchscreen device.

Apparatus

We configured our mobile application to collect patient data and capture diagnostic tests for malaria. Since our study deals with sensitive personal health information, all the data we collected was anonymized. We recorded the patient's gender, age, location, test type, date, time, the health worker's visual diagnosis and the system's computed diagnosis. The application was also instrumented to record the times at which the test started and ended and the time at which the captured data was transmitted to the server.

We deployed the mobile system on Samsung Galaxy XCover 2 Android devices because they were readily available and moderately priced in Zimbabwe. The XCover 2 has a 1 GHz dual core processor, a 5 mega-pixel built-in camera, a 4-inch capacitive touchscreen and a ruggedized plastic cover to protect the device from dust and moisture. Each device was loaded with a 2GB external SD card and a local SIM card with a 300MB prepaid data bundle. As shown in Table 1, the provincial and district hospitals received four devices each, with one device placed in each of the clinical departments that perform the most malaria diagnostic tests (the maternal and child health department, the outpatient department, the maternity department and the opportunistic infections department). Since rural clinics typically employ fewer health workers than hospitals, each clinic was issued with one or two devices to cover their patient load. In total, we deployed 15 devices across five study sites. Each device was accompanied by a simple plastic stand (see Figure 4) designed to provide stability and a fixed focal length for the device's camera. In addition, we distributed printed user manuals to provide support and guidance for participants in the absence of the researchers.

Figure 4: A simple plastic stand holds the mobile device in a convenient position above the diagnostic test.

Procedure

Preliminary Site Assessment

Before beginning the study, we visited Zimbabwe to assess the selected study sites, understand the health workers' current routines, and gain insight into how a mobile system might be successfully integrated into the clinical workflow. We also tested the network connectivity and data transmission rates at each site.

Adding Diagnostic Tests to the System

During our site assessment, we obtained the details of the three brands of malaria tests currently in use at the sites. Following this, a clinical expert at Global Solutions for Infectious Diseases (San Francisco, USA) [10] created the test description files to add the malaria tests to the system. Encouragingly, creating new test description files was relatively quick, taking roughly 30 minutes per test. Rigorously testing the system parameters for each test took longer, typically several days, since it was necessary to prepare and administer series' of tests with differing dilutions of malaria infected blood under a variety of lighting conditions, before empirically determining the parameters that resulted in the optimal sensitivity and specificity. To make the optimization phase easier, we created a version of the system that could batch process a variety of different test images and parameters at once, which allowed the clinical expert to more quickly narrow in on the optimal algorithm parameters for each test.

Training and Deployment

After adding the malaria tests to the system, we returned to Zimbabwe and conducted participant training sessions at each study site. Each training session lasted approximately 60 minutes and began by demonstrating the system to a group of participants. Participants were then divided into pairs and given time to read user manuals, ask questions and practice capturing and reporting diagnostic test data. The amount of practice that participants required varied depending on their familiarity with technology, although all participants mastered the system within the training session. During each session, we explained that this was the first time the mobile system was being field tested and that all patient care should continue to be based on participants' visual diagnoses. At the end of each session we collected demographic data regarding each participant's age, work experience and familiarity with technology.

After the training sessions, participants were asked to integrate the system into their daily patient care routine and to capture and transmit data every time that they performed a malaria test. At the hospitals, we explained that each of the main clinical departments that conducted malaria tests would receive one device and that participants working in these departments would share the device accordingly. We conducted follow-up visits at each site one week and six weeks after the training sessions to assess how easily participants were able to integrate the technology into their workflow. During these visits, we observed participants using the system and conducted semi-structured interviews to collect data about participants' opinions and experiences.

RESULTS AND DISCUSSION

Q1: Integrating the System into the Clinical Workflow

Our first research question explores the extent to which participants were able to integrate the system into their clinical workflow, including how they used the system and what errors they made when interacting with the system.

Analyzing System Usage and Data Collection

Health workers reacted positively to the introduction of the mobile system into their daily clinical workflow. Most participants were eager to learn the new technology and excited that the system represented progress for Zimbabwe's national health system. Figure 5 shows the number of tests captured by each site. In total, participants captured and transmitted 1828 malaria tests over the two-month study period. The median time that it took to capture and process a test and record patient data was 1.5 minutes (ranging from an average of 2.5 minutes at Mutare to 1.3 minutes at Nyanga).

In addition, a large proportion of participants immediately identified the potential for the system to reduce their paperwork burden by automatically generating aggregated monthly reports. Participants at the small clinics told us that they were currently responsible for manually maintaining a large number of paper-based patient registers, several of which required the same data to be repeated and most of which required a written monthly report. They were eager to show us which registers could be replaced by collecting data using the mobile system, and they requested that the system be extended to enable them to collect a broader range of health information beyond diagnostic test data.

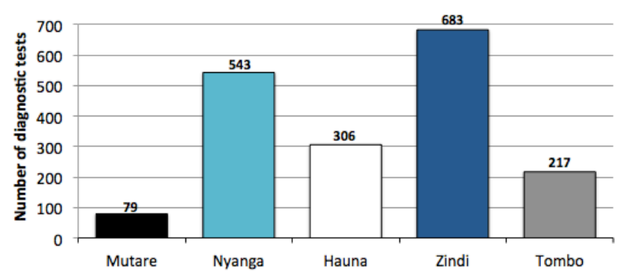

Figure 5: Total number of malaria diagnostic tests captured by each study site.

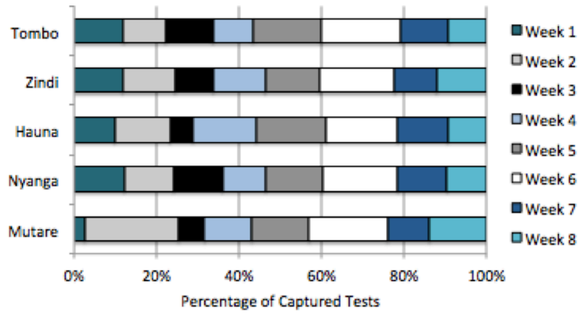

Figure 6: Proportion of each site's total tests that were collected during each week of the study.

Figure 6 shows the proportion of each site's captured tests that were collected during each week of the study. Encouragingly, participants used the system regularly to collect data for the duration of the study, which suggests that they were generally able to integrate the system into their daily clinical routines. In addition, several participants mentioned that patients responded favorably to the new system, believing that it was important for their diagnostic test data to be reported directly to the Ministry of Health.

Somewhat surprisingly Zindi, one of the small rural clinics, captured the largest number of tests (see Figure 5). This could be explained by several factors. First, Zindi is located in the most malaria endemic region of all the study sites and typically performs a large number of diagnostic tests during the malaria season. In addition, small clinics are usually a patient's first point of contact with the health system and represent an ideal place at which to quickly diagnose and treat malaria. Finally, we found that the health workers at Zindi set up a dedicated desk that they used exclusively for performing malaria tests. Placing the mobile device on the same desk made it easy for them to integrate the system into their current workflow and capture tests as necessary.

By contrast, Mutare, the largest hospital in the study, captured the smallest number of tests (see Figure 5). There are several possible explanations for this. First, since visiting the hospital is substantially more expensive than visiting a clinic, many patients only visit the hospital if they are referred there by a clinic. In these cases, patients often already received a diagnostic test at the clinic and there is no need to perform another test at the hospital. In addition, the clinical departments at Mutare hospital are much larger than those at the district hospitals and have many patient exam rooms, and it was challenging for participants to coordinate usage of a single device across multiple rooms. Several participants explained that sometimes when they needed to use the device, it was already being used in another room. Thus, for large hospitals, a better deployment model may be to provide one device per exam room.

We also discovered that the large number of patients at the hospital increased participants' fear that the devices would get stolen if left in a room unattended. As a result, participants frequently locked the device in a cabinet when they left the room. Other participants were then unable to access the device until they found the health worker with the key. During a follow-up visit, we explained to participants that it would be better to use the system and risk theft of the device than keep it locked away. This encouragement may have been responsible for the larger number of tests collected in Mutare in the second week of the study (see Figure 6) although the number of tests decreased again in subsequent weeks. At the end of the study several participants suggested that it would be better for each health worker to have a device and carry it around the hospital. However, this would substantially increase the number of devices and overall cost of deploying the system.

Analyzing User Errors

To analyze the errors that participants made when they interacted with the system, we manually examined all of the captured images and coded any anomalies. User errors fell into five categories: (1) *no image captured* - users collected patient data but failed to capture an image of a test; (2) *captured image is not of a diagnostic test* – in these cases the image was usually of the desk or stand and probably captured accidentally; (3) *incorrect test selected* – the type of test shown in the image is different to the test type selected by the user; (4) *test placed upside-down* – the test in the image was placed the wrong way up; and (5) *unusable image* – the captured image is not usable for analysis because it is out of focus, overexposed or the device was not positioned in the stand correctly.

Figure 7 shows the percentage of each site's captured tests that contained each type of error. Overall, we identified errors in 114 out of 1828 tests (6.2%). The most common error was placing the test upside-down (53/1828 or 2.9%), followed by incorrectly selecting the test type (31/1828 or 1.7%). The other error types were more infrequent, with users failing to capture an image in 11/1828 or 0.6%, capturing an image of something other than a test in 6/1828 or 0.3%, and placing the device incorrectly in 13/1828 or 0.7%. Zindi had the lowest error rate at 2.2%, while Tombo had the highest error rate at 11.5%. One reason for the high error rate at Tombo was due to one participant repeatedly placing the device incorrectly in the stand (see Figure 7).

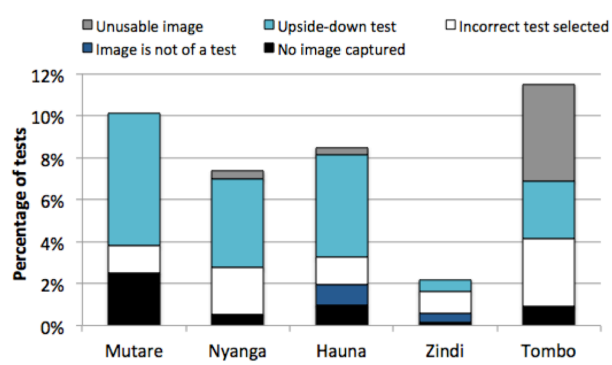

Figure 7: Percentage of each site's total tests that contained each type of user error.

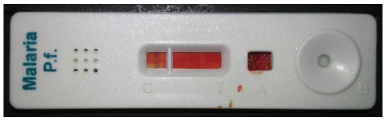

Figure 8: A diagnostic test that contains too much red background to yield a valid diagnosis.

We plan to address several error types through additions to the algorithm. We have already updated the system to make it impossible for users to submit data without capturing an image. In addition, for tests that contain any identifying markings (such as a brand name), it should be relatively straightforward to automatically detect the test type and/or if the test is upside-down. Finally, additional or repeated training sessions would likely also decrease user errors.

Using the System for Quality Control

Our findings also show that the system has the potential to be a valuable quality control tool. Early in the study, our analysis revealed a series of transmitted images in which the tests had been incorrectly administered since the test strip had a dark red background (see Figure 8). This could be the result of health workers either using too much blood on the test or of reading the test result too early. In either case, the dark red background could easily obscure a weak positive result and the test should be discarded. Instead, health workers were incorrectly reporting these tests as valid, usually negative, diagnoses. This problem occurred in a total of 58 tests (3.2%), of which 34 were from Tombo clinic. Based on this data, we alerted the relevant supervisor who reviewed the test manufacturer's guidelines with the health workers. In the future, it will be relatively simple to automatically identify tests that have a dark red background and instruct the user to discard the test and run another test.

Q2: Infrastructure Challenges

Our second research question explores the infrastructure challenges faced by participants using the system. None of the devices were broken, lost or stolen during the study. In addition, participants at each study site were able to keep the devices sufficiently charged despite experiencing frequent electricity outages, which confirms the benefits of deploying portable, battery-charged devices.

We observed substantial variations in network connectivity between the study sites. Mutare and Nyanga had fairly reliable 3G connections, Hauna and Zindi had much slower, unreliable 2G connections, and although Tombo appeared to have a 2G connection, the distance from the clinic to the cell tower made transmitting any data extremely challenging. Thus, although we initially planned to transmit relatively high quality images (roughly 300KB), we decided to instead transmit low-quality, compressed images, which decreased the amount of data to about 70KB per test.

Our findings show several interesting differences in data transmission between the sites. Figure 9 shows the proportion of each site's total tests that were transmitted during each week of the study. Mutare and Nyanga were able to reliably transmit data to the server for the duration of the study. By contrast, Tombo was unable to transmit any data for the first two weeks of the study. To overcome this issue, a health worker from Tombo traveled by bus every few weeks to a town roughly 40 minutes away to transmit data. This behavior is illustrated by the high percentages of data transmitted by Tombo in weeks 3 and 6 of the study (see Figure 9). Zindi was also unable to transmit data for the first few weeks of the study. We are unsure what caused the changes in connectivity that enabled Zindi to transmit large amounts of data in weeks 6-8 of the study, although we hypothesize that the network provider installed a new cell tower or signal-boosting hardware in the area. Finally, health workers at Hauna also struggled initially to transmit data. However, after noticing that connectivity seemed to be more reliable at night, health workers began leaving the devices on to transmit overnight. In total, Hauna transmitted 34% of their tests between 10pm and 4am, compared to 4% by Mutare, 2% by Tombo and 0% by Zindi and Nyanga during the same time frame.

Our findings also show large variations between sites in the delays between capturing tests and transmitting them to the server. Mutare and Nyanga experienced the smallest delays in transmission, with geometric means of 33 seconds (SD = 23 hrs) and 45 seconds (SD =17 hrs) respectively, compared to mean delays of 2 hrs (SD = 129 hrs) at Hauna, 2 hrs (SD = 250 hrs) at Zindi and 6 hrs (SD = 140 hrs) at Tombo.

These findings can inform the design of other mobile health systems. Many systems that target low-resource settings face challenges relating to network connectivity, and our study highlights the importance of supporting asynchronous data transmission. Depending on the site, it took anywhere from a few seconds to a few weeks for a captured test to be transmitted to the server, which also highlights the need for mobile systems to process data locally rather than transmit it for off-site analysis. Furthermore, several of the strategies that our participants developed for transmitting data could be generalized to other contexts.

Q3: Comparison of System and Human Diagnoses

Our final research question explores the extent to which the system's diagnoses agreed with the health workers' diagnoses. Although the Ministry of Health informed us prior to the study that the sites used three brands of malaria

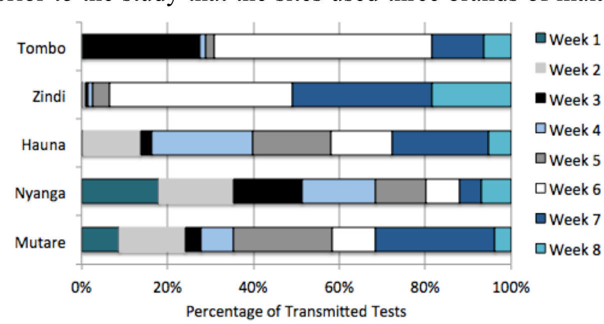

Figure 9: Proportion of each site's total tests that were transmitted during each week of the study.

Table 2: Summary of agreement between (A) the system and visual diagnoses, (B) the review and visual diagnoses and (C) the review and system diagnoses.

A

Visual \ System	Positive	Negative	Kappa (95% CI)
Positive	345	26	
Negative	70	1273	0.84 (0.81-0.87)

B

Visual \ Review	Positive	Negative	Kappa (95% CI)
Positive	352	19	
Negative	22	1321	0.93 (0.91-0.95)

C

System \ Review	Positive	Negative	Kappa (95% CI)
Positive	367	50	
Negative	7	1290	0.91 (0.88-0.93)

tests, we found that in practice all the tests they performed were of a single brand (Paracheck Pf®, see Figure 2). Our analysis of agreement did not include tests that contained user errors (shown in Figure 7). Thus, our dataset consisted of 1714 Paracheck tests. As shown in Table 2(A), the system and visual diagnoses agreed in 1618 tests (94.3%). We used Cohen's Kappa [22] to compute a statistical level of agreement between the two diagnoses. The measured Cohen's Kappa for these results was 0.84 (95% CI:[0.81,0.87]), indicating strong agreement.

Further analysis revealed several sources of disagreement. In some cases, the correct diagnosis was clear and either the system or the health worker made an obvious error. Unfortunately, we cannot determine if visual errors were the result of data entry mistakes or if the participants in fact diagnosed a patient incorrectly. In addition, a large portion of tests (430 or 25%) contained at least some amount of red background coloring, which caused a substantial number of the errors made by the system (often these were false positive results). As discussed previously, the red coloring may result from health workers using too much blood on the test or reading the test result too early. In testing the system prior to the study, we always used the correct amount of blood and waited the right amount of time before reading the result. Thus, we did not test the system on images with colored backgrounds. Further adjustments to the algorithm may be required to deal with this issue. Finally, there were tests with equivocal results at the limit of detection. In these cases, either the system or the health worker reported a positive result. However, since our study did not include microscopy-based analysis of blood samples by a clinician (considered the gold standard of malaria diagnostics), we do not know which diagnosis was correct.

To further analyze differences between the system and visual diagnoses, we created a *review* diagnosis for each test. We asked three researchers to examine each test image and provide a diagnosis, and we recorded their majority opinion as the *review* diagnosis. As shown in Table 2(B), the review and visual diagnoses matched in 1673 tests (97.6%). The measured Cohen's Kappa for these results was 0.93 (95% CI: [0.91, 0.95]), indicating a stronger agreement than observed between the system and visual diagnoses (Table 2(A)). Finally, the review and system diagnoses agreed in 1657 tests (96.7%) (see Table 2(C)). The measured Cohen's Kappa for these results was 0.91 (95% CI: [0.88, 0.93]), which is only slightly lower than observed between the review and visual diagnoses (Table 2(B)).

We were interested to see how our results compared to those obtained in the clinical study of the FIO mobile diagnostic system [19]. Their study was conducted at a single site in Tanzania and analyzed SD Bioline malaria tests. Our study took place at five sites in Zimbabwe and analyzed Paracheck malaria tests. The different test brands used somewhat limits the comparisons that can be made between the two systems, since variations could be due to differences in the tests. In addition, their study was able to compare their system's results to a microscopy-based gold standard, finding that the system had an overall accuracy of 95% (95% CI: [93.7, 96.0]). In our deployment, we were not able to obtain additional blood tests for every patient and perform microscopy-based analysis. Indeed, one reason that rapid diagnostic tests are currently used at our study sites is because microscopy-based analysis is infeasible. Finally, their study focuses solely on the medical outcomes of using the system, whereas our study also discusses findings related to health workers' usage of the system and strategies to overcome infrastructural challenges.

The strength of our findings is highly encouraging and will allow us to explore several future deployment scenarios. For example, if health workers are well trained, the system could simply focus quality control efforts on tests in which the human diagnosis differs from the system diagnosis. Alternatively, in situations where health workers are less experienced, the system could provide a second opinion and instruct health workers to run another test if there is disagreement. Finally, in situations where health workers are untrained, the system could perform the diagnosis and inform the health worker of the outcome.

Deployment Costs

The cost of deploying and sustaining a new technology is an important consideration in low-resource settings. The XCover devices that we deployed cost roughly USD $250 each in Zimbabwe. The devices were loaded with 2GB SD cards that cost $5 each and 300MB data bundles that cost $15 each. The plastic stands to hold the devices cost $10 each. Transmitting a single test required roughly 70KB of data, which equates to a transmission cost of 0.35 cents per test. Storing data from all 15 devices in an online database cost less than $10 per month. Health workers were trained during regular work hours and did not receive additional compensation for participating in the study. Thus, the entire deployment cost approximately $4200, or $280 per device. In fact, the travel, transport and lodging costs for four researchers and two accompanying ministry of health staff were 4 to 5 times more than the costs of the deployment.

Future Work and Limitations
Our study has several limitations that provide opportunities for future work. We only analyze tests for malaria and would like to field test the system with tests for a variety of other diseases. In addition, our participants were well trained and we want to explore the potential for the system to aid health workers that have less medical background. Our system also uses a stand to hold the device in position above the test. Systems that require users to capture images using a handheld device may be more convenient to use, but will likely experience additional image capture issues. Finally, we focus on a camera-based mobile system. Although some of our findings apply to sensor-based systems in general (such as data transmission issues), future research will likely expose additional challenges for systems that use external sensors attached to the device.

CONCLUSION
The availability of mobile devices in low-resource settings has led to the creation of mobile health systems that target disadvantaged populations. This paper presents findings from an eight-week field deployment of a mobile system that captures and analyzes rapid diagnostic tests. Health workers in Zimbabwe used the system for the duration of the study to analyze thousands of tests for malaria. In addition, they employed a variety of strategies to overcome poor network connectivity and transmit test data to a server. Finally, the system's computed diagnoses strongly agreed with the visual diagnoses provided by trained health workers. Taken together, our findings show the potential for mobile systems to aid the delivery of healthcare in low-resource settings and provide valuable insights for HCI researchers working in similar environments.

ACKNOWLEDGEMENTS
This work was funded by Econet Wireless Zimbabwe, NSF Grant 1111433 and a Facebook Graduate Fellowship. We thank the Zimbabwe Ministry of Health and Child Care, Global Solutions for Infectious Diseases (GSID), Dimagi Inc. and all of our study participants.

REFERENCES
1. Boyle D., Hawkins K., Steele M., Singhal M., and Cheng X. (2012) Emerging technologies for point-of-care CD4 T-lymphocyte counting. *Trends in Biotechnology*, 30(1):45-54
2. Chaudhri, R., Vlachos, D., Borriello, G., Israel-Ballard, K., Coutsoudis, A., Reimers, P., and Perin, N. (2013) Decentralized human milk banking with ODK sensors. *Proc. DEV '13* New York: ACM Press p. 4.
3. Dell, N., and Borriello, G. Mobile Tools for Point-of-Care Diagnostics in the Developing World. *Proc. DEV '13.*
4. Dell N., Venkatachalam S., Stevens D., Yager P., and Borriello G. (2011) Towards a Point-of-Care Diagnostic System: Automated Analysis of Immunoassay Test Data on a Cell Phone. *Proc. NSDR '11*.
5. DeRenzi B., Borriello G., Jackson J., Kumar V, Parikh T., Virk P., and Lesh N. (2011). Mobile phone tools for field-based health care workers in low-income countries. *The Mount Sinai Journal of Medicine*, 78(3):406-418.
6. DeRenzi, B., Lesh, N., Parikh, T., Sims, C., Maokla, W., Chemba, M., Hamisi, Y., Hellenberg, D., Mitchell, M. and Borriello, G. (2008) e-IMCI: Improving Pediatric Health Care in Low-Income Countries. *Proc. CHI '08*. pp. 753-762.
7. FIO Corporation. http://www.fio.com/.
8. FrontlineSMS. http://www.frontlinesms.com/.
9. Goel, M., Larson, E., Borriello, G., Heltshe, S., Rosenfeld, M., Patel, S. (2012). SpiroSmart: Using a Microphone to Measure Lung Function on a Mobile Phone. *Proc. Ubicomp '12*.
10. Global Solutions for Infectious Diseases. http://gsid.org/
11. Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W. and Borriello, G. (2010) Open Data Kit: Tools to Build Information Services for Developing Regions. *Proc. ICTD '10.* New York: ACM Press. Article 18.
12. List of known commercially-available antigen-detecting malaria RDTs. http://209.61.208.233/LinkFiles/Malaria_in_the_SEAR_manufat_rdt3.pdf.
13. Martinez A., Phillips S., Whitesides G., and Carrilho E. (2010) Diagnostics for the Developing World: Microfluidic Paper-Based Analytical Devices. *Analytical Chemistry*. 82(1):3-10.
14. Matthews J., Kulkarni R., Gerla, M., and Massey T. (2012) Rapid Dengue and Outbreak Detection with Mobile Systems and Social Networks. *Mobile Netw. and Appl.* 17(2):178-191.
15. Mudanyali O., Dimitrov S., Sikora U., Padmanabhan S., Navruz I., and Ozcan A. (2012) Integrated Rapid-Diagnostic-Test Reader Platform on a Cellphone. *Lab on a Chip* 12(15):2678-86.
16. Pamplona V., Mohan A., Oliveira M., and Raskar R. (2010) NETRA: interactive display for estimating refractive errors and focal range. ACM Trans. on Graphics. 29(4), Article 77.
17. Ramanathan, N., Lukac, M., Ahmed, T., Kar, A., Siva, P., Honles, T., Leong, I., Rehman, I, Schauer, J. & Ramanathan, V. (2011) A cellphone based system for large scale monitoring of black carbon. *Atmospheric Environment*. 45:4481-7.
18. Rennie W., Phetsouvanh R., Lupisan S., Vanisaveth V., Hongvanthong B., Phompida S., Alday P., Fulache M., Lumagui R., Jorgensen P., Bell D., and Harvey, S. (2007) Minimizing human error in malaria rapid diagnosis: clarity of written instructions and health worker performance. *The Royal Society of Tropical Medicine and Hygiene* 10(1):9-18.
19. Shekalaghe, S., Cancino, M., Mavere, C., Juma, O., Mohammed, A., Abdulla, S., & Ferro, S. (2013). Clinical performance of an automated reader in interpreting malaria rapid diagnostic tests in Tanzania. *Malaria Journal*, 12(1):141.
20. Tapley A., Switz N., Reber C., Davis J., Miller C., Matovu J., Worodria W., Huang L., Fletcher D., Cattamanchi A. (2013) Mobile digital fluorescence microscopy for diagnosis of tuberculosis. *Clinical Microbiology*. 51(6):1774-8.
21. World Health Organization, 2013 World Malaria Report. http://www.who.int/malaria/publications/world_malaria_report_2013/report/en/index.html.
22. Cohen, J. (1960). "A coefficient of agreement for nominal scales". *Educ. Psychol. Meas.* 20 (1): 37–46.