Including the Voice of Care Recipients in Community Health Feedback Loops in Rural Kenya

FABIAN OKEKE, Information Science, Cornell Tech, USA
BEATRICE WASUNNA, Medic Mobile, Kenya
MERCY AMULELE, Medic Mobile, Kenya
ISAAC HOLEMAN, Medic Mobile, USA
NICOLA DELL, Information Science, the Jacobs Institute, Cornell Tech, USA

Community health programs in low-resource settings (like rural Kenya) aim to provide essential health services to vulnerable populations. However, to date, there has been limited research that explores the design of mechanisms that enable care recipients to provide feedback regarding their satisfaction with the services they receive. Such feedback has the potential to increase the motivation of community health workers (CHWs), enhance training procedures, detect fraudulent behavior, and inform key performance indicators for health programs. Our paper explores the design and deployment of a USSD-based system that allows anyone who possesses a basic mobile phone to provide feedback regarding the health services and quality of care they received from a CHW or during a hospital visit. Our system was designed through iterative fieldwork in rural Kenya that engaged with multiple stakeholder groups, including care recipients, CHWs, and high-level decision makers. After designing and testing the system, we deployed it for seven weeks in Siaya, Kenya, collecting both quantitative system usage data and qualitative data from six focus groups with 42 participants. Findings from our deployment show that 168 care recipients engaged with the system, submitting 495 reports via USSD. We discuss the broader factors impacting deployment, including the feasibility of USSD, actionability of feedback, scalability, and sustainability. Taken together, our findings suggest that USSD is a promising approach for enabling care recipients to submit feedback in a way that balances privacy, equity, and sustainability.

CCS Concepts: • **Human-centered computing → Field studies**.

Additional Key Words and Phrases: HCI4D; ICTD; mHealth; beneficiary feedback; QA; USSD

ACM Reference Format:
Fabian Okeke, Beatrice Wasunna, Mercy Amulele, Isaac Holeman, and Nicola Dell. 2019. Including the Voice of Care Recipients in Community Health Feedback Loops in Rural Kenya. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 71 (November 2019), 20 pages. https://doi.org/10.1145/3359173

1 INTRODUCTION

Community health programs in low-resource environments (like rural Kenya) provide essential health services to vulnerable populations. Well-functioning community health programs receive input from and pay attention to the needs of multiple groups of stakeholders, including community health workers (CHWs), supervisors, government ministries, NGOs, and, of course, the communities of people who receive care. Unfortunately, prior work has shown that the quality of the community

Authors' addresses: Fabian Okeke, Information Science, Cornell Tech, USA; Beatrice Wasunna, Medic Mobile, Kenya; Mercy Amulele, Medic Mobile, Kenya; Isaac Holeman, Medic Mobile, USA; Nicola Dell, Information Science, the Jacobs Institute, Cornell Tech, USA.

2573-0142/2019/11-ART71 $15.00
https://doi.org/10.1145/3359173

health services that are delivered is not always on par with what care recipients should receive. For example, a study at a referral hospital in Tanzania found that 70% of women experienced disrespect or abuse during labor and delivery [45]. To try and address such problems, health programs are interested in collecting data that might promote accountability, transparency, and equity [21]. A first step in this direction is to design new tools that encourage individual care recipients within target communities to voice their opinions about the health services they receive.

The CSCW community has a growing interest in understanding the complex socio-technical systems that impact communities in the Global South [12, 51, 56] and the design of technologies that engage multiple stakeholders in these communities [34, 46, 53, 56]. In the community health literature specifically, prior research has looked at how to gather feedback from CHWs, supervisors, and high-level decision makers [12, 13, 32], but there remains a need for research that examines how to collect feedback from *care recipients*. Recent research reinforces this need by discussing the potential benefits and opportunities for new feedback systems that target care recipients [35].

Our work begins to fill this gap in the community health literature by designing a new system that enables community members in rural Kenya who possess only a basic mobile phone to submit feedback about care received during a CHW or hospital visit. The system, which is available in three languages (English, Kiswahili, and Dholuo), was designed through an iterative, stakeholder-engaged approach that included the opinions of care recipients, CHWs, and high-level decision makers.

We implemented the system using USSD, a universal communication channel available on any mobile phone (e.g. users dial *144# to check their airtime balance on the Kenyan network Safaricom). Although USSD is universally available and not a new technology, it has been surprisingly overlooked as a potential channel for engaging underserved communities in the Global South [38]. To date, USSD has primarily been utilized only for carrier service requests and mobile money transactions [40]. Our research expands the limited literature on USSD to a new context and, to the best of our knowledge, is the first to apply USSD to the domain of community health.

After designing the system, we deployed it for seven weeks in Siaya, a rural county in western Kenya, collecting quantitative system usage data as well as qualitative data from six focus groups with 42 participants. Our findings show that 168 care recipients engaged with the system during our deployment, submitting 495 reports. Most of the feedback received was positive and submitted in Dholuo, the local language. We show how sending SMS reminders to care recipients triggered engagement, and highlight a range of socio-technical factors that impact our system, including the importance of preserving user privacy and promoting equity by making the system free to use. Finally, we discuss key themes to address as we move forward: the feasibility of USSD as a mechanism for collecting feedback from communities in low-resource settings, the actionability of such feedback, how we might scale the system, and challenges impacting sustainability. To summarize, our contributions to the CSCW community are:

- We describe the design and deployment of a mobile phone-based system that collects feedback from *care recipients* in community health programs in the Global South. In doing so, we address an important gap in the community health literature, which has thus far focused on collecting feedback from CHWs, supervisors, and decision makers.
- We expand prior literature on USSD [38] by being the first to apply USSD to the domain of community health. Our findings suggest that USSD is a feasible mechanism for collecting health-related feedback from rural communities in the Global South.
- We discuss a diverse range of socio-technical factors that impacted the design and deployment of our feedback system, revealing important tensions that arose as we strove to create a system that balances privacy, equity, and sustainability. These insights could inform how researchers and practitioners adopt and use USSD in low-resource contexts.

2 RELATED WORK

The design of tools to collect feedback within socio-technical systems is a popular research topic in CSCW and HCI. For example, prior work has examined the role of feedback in the so-called sharing economy, including AirBnB, eBay, and Uber [19, 25, 27, 29]. A survey of online services that rely on trust and reputation found that reviews and ratings play a crucial role in decision making [25]. For example, AirBnB hosts in Cuba felt obligated to pretend to enjoy conversations with guests because *"the ratings are so important"* [34]. On Uber, drivers undergo considerable emotional labor to get good reviews from passengers because low ratings could kick a driver off the platform, while passengers who have low ratings could be refused rides [19]. Thus, feedback in the form of ratings and reviews have real consequences for both service providers and clients.

In the health domain, researchers have explored technologies that solicit feedback on the quality of healthcare [3, 6, 16, 17, 23]. Dow et al. [16] designed a platform for care organizations to collect and respond to feedback from users. Feedback collected in a four-month deployment of the system with four not-for-profit organizations revealed a mismatch in the values of organizations and their everyday practices. Clement et al. [6] discovered that integrating features for user feedback in a mobile app for patients with lower back pain contributed to frequency of app use. However, this prior work has primarily focused on Western contexts, not communities in the Global South.

The CSCW community's interest in how technology impacts the work and lives of marginalized communities around the world has grown in recent years [10, 33, 46, 53, 56]. Researchers have investigated how digital technologies improve transparency and accountability in low-resource settings via good governance [31, 41, 47]. For example, grievance redressal systems have been deployed as part of government accountability and transparency initiatives to collect citizen responses through telecenters [41], web portals [31, 42] and IVR systems [5, 47]. In the health sector, a recent review identified that many good governance interventions have focused on digital tools that collect information and promote transparency in community programs [21].

In community health programs specifically, studies have designed tools to gather feedback from CHWs. Molapo et al. found that CHWs in Lesotho who were equipped with a mobile app reported more experiences and challenges compared to the old approach of face-to-face supervisor check-ins [32]. DeRenzi et al. used voice and web-based feedback to engage CHWs in India, which led to over 20% increase in the number of client visits they performed [12, 14]. Using personalized performance dashboards during face-to-face supervision with CHWs in Mali increased household visits while still maintaining the quality and speed of care provided [55]. However, these prior studies have focused primarily on CHWs. There is a need for research that studies how to include *care recipients* in community health feedback loops, which is a gap that our research begins to fill. One notable exception is the MomConnect initiative in South Africa that enables pregnant women and new mothers to submit compliments and complaints about the program via SMS [1].

A wealth of literature has examined the design of tools that use only the functionality provided by basic mobile phones (i.e., not smartphones), such as voice calls and short message service (SMS) [13, 37, 52]. However, one universal communication channel that remains relatively unexplored is USSD (Unstructured Supplementary Service Data). Also called short-codes or quick-codes, USSD is commonly used for carrier service requests (e.g. users dial *144# to check an airtime balance with Safaricom in Kenya) and mobile money transactions (e.g. Telenor's Easypaisa service in Pakistan). USSD has a number of advantages over SMS. For example, USSD preserves user privacy as interactions leave no visible traces on the device, and its user interface dialog visually lets a user know when an interaction begins and ends. Also, in contrast to SMS, which is an asynchronous communication channel, USSD supports stateful, synchronous communication, and its interactions happen in real time [38]. Researchers have studied USSD in mobile money interventions [40] and

information collection [54]. However, prior work on USSD is very limited compared to SMS and voice-based systems. Our works expands this literature by exploring a USSD system that collects feedback from care recipients about community health services in rural Kenya.

3 SYSTEM DESIGN

The goal of our research is to explore the design of a system that enables individual care recipients, or beneficiaries, to provide feedback on the community health services they receive. Although the target users for the system are beneficiaries, our design process involved multiple stakeholder groups including beneficiaries, CHWs, supervisors, decision makers, and more. Iterating with multiple groups of stakeholders nurtured a spirit of community collaboration as each group learned the other groups had shared their views on the importance of a tool for care recipients.

We worked in Siaya, a rural area in southwest Kenya with a population of about 850,000 people. Communities in Siaya primarily receive health services in two ways: (1) when they visit a health facility; and (2) when CHWs conduct household visits (e.g. to check on a pregnant woman or provide medication). Our fieldwork team consisted of two women from East Africa and one man from West Africa who all had experience working in Africa. Two team members were very familiar with Siaya: one spent the previous year in Siaya cultivating relationships with stakeholders; the other had previously worked in Siaya. All research activities were approved by our university's IRB and by local Ministry of Health authorities (Director of Siaya County and other County leadership).

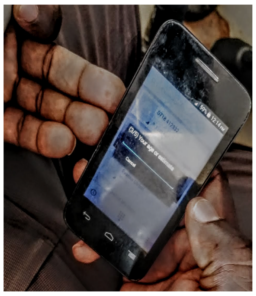
3.1 Design Goals and Challenges

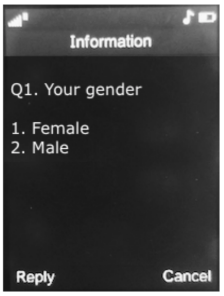
Our design goals focused on creating a tool that takes into consideration the needs of all stakeholders, including care recipients and decision-makers (e.g., government or organization leadership). We focused on two contexts for care recipients to provide feedback: (i) during a hospital visit; and (ii) after receiving care from a CHW who visited a household. Throughout the paper, we refer to the avenue for providing feedback about visits to health facilities as the *Hospital Line* and for CHW household visits as the *Household Line*. The initial goals and challenges described below were informed by prior research on community health feedback tools aimed at care recipients [35], design recommendations from experienced stakeholders who worked closely with our target users, and our prior work on cultivating a human-centered design process [22].

Access and Equity. Providing access is a critical part of designing for low-resource settings. Many technical tools for underserved communities build on universal communication channels available on any basic phone, including phone calls [52], SMS [37] and USSD [38]. Communities are often familiar with these channels, which reduces the training required to deploy applications. Users also do not have to install custom software to use tools implemented via these channels. Our work uses one such universal channel: USSD. It is also important that our tool work in areas with relatively poor connectivity. Similar to phone calls and SMS, USSD does not require Internet connectivity to work, thereby making our tool available to any mobile phone user in Kenya with cellular coverage.

Promoting equity works in tandem with providing access. For example, many tools deployed in prior research in low-resource settings provide free access, since charging users discourages engagement [51]. In addition, prior work recommended that feedback tools for underserved communities should not be limited to only a few community members because the feedback will not be representative of the community and may further marginalize vulnerable populations [35]. We chose to make our feedback tool free to use so it is accessible to as many users as possible.

Privacy and Anonymity. Prior work on CHW programs suggests that feedback tools need to preserve the privacy of care recipients to protect them from potential backlash by CHWs or others who may not be happy with the feedback [35]. In addition, research has shown that providing users





Hospital Line (10 Hops)	Household Line (8 Hops)
1. USSD initialization (dial in)	1. USSD initialization (dial in)
2. Choose language	2. Choose language
3. Introduction – continue/exit	3. Introduction – continue/exit
4. Your gender (Q1)	4. Your gender (Q1)
5. Your age (Q2)	5. Your age (Q2)
6. Select the hospital (Q3)	6. NPS question (Q3)
7. Select the department (Q4)	7. Your reason (Q4)
8. NPS question (Q5)	8. End: Thank you
9. Your reason (Q6)	
10. End: Thank you	

Fig. 1. On the left, a user dials into the system and answers a question about gender as shown in the next image. Completing all steps in the *Hospital* line involves 10 hops while the *Household* line involves eight hops.

with anonymity can increase self-disclosure and empower people to better express themselves [2, 28]. These design recommendations reinforced our decision to use USSD instead of creating an SMS or phone-call based system, since both SMS logs and call logs would reveal a person's usage of the system. By contrast, USSD interactions immediately disappear the moment a user completes their interaction with the system and leave no visible traces behind.

Generalizability. Another key design goal is collecting feedback about health services in a way that is useful to decision makers in community health programs. We sought to use standardized feedback metrics that fit our research context in Siaya while also serving as a relevant metric on a global scale, so that the system might be generalized to other contexts relatively easily. To achieve this, we chose to use the Net Promoter Score (NPS) metric [43] (*On a scale of 0 to 10, how likely is it that you would recommend [company, product, or service] to your friends or colleagues?*). This is widely used across global industries [44] and has been adapted into nonprofit settings [9, 20, 26].

Sustainability and Affordability. Another important design goal was to create a tool that decision makers could afford and potentially sustain over a long period of time. In Kenya, deploying a free USSD-based system is about two times the cost of creating an interactive SMS-based system for collecting feedback (i.e. about $500 more per year). However, deploying a phone-call based system is over 10 times more expensive when compared to a USSD system. Moreover, setting up an IVR system requires multi-weeks steps with official regulatory authorities in Kenya unlike SMS and USSD systems that could be approved in a few days. These financial projections further reinforced our decision to use USSD to create a feedback tool in place of adopting a phone-call based system.

3.2 Implementation Details

Based on these design goals and prior work [35], we created an initial USSD prototype that used Africa's Talking [49], a third party service provider based in Kenya. The system was made available to all users on Safaricom, the most dominant mobile network in Kenya [8]. We implemented the system as open-source using Django web framework with an API endpoint for receiving USSD requests from users and a Postgres database deployed on Amazon Web Services. When a user dials the USSD code, it is converted to a network request by Safaricom, which redirects the request to Africa's Talking, who forwards the request to our web application. The web app generates a formatted response that is sent back to Africa's Talking, forwarded to Safaricom, and rendered on the user's mobile phone. This process happens for each question until the network forwards an "END" signal that originates from the web app after the last question has been answered.

3.3 Iterative Design through Fieldwork

After creating an initial prototype, we conducted iterative design with five stakeholder groups: care recipients, CHWs, supervisors, decision makers (county and sub-county leadership), and professionals at Medic Mobile (an NGO we worked with). In total, we did 7 rounds of iteration with 45 participants: 19 care recipients, 7 CHWs, 3 supervisors, 4 hospital leaders, and 12 NGO staff.

We began by explaining the goal of the system. Then we asked participants to tell us what mobile networks they used on their device, if they knew how to check the airtime balance on their devices, and how they did it (checking airtime is one of the common use cases of USSD). This gave us a sense of how many participants were familiar with USSD and could test on their personal devices. We also bought a basic phone with an active Safaricom number in case participants did not have a mobile device or a Safaricom line. Two of the 45 participants in our design process did not have a mobile device (both care recipients). Everyone who did have a phone also had a Safaricom line.

We asked participants (except care recipients) to imagine a scenario from the perspective of a care recipient. For example, for hospital feedback, participants imagined completing a hospital visit before dialing the USSD code. After completing the scenario on providing feedback, participants shared how they felt about using the USSD system, the questions they had answered, and challenges they encountered or they anticipated care recipients would face. We also followed up on any observations we had that they did not bring up (e.g. getting stuck on a question). For decision makers, we asked the ways in which the system could be integrated into existing workflows at the hospitals. We used a similar procedure to test the system for collecting feedback about CHW home visits, asking participants to imagine themselves as care recipients who had been visited by a CHW at their home. Figure 1 shows examples of USSD questions and a user submitting feedback.

We tested with 14 care recipients at hospital facilities and five at their homes. At the hospital, we approached people at the exit gates and waiting areas, spontaneously asking them if they were willing to spend a few minutes to test a system. If they agreed, we shared the goal of the project and asked those who had their mobile devices to test the system by dialing in using their own phones. If they preferred, we provided them with a basic phone that we carried with us. We followed a similar procedure for testing with people at their homes.

After each round of testing, the research team met and combined notes. We also analyzed the system logs during and after testing to understand the interactions that occurred (e.g. how long users spent on each screen). Based on these insights, we decided on any changes that should be made before further testing. We now discuss lessons learned through our design iterations.

3.4 Findings and Lessons from Design Iterations

Navigate mobile network constraints. A recurring theme in our fieldwork was that users who spent more than one minute on a USSD screen (single question) received a timeout error that was automatically generated by the mobile network, which required the user to restart the submission from scratch. For example, in an early version of the system, participants frequently timed out before they could type a response to *"Any other comment?"* To combat this issue, we changed our design so that all questions only required numeric input (rather than text). We also ideated about how we might inform users at the beginning of a USSD session that questions should be answered as quickly as possible, but ultimately decided against this since it might compromise the quality or accuracy of feedback submitted. Interestingly, our design iterations did reveal that many users were familiar with timeout errors, and usually redialed into the system when they happened.

Another constraint was that each USSD screen could have at most 154 characters (including invisible whitespace). Any text exceeding this length led to a confusing user experience where the extra characters were split into a next screen that could only be viewed by typing '98' or '0'

to go back. As a result, we worked to design each of our USSD screens to have fewer than 154 characters, which made it challenging to balance creating questions that were long enough to not lose meaning after translations yet short enough to not be automatically split into other screens.

Shorten codes. Our prototype used the codes **384*11100#* for the *Hospital Line* to collect feedback about health facilities and **384*99900#* as the *Household Line* for feedback about CHWs who visited care recipients in their homes. We used two different codes to separate responses for the two contexts. However, testing with stakeholders quickly yielded suggestions that we shorten the codes since community members are used to dialing shorter codes (e.g. **144#* to check airtime balance on Safaricom). Thus, we worked with the USSD service provider to find shorter codes, specifically **384*777#* for the *Hospital Line* and **384*888#* for the *Household Line*. We were unable to make the codes shorter because (i) setting up a 3-digit code is over 10 times more expensive than setting up a 6-digit code; (ii) it takes multiple weeks to get approval from the telecom networks for 3-digit codes compared to a few hours for 6-digit codes; and (iii) it is non-trivial to find an available 3-digit code that will not be confused with other USSD services in Kenya.

Balance stakeholder suggestions. We frequently received conflicting suggestions from stakeholders. For example, CHWs were interested in knowing which community members submitted feedback about them, conflicting with care recipients desire for anonymity. As another example, decision makers often asked us to include lengthy and complex questionnaires, which would have been very cumbersome via USSD. Although we did not incorporate every stakeholder suggestion, we did manage to include some suggestions, such as adding two questions to the *Hospital Line* that asked care recipients to specify the specific facility and health department they visited.

Adapt standardized questions. Feedback from CHWs and care recipients revealed that the initial feedback question based on the Net Promoter Score [9] was challenging for users to fully comprehend. For example, asking *"How likely are you to recommend this hospital to someone else? Use a scale of 0 (highly unlikely) to 10 (highly likely)"* resulted in users selecting only 0 or 10 because they did not understand what the numbers in between stood for. To address this problem, we used cultural analogies such as "*think of it as marks you earn in school and give a score where 0/10 means bad and 10/10 means excellent"*. However, users said that a *"score"* made them think of a soccer match and not a hospital or a CHW. After many iterations, we eventually changed *"the scale of 0 to 10"* to three options: *"definitely yes"*, *"maybe"*, and *"definitely no"*, which was well received.

Integrate language preferences. During our fieldwork, we asked participants their preferred language and received three answers: English, Kiswahili, and a local language in Siaya (Dholuo). There was no consensus, with different individuals preferring different languages. For example, some explained that they did not want the system in Kiswahili because *"it is not our mother tongue, we don't speak it."* To account for these varying preferences, we redesigned the system so that questions could be rendered in any of the three languages, but without any single question in any language exceeding 154 characters in length. This process was quite challenging, with translations of English words often being longer in the other two languages. In addition, Dholuo had varying dialects, so substantial effort was put into using words that were well known in the local dialect. We validated our translation with people from different communities to find the most common terms and tested the translations with CHWs and care recipients during our iterations.

4 FIELD DEPLOYMENT

After converging on our final design, we deployed the tool for seven weeks in Siaya. We trained five CHWs to use the tool and they in turn trained community members to provide feedback in two contexts: at the hospital and during home visits. All CHWs trained care recipients during household

visits, while only one CHW trained care recipients at a health facility. Each CHW reported being responsible for 50-100 households, making between three and 22 household visits per week. After two weeks, we held focus groups with stakeholders to understand their experiences using the system and adoption challenges. We now provide details of our deployment.

Procedure. We initiated our deployment by training five CHWs to use the system, and they in turn trained community members. We limited the training to only five CHWs because we wanted to get a sense of how the system might work at a relatively small scale and address any potential issues that arose before scaling to a larger group. We started by asking CHWs if they knew how to check their airtime balance, and everyone indicated that they were proficient with doing this. Then we discussed the potential to improve service delivery by collecting feedback, and introduced our USSD tool as a mechanism for care recipients to become connected to the health feedback loop. We then observed CHWs as they used the USSD system several times for each of the two codes: for hospital feedback and household feedback, and answered any questions or issues that came up. Finally, we asked CHWs if they felt comfortable teaching care recipients to use the tool, and ensured that they emphasize that it is free to use. During training, we were careful to explain that any feedback collected would not affect the CHWs' jobs and could not be traced back to them.

After training, the CHWs were free to suggest the system to any care recipients that they interacted with. All CHWs told us they advertised the system to households they visited, and one advertised it at a local hospital. We are aware that asking CHWs to recruit people to submit feedback regarding the CHWs' own work may lead to bias [11], such as CHWs telling people to submit positive feedback. We ideated on ways to try and mitigate such bias. For example, we focused on training CHWs to (1) teach beneficiaries to use the USSD system on their own, and (2) emphasize that beneficiaries could submit feedback at any time (such as after the CHW departed).

To further encourage participation, we also sent SMS reminders to care recipients. To do this, we collected the phone numbers of care recipients that CHWs visited and asked the CHWs to inform these care recipients that they would, after the home visit, receive an SMS reminder encouraging them to provide feedback about their CHW visit. An example of an SMS reminder is, *"Hello, you were recently visited by a CHW. Please dial *384*888# to provide feedback about the visit. This service is free. Thank you".* We collected roughly 20 new phone numbers every two weeks, sending a total of 80 SMS reminders during the deployment to care recipients who had received a CHW visit.

Focus Groups. Two weeks into the deployment, we conducted six focus groups with 42 participants: 24 care recipients, five CHWs, three supervisors, and 10 hospital decision makers, to collect qualitative feedback about the deployment. Two CHWs, all supervisors and one decision maker had also taken part in the design phase; however, the remaining participants were new to the study. We recruited participants through supervisors who reached out to their CHWs and CHWs in turn reached out to their care recipients. In addition, one of the research team members who resides in Siaya invited decision makers to partake in our study. Then all interested participants who came forward participated in different focus groups. Focus groups lasted for about an hour for care recipients and 30 minutes for other stakeholders and questions were tailored to each focus group. For care recipients, we aimed to understand their experiences and challenges of using the USSD system; for CHWs, we focused on how they perceived the feedback and their experience informing care recipients about the system; and for decision makers, we inquired about the role and impact of feedback on community health programs. All participants except decision makers were compensated $5.00 to cover their transportation to the focus group. Table 1 shows the demographic details of participants. All care recipients we interacted with were literate and had their own devices. Only two people said they shared their phones. Focus groups were conducted in English and Dholuo by three researchers. In total, we had five hours of focus group discussions.

42 participants	**Beneficiaries**: 24, **CHWs**: 5, **Supervisors**: 3, **Decision Makers**: 10	
Age	**Beneficiaries**: Min: 20, Max: 49, Avg: 31; **CHW**: Min: 33, Max: 46, Avg: 40	**Supervisors**: Min: 30, Max: 40, Avg: 35 **Decision Makers**: Min: 30, Max: 60, Avg: 40
Gender	**Beneficiaries**: Female: 21, Male: 3; **CHW**: Female: 3, Male: 2	**Supervisors**: Male: 3 **Decision Makers**: Female: 6, Male: 4
Education	**Beneficiaries**: form two - diploma; **CHW**: form two - college	**Supervisors**: diploma - masters **Decision Makers**: diploma - medical degree
Phone	**Beneficiaries**: basic phone: 17, smartphone: 6, shared: 2; **Other participants**: smartphones	
Beneficiary Occupation	Farmer: 10, trader: 4, house wife: 5, teacher: 2, business owner: 1, tailor: 1	

Table 1. Demographic characteristics for focus group participants.

Data Collection and Analysis. Our qualitative data consisted of focus group discussions and 22 pages of notes that we took during our fieldwork. We audio-recorded and transcribed our focus groups. We then performed thematic analysis [48] on the transcripts and field notes, beginning with a close reading of the transcripts and allowing codes to emerge from the data. Multiple passes through the data resulted in 39 codes (e.g., *bring about change* and *concerns about negative feedback*). We clustered related codes into high-level themes (e.g., *promoting equity* and *sustainability*) and organized them in a codebook. After multiple discussions and iteratively refining the codes, we arrived at a final set of themes that comprehensively represents the data.

Our quantitative data consisted of system usage logs recording when and how often users dialed into the system. For each user we recorded the choices selected, any errors made (and types of errors), time spent on each screen, languages chosen, types of feedback submitted, and timestamps of all events. All data logged during training sessions were removed from the analysis. We also recorded when SMS reminders were sent and when users responded, if they did.

5 RESULTS

We now discuss our deployment results as they relate to two key themes: (1) how the system was used by care recipients, including the kinds of feedback submitted, the effects of training, and the impact of SMS reminders; and (2) the socio-technical factors impacting our deployment.

5.1 System Usage

Overview. During the deployment, we observed varying levels of engagement with the *Hospital Line* and the *Household Line*. In total, we recorded 495 sessions and 2602 hops. A session is defined as a period from when a user initiated a code to when it ended, while a hop is a USSD request that shows the user a single question. Over the deployment, 168 unique phone numbers dialed into the system (79 into the *Hospital Line*, 124 into the *Household Line*, and 35 into both lines). Figure 2 shows the number of times users dialed into the system. Most users (n=64) dialed in only once. The median number of dial-ins was two times, and the maximum was 17 times.

It is important to note that users who dialed into the system did not necessarily complete the entire survey. Indeed, Figure 2, which shows the number of hops completed by each user who initiated a session, reveals that the number of submissions decreases as the number of hops increases. The tendency of completion rates to decrease as questionnaire length increases is well-documented in the literature (e.g., [18]). In our case, the dropoff may be due to several factors, including USSD timeout errors or people running out of time, losing interest, or dialing in out of curiosity. Nevertheless, a benefit of USSD is that the system records data from partially completed

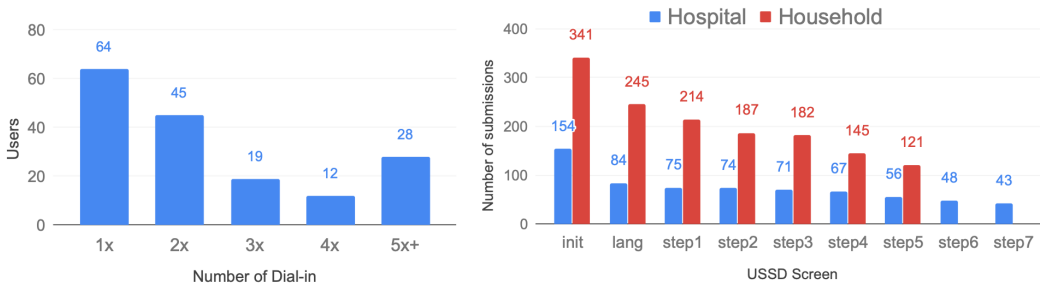


Fig. 2. **Left**: Frequency of individual user dial-in (median = 2). **Right**: Submissions for each USSD screen. The user starts from initializing the code (init) to selecting a language (lang) and then going through seven or five steps for the hospital and household lines respectively.

USSD Line	Sessions	Phones	Women	Men	Avg Age	Min Age	Max Age
Hospital	154	79	92	34	30.8	19	70
CHW	341	124	179	59	33.3	18	70
Total	495	168*	271	93	32.1	18	70

Table 2. Details of USSD submissions: *"Sessions"* refer to each unique time that a user dials in; *"Phones"* refers to the unique phone numbers that dialed in (some phone numbers (n=35) submitted to both the Hospital and Household lines). *"Women"* and *"Men"* show gender indicated on submissions; Age is shown in years.

surveys (any completed hops). As a result, our total data set is larger for some questions than others.

The decrease in question completion rates also suggests it is advantageous to place the most important questions first. However, in our design, we asked simple, demographic questions first (e.g. gender, age) so that people would find the system easy to use and feel encouraged. This meant that the important feedback rating question was positioned towards the end of the survey (see Figure 1) and completed by fewer people. We plan to correct this in future design iterations.

Error rate. Our system was designed to minimize erroneous user entries. Whenever a user entered an invalid input, they received a message that gave them the range of values allowed. For example, when a user is asked *"Your age"* and enters non-numeric text, the system prompts: *"Please enter a number between 18 to 99"*. This prompt persists until the user enters a valid value. Surprisingly, the error rate in the system was negligible: 0.6% (16 out of 2602 hops) from eight phone numbers.

All errors were of two kinds: (i) entering text when only numeric responses are allowed; and (ii) submitting numeric data outside the input range allowed. The overall low error rate is in part due to: (i) familiarity with USSD—participants said they were familiar with USSD; and (ii) understandable questions—participants said that the questions were straightforward and easy to understand.

The kinds of feedback received. We received a total of 193 responses to the Net Promote Score (NPS) question: 135 (69.9%) positive, 33 (17.1%) negative, and 25 (13.0%) neutral (see Fig. 3). Most feedback (145 submissions) came from households. We computed the NPS [9] for both lines (% positive feedback - % negative feedback). For hospital feedback, 58.3% (n=28) were positive and 25% (n=12) were negative, yielding an NPS score of 33.3. For household feedback, 73.8% (n=107) were positive and 14.5% (n=21) were negative, yielding an NPS score of 59.3. Thus, household feedback was more positive than hospital feedback. There are several possible explanations for this. For

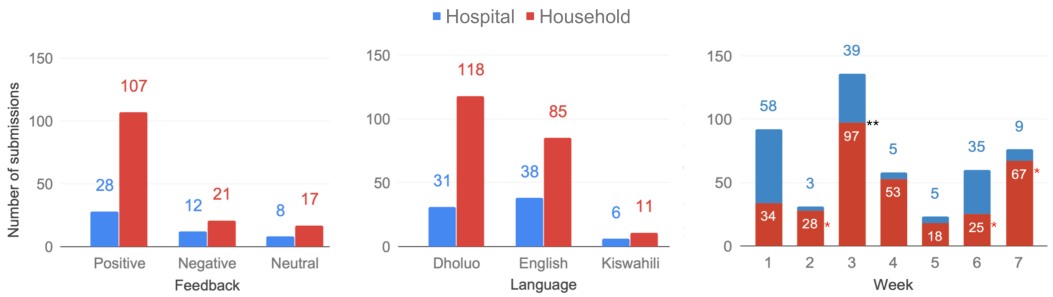


Fig. 3. (Left) The kinds of feedback submitted; (Middle) Language selection; (Right) Submissions per week. SMS reminders were sent to households in weeks two, six and seven (*). In week three (**), we conducted focus groups, and observed a substantial increase in submissions on both the *Hospital* and *Household* lines.

example, for household visits, CHWs make the effort to travel to beneficiaries' homes, which is easier and more convenient for beneficiaries than traveling to and waiting in line at the hospital. Thus, a household visit may actually be a more positive experience than a hospital visit. In addition, CHWs often have personal relationships with beneficiaries, which is not necessarily the case for staff at hospitals. Finally, household visits may have been more prone to response bias [11] since CHWs were soliciting feedback about their own services.

Training leads to engagement. Recall that CHWs were trained to guide care recipients on how to submit feedback via the USSD system. Our findings show that CHWs were comfortable training care recipients. On the first day of deployment, our team accompanied a CHW to the laboratory department of a hospital and observed from a distance as the CHW spoke to a group of hospital attendees about how they could provide feedback after their hospital visit. Within three hours, 16 users had dialed into the system and made 20 submissions, demonstrating that CHWs may be a good channel through which to reach care recipients. All CHWs said that it only took a few minutes for them to explain to care recipients how to use the system. In addition, they emphasized that care recipients could submit feedback at any time. One CHW shared,

> *"Training care recipients was OK. I did it and it was easy for me and my households. But there was a challenge. They asked me 'Are we going to be given so many questions to answer and questions that we don't know how to answer?'. So I just tried to tell them 'you'll be given questions that you will be able to answer.'"* (P27, Female, CHW)

In addition, care recipients shared that after they were informed about the system, they were able to dial into the system by themselves and submit feedback because they were familiar with how to use USSD. However, CHWs also said that, occasionally, they visited care recipients who did not possess mobile devices. In these cases, they did not mention the feedback system. As a result, training only occurred when CHWs felt it was appropriate. One CHW shared,

> *"There are some households where you find that the care recipient is an old mama and she doesn't have a phone so you find that there is no need training her about the USSD tool or her submitting feedback." (P25, Male, CHW)*

SMS reminders renew engagement. We sent SMSs to household care recipients reminding them to submit feedback about recent CHW visits. We sent a total of 80 SMS reminders during the deployment. SMSs were sent on weekdays, between 10am and 4pm, the week after care recipients had been visited by a CHW. Of the 80 SMSs sent, 75 were successfully delivered, while five failed because of network issues or the numbers being out of service. Figure 3 shows that, on average,

system usage increased after SMS reminders were sent. We received feedback submissions from 20 out of the 75 people who received an SMS reminder (27%). The fastest response was received 3 mins after the SMS was delivered, and the slowest 21 hours after the SMS was delivered (average: 2.5 hours, median: one hour). However, not all SMS reminders triggered engagement for a variety of reasons, such as the SMS not being delivered or care recipients not wanting to submit feedback, perhaps because they did not have time or felt indifferent about the services received.

5.2 Socio-Technical Factors Impacting Deployment

Having described how care recipients engaged with the system, we now discuss socio-technical factors that affected our deployment. We uncovered five such factors: (1) trust and accountability, (2) privacy and anonymity, (3) equity (4) leadership buy-in, and (5) feedback granularity.

Trust and Accountability. Our findings showed that feedback from care recipients fostered a new layer of social interactions in the community, helping to promote accountability and build trust in the way health services are delivered. All care recipients and CHWs expressed that the availability of the *Hospital Line* and the *Household Line* served, in combination, as an empowerment platform for care recipients to have their voices heard in ways that were not possible before. For example, participants described how, previously, when they were dissatisfied with services at hospital facilities, they bottled up their complaints because they did not have a way to share their experiences. Participants were now happy to have a way to report feedback. One said,

> *"The majority are happy with the USSD tool and they accept it. They are really happy with it, more so in the hospital. They really appreciated that because they normally face people who have got many characters. For example, some people mistreat care recipients in the hospital laboratory department. Or at another department they normally take long. But if there's something that is being done in the community such that the client reports it and it can be followed up, then at least the people responsible are going to change."* (P28, Female, CHW)

Hospital decision makers explained that they wanted to collect client feedback to gain insights into how people were treated across hospital departments. They thought that although some clients might have poor experiences at hospitals, these could perhaps be attributed to one or a few specific departments. Feedback might help to identify such departments and hold them accountable.

With respect to feedback about CHWs at the household level, many participants (n=13) expressed that health was an important matter that required attention. As such, feedback might be useful in discouraging CHWs from reneging on their responsibilities. One participant said,

> *"[CHWs] have to do a good job because if they don't, they know we shall be submitting feedback about it…Since the work of CHWs are issues that are concerned with health, it is important they do a good job. With good health my life is okay, so I should be there to speak the truth about their work."* (P17, Male, Care recipient)

However, CHWs worried that they might receive unwarranted negative feedback from malicious users in their communities. Some participants (n=7) shared that even though CHWs worked hard, they sometimes encountered difficult community members who did not value their contributions. CHWs worried that such individuals could easily use the system to leave negative reviews. A CHW said, *"It's like we are selling what is maybe going to kill us."* (P25, Male, CHW). We assured CHWs that our research was exploratory and made clear that any feedback received during our study would not affect their jobs. However, these tensions raise open questions for future research.

All CHWs (n=5) also felt that, although positive feedback encourages them, genuine negative feedback could help them know how to improve their work. They described wanting to be able to personally review all the feedback received about them. One CHW shared,

> *"It's now upon me as a CHW to take all the positive and negative criticism and work on them either to scale up my work or to improve on what I have been doing. Even when we were telling our fellow CHWs, we told them that it is not a tool that has been introduced to destroy what we have been doing. It's just to monitor and to make us aware of our weaknesses and to help us improve on those weaknesses. One or two people might use it negatively against us but not everybody will do that. I hope it will make us improve on what we are doing."* (P25, Male, CHW)

As discussed above, we received much more positive feedback than negative feedback (see Fig 3), which suggests that care recipients are perhaps more likely to submit positive feedback.

Privacy and Anonymity. Our analysis revealed that privacy played an important role when providing feedback. All participants said they were confident they could provide feedback without someone else knowing they had done so. One participant shared,

> *"When I am done answering the feedback questions, the [dialog on the] screen disappears—it's already gone and no longer on the phone."* (P21, Male, Care recipient)

When a user interacts with the system, the pop-up dialog on their screen is immediately removed when (i) they complete the last USSD question; (ii) they press the cancel button to exit before the last screen; or (iii) the mobile network automatically makes the screen timeout if the user has taken longer than one minute. Once the dialog has been removed there is no visible trace on the user's device that indicates they dialed into the system. As such, USSD automatically makes users' interactions private compared to systems that use SMS (where users would have to delete each SMS to hide their interactions) or phone calls (where users would have to delete their call logs).

Hospital decision makers and a few CHWs (n=3) inquired if we could tell which care recipients submitted feedback. We explained that we could not because the system further protects user privacy at the backend. All feedback received is stored in a database that does not contain any user names. We do, however, store user phone numbers as a way to distinguish return users from new users. As a result, an administrator who looks at the data cannot infer who submitted feedback unless they have a database that connects phone numbers with care recipient names.

Equity. Our design promotes equity by making the system free to use by anyone who dials in. All our participants had access to a personal or shared phone and Safaricom SIM card. Several participants told us that they deliberately checked that the system was indeed free to use. One said,

> *"It is free and I didn't use any money. I know because after dialing into the system, I checked my airtime balance."* (P3, Female, Care recipient)

If a user dials into the system from a network other than Safaricom, they receive an error message saying that the code they dialed is invalid. A few participants (n=5) asked if it was possible for the system to work on all networks. We explained that since we were in the exploratory phase of our work, we limited the system to Safaricom for financial sustainability (discussed more in Section 6).

In addition, participants shared that the opportunity to choose their own language upon dialing into the system made the system more accessible and easier to use. Figure 3 shows that Dholuo was the most preferred language, followed by English and then Kiswahili. The low usage of Kiswahili

was consistent with participants' feelings during our focus groups, where they described how they were not comfortable speaking or reading Kiswahili because it was not their mother tongue.

Leadership buy-in. Our design approach purposefully involved hospital and community leadership from the start. At the beginning of our project, we reached out to the Siaya County Ministry of Health leadership requesting to test the system at one hospital, approaching clients as they exited the hospital. In response, we were invited to test not only at the exit of the hospital but also in the waiting area inside a hospital department. This perhaps indicates these stakeholders' enthusiasm for obtaining client feedback. We kept these stakeholders updated about subsequent iterations of the system, and were subsequently invited to partner and expand the deployment to ten hospital facilities in Siaya. Although the hospital decision makers were familiar with well-known data collection tools such as ODK (Open Data Kit) [4], they were enthusiastic about using USSD for collecting community feedback because it was accessible by community members with basic mobile phones and, unlike ODK, it did not require extensive technical support or training to deploy. We are now in the process of coordinating with stakeholders to expand the system to multiple facilities.

Granularity of feedback. A key finding in our research is that it is complex to collect granular feedback. Although it is relatively straightforward to collect feedback about how communities view CHWs in general, understanding which communities submitted feedback would require users to identify their community (e.g., village) in their submission. However, as discussed in Section 3, we eliminated text-based responses to avoid frequent timeout errors. Alternatively, users could choose from multiple choice options of villages, but there are hundreds of villages and the list would not fit on a single screen. Compounding the challenge of getting granular feedback is accounting for scenarios in which multiple CHWs visit one household. CHWs explained that although each CHW is in charge of a set of households, it was common for a household to receive services from CHWs that work in different health organizations. One CHW explained,

> *"We go across each other's borders. You might find that there's a CHW dealing with a household on a different issue. Maybe there's a different organization that has come and recruited their own CHWs . . . Does the feedback collected apply to all CHWs or only the particular ones in our area? Can the care recipient give feedback since the tool is not particular about who specifically visited the recipient?"* (P25, Male, CHW)

This suggests that separating feedback about multiple CHWs that attend to a household will be challenging if there is no unique identifier that links specific CHWs to the feedback submitted.

6 DISCUSSION

Having analyzed how care recipients engaged with the USSD system and the socio-technical factors that impacted our deployment, we now discuss four key themes to address as we move forward: the feasibility of using USSD to collect feedback from communities, how the feedback may be used, how we might scale the system to larger deployments, and challenges impacting sustainability.

Feasibility of USSD. Our analysis suggests that USSD has a strong prospect of becoming an effective mechanism for collecting community health feedback from care recipients and contributes to the limited literature available on USSD-based systems [38]. As a universal communication channel embedded in all mobile devices, it is accessible on any basic feature phone, and access is not dependent on Internet connectivity. Since our participants in Kenya were already familiar with using USSD for mobile carrier services (e.g., checking their airtime balance), it was easy for them to learn how to apply the same approach in providing community feedback. Surprisingly, our data revealed an extremely low error rate (less than 1%), which corroborates our qualitative findings

that users did not struggle to use the system. This finding contrasts prior research that describes high error rates in SMS-based systems in low-resource contexts [36, 39].

Our stakeholder-engaged approach led to a number of key design decisions that aided deployment. For example, making the system free to use promoted equity by enabling anyone with a basic mobile phone and a Safaricom SIM to submit feedback. Providing the system in multiple languages further aided access; we saw usage in all three languages, with Dholuo (the local language) being most popular. These findings support prior work on the importance of local-language settings [12, 37, 38, 54] and extend this literature by showing the benefits of supporting *multiple* languages.

Although USSD is a promising approach for collecting community feedback, it also has limitations and challenges that could impact adoption. Clearly, it requires users to read the questions on the screen and thus may be inappropriate for people who are illiterate. Beyond this, our analysis revealed technical challenges with adopting USSD. For example, each screen was automatically limited to an interaction timeout of one minute, and the maximum length of a single screen was 154 characters. We accounted for these constraints through careful design, although coming up with a design in which each screen was less than 154 characters regardless of language was challenging.

Actionability of feedback. This study is part of a larger program of work on precision performance management in community health, which aims to establish new, data-driven approaches to supervision and health system leadership. While our study explored a system for collecting beneficiary feedback, in our interviews with health system leaders we discussed possible channels for making this feedback actionable within existing infrastructure. We now share ideas about how such a system might be useful and what actions may be taken based on the feedback.

Prior studies have shown that giving CHWs access to feedback about their work can lead to performance improvements [14, 55]. For example, showing personalized performance feedback to CHWs via a dashboard during face-to-face meetings with their supervisors led to improvements in CHW performance in Mali [55]. However, the feedback shown to CHWs did not include subjective feedback from care recipients. Thus, a concrete use case would be to add to the dashboard aggregate feedback (e.g., ratings) submitted about the CHW. Supervisors and CHWs could then discuss this feedback and any actions to take during their face-to-face meeting. In addition, we worked within a CHW program that already presents CHWs with targets (e.g. targets for facility-based births among women enrolled in antenatal care) built using the open source Community Health Toolkit [7]. Thus, another use case would be to create new CHW targets for care recipient feedback (e.g., number of feedback reports the CHW receives, target quality ratings, etc.) CHWs could then view aggregate feedback submitted by their care recipients and use it to improve their work (e.g., by making more household visits or improving the quality of the visits that they perform).

To make the feedback actionable for community health program leadership, we note that the decision makers we worked with already use the open source health management information system DHIS2 [15]. Using DHIS2, decision makers routinely view a range of dashboards designed for high-level health system management. We anticipate building a care recipient feedback view within these dashboards that would allow decision makers to see aggregate feedback about CHWs and health facilities. By including questions that ask care recipients who submit negative feedback to choose from a list of common reasons (such as *"waiting time is too long"*, *"medications are too expensive"*, *"facility staff are rude"*, etc.), the system could provide decision makers with possible next steps to take based on the feedback (such as more staff training, increasing stocks of pharmaceutical supplies, subsidizing the cost of medications, etc.). Further, decision makers could update the survey questions periodically to seek feedback on specific topics of interest, such as asking *"were you treated with respect"*, *"did you face discrimination"* and so on. New questions would still need to be carefully designed and pilot tested prior to large-scale deployment.

For each of these scenarios, it will be important to carefully consider privacy, who has access to the feedback, and how it might impact CHW and hospital staff employment (e.g., if they could lose their job). This will be particularly important as we explore opportunities to link feedback received to individual CHWs or health facilities. One technique for linking feedback in this way would be to assign unique ID numbers to CHWs and facilities, which beneficiaries would enter into the USSD system. This approach is similar to how mPesa customers submit a unique number into mPesa in order for mPesa-accepting business points to be identified [24]. The IDs for health facilities could be displayed on posters in each facility and distributed to care recipients by hospital staff; while CHWs could share their IDs with care recipients during household visits. Implementing identifiers in this way would open up new opportunities for synthesizing insights across different data sources. For example, the IDs could be used to connect aggregate feedback data with other performance information such as the speed with which CHWs reach sick children. Our next study plans to explore how to assign unique identifiers to CHWs and health facilities that are shared with care recipients as part of a broader aim to incorporate beneficiary feedback into existing infrastructure for precision performance management.

Scalability. We conducted a small-scale deployment, training only five CHWs who trained community members. We did this because we wanted to understand the impact and consequences of the system before proceeding to a large-scale deployment. For instance, CHWs were worried that beneficiaries might maliciously submit negative feedback that would impact their employment, which our small-scale deployment suggests is perhaps not a major concern (at least for now).

Our findings show that sending SMS reminders often prompted care recipients to engage with the system and submit feedback. This finding corroborates prior research that SMS reminders trigger engagement in community health programs [13, 37]. Although manually sending SMS reminders to a few households per week was manageable in our small-scale deployment, this approach would not be suitable at scale. In addition, we sent reminders at the beginning of the week after care recipients received a CHW visit because we did not want to send reminders to care recipients if they had not yet been visited. However, this led to scenarios where care recipients received reminders to submit feedback several days after a household visit.

As we consider scaling the system, we intend to create an automated SMS management system that is integrated with the existing information system that CHWs already use to collect data about clients they visit (in our case, Medic Mobile [30]). This would enable SMS reminders to be automatically sent to care recipients immediately after CHWs complete their household visits and submit the data to Medic Mobile. This would ensure that the visit is still fresh in the care recipient's mind when they provide feedback. A few days after prompting for feedback, the SMS system could check the USSD system to see if care recipients submitted any response and if not, send an additional reminder message. We elected not to integrate an automated SMS reminder system in our initial deployment because we were still unsure about the feasibility of the USSD system.

Finally, we conducted a small-scale deployment with limited participants in one county in Kenya. More research is needed to evaluate the system at a larger scale and in different contexts.

Sustainability. A key part of scaling our USSD system is sustaining the deployment over a long period of time. Sustainability remains a longstanding problem for researchers and practitioners who deploy projects in low-resource settings [37, 51, 52, 54]. Our analysis showed that setting up and maintaining a USSD system in Kenya is significantly cheaper than setting up an IVR system (by a factor of 10) but slightly more expensive than setting up an SMS system (by a factor of two). We estimate the cost of setting up two USSD lines to run for a year to be $1700. This would account for 20,000 free feedback submissions to the USSD system. Based on these figures and in collaboration

with our research partner, Medic Mobile, we plan to deploy the system at a larger scale and in a broader set of communities. In this deployment and beyond, Medic Mobile plans to support the system by integrating it into their existing, widely-used open source software platforms [7, 30].

Beyond financial sustainability, effort will be needed to consistently publicize the system within communities and train users on how to submit feedback. Taylor et al. [50] emphasize that usage issues are a primary challenge in the handover of community health technologies because *"a technology can be given to the community but might not be used"*. As a result, sensitizing communities about the need to provide feedback is important for sustainability. Creating awareness could happen at hospital facilities through ongoing health education activities where community members gather to learn about how to improve their health. Catchy posters could also be put up in strategic places, such as waiting areas and exit gates of hospital facilities. At the household level, we worked with only five CHWs who reached households in Siaya. Effort would be needed to train more CHWs, who in turn could train community members. Communities could also learn about the system through community events and meetings that take place in many villages, word of mouth at household levels, and radio stations. However, these combined efforts to sensitize communities on the need to provide feedback on their health programs would incur additional expense.

7 CONCLUSION

This paper explored the design and deployment of a USSD system that enables care recipients to provide feedback about the health services they receive. Our design approach highlighted the importance of incorporating the views of multiple stakeholder groups. Quantitative and qualitative data from our seven week deployment showed that 168 care recipients engaged with the system and submitted 495 reports via USSD, suggesting that USSD is a viable mechanism for a beneficiary feedback system. We also discussed a range of sociotechnical challenges that arose during our study as we explored the feasibility, equity, privacy, and sustainability of using such a system to provide vulnerable populations with a voice in their own community health programs.

8 ACKNOWLEDGMENTS

We are deeply grateful to our research participants who made the time and effort to give us valuable and detailed responses. We also thank the anonymous reviewers for their helpful comments and suggestions. This work was supported by the Children's Investment Fund Foundation UK. The funder played no role in conducting the research or the decision to publish.

REFERENCES

[1] Peter Barron, Yogan Pillay, Antonio Fernandes, Jane Sebidi, and Rob Allen. 2016. The MomConnect mHealth initiative in South Africa: Early impact on the supply side of MCH services. *Journal of public health policy* 37, 2 (2016), 201–212.

[2] Michael S Bernstein, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich, and Gregory G Vargas. 2011. 4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community.. In *ICWSM*. 50–57.

[3] Gavin Brookes and Paul Baker. 2017. What does patient feedback reveal about the NHS? A mixed methods study of comments posted to the NHS Choices online service. *BMJ open* 7, 4 (2017), e013821.

[4] Waylon Brunette, Mitchell Sundt, Nicola Dell, Rohit Chaudhri, Nathan Breit, and Gaetano Borriello. 2013. Open data kit 2.0: expanding and refining information services for developing regions. In *Proceedings of the 14th workshop on mobile computing systems and applications*. ACM, 10.

[5] Dipanjan Chakraborty and Aaditeshwar Seth. 2015. Building citizen engagement into the implementation of welfare schemes in rural India. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. ACM, 22.

[6] Innocent Clement, Andreas Lorenz, Bernhard Ulm, Anne Plidschun, and Stephan Huber. 2018. Implementing Systematically Collected User Feedback to Increase User Retention in a Mobile App for Self-Management of Low Back Pain: Retrospective Cohort Study. *JMIR mHealth and uHealth* 6, 6 (2018).

[7] Community Health Toolkit. 2019. Retrieved April 4, 2019 from https://communityhealthtoolkit.org/.

[8] Abdi Latif Dahir. 2019. Airtel and Telkom are uniting to take on Safaricom in Kenya. Retrieved March 26, 2019 from https://qz.com/africa/1548281/airtel-and-telkom-are-uniting-to-take-on-safaricom-in-kenya/.
[9] Sophie Sahaf Reem Rahman David Bonbright, Britt Lake and Renee Ho. 2015. Net Promoter Score for the Nonprofit Sector: What WeâĂŹve Learned So Far. Retrieved March 18, 2019 from https://feedbacklabs.org/blog/net-promoter-score-for-the-nonprofit-sector-what-weve-learned-so-far/.
[10] Nicola Dell, Trevor Perrier, Neha Kumar, Mitchell Lee, Rachel Powers, and Gaetano Borriello. 2015. Digital Workflows in Global Development Organizations. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 1659–1669.
[11] Nicola Dell, Vidya Vaidyanathan, Indrani Medhi, Edward Cutrell, and William Thies. 2012. "Yours is Better!": Participant Response Bias in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1321–1330. https://doi.org/10.1145/2207676.2208589
[12] Brian DeRenzi, Nicola Dell, Jeremy Wacksman, Scott Lee, and Neal Lesh. 2017. Supporting community health workers in India through voice-and web-based feedback. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2770–2781.
[13] Brian DeRenzi, Leah Findlater, Jonathan Payne, Benjamin Birnbaum, Joachim Mangilima, Tapan Parikh, Gaetano Borriello, and Neal Lesh. 2012. Improving community health worker performance through automated SMS. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*. ACM, 25–34.
[14] Brian DeRenzi, Jeremy Wacksman, Nicola Dell, Scott Lee, Neal Lesh, Gaetano Borriello, and Andrew Ellner. 2016. Closing the feedback Loop: A 12-month evaluation of ASTA, a self-tracking application for ASHAs. In *Proceedings of the Eighth International Conference on Information and Communication Technologies and Development*. ACM, 22.
[15] DHIS2. 2019. District Health Information Software 2. Retrieved April 4, 2019 from https://www.dhis2.org/.
[16] Andy Dow, John Vines, Toby Lowe, Rob Comber, and Rob Wilson. 2017. What Happens to Digital Feedback?: Studying the Use of a Feedback Capture Platform by Care Organisations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 5813–5825.
[17] Bassey Ebenso, Rumana Huque, Zunayed Azdi, Helen Elsey, Shammi Nasreen, and Tolib Mirzoev. 2017. Protocol for a mixed-methods realist evaluation of a health service user feedback system in Bangladesh. *BMJ Open* 7, 6 (2017). https://doi.org/10.1136/bmjopen-2017-017743 arXiv:https://bmjopen.bmj.com/content/7/6/e017743.full.pdf
[18] Mirta Galesic and Michael Bosnjak. 2009. Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public opinion quarterly* 73, 2 (2009), 349–360.
[19] Mareike Glöss, Moira McGregor, and Barry Brown. 2016. Designing for labour: uber and the on-demand mobile workforce. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 1632–1643.
[20] DF Hamilton, Judith V Lane, P Gaston, JT Patton, DJ Macdonald, AHRW Simpson, and CR Howie. 2014. Assessing treatment outcomes using a single question: the net promoter score. *The bone & joint journal* 96, 5 (2014), 622–628.
[21] Isaac Holeman, Tara Patricia Cookson, and Claudia Pagliari. 2016. Digital technology for health sector governance in low and middle income countries: a scoping review. *Journal of global health* 6, 2 (2016).
[22] Isaac Holeman and Dianna Kane. forthcoming. Human-Centered Design for Global Health Equity. *Information Technology for Development* (forthcoming).
[23] Sophie Yahui Hsieh. 2011. Healthcare complaints handling systems: a comparison between Britain, Australia and Taiwan. *Health services management research* 24, 2 (2011), 91–95.
[24] Nick Hughes and Susie Lonie. 2007. M-PESA: mobile money for the âĂIJunbankedâĂİ turning cellphones into 24-hour tellers in Kenya. *Innovations: technology, governance, globalization* 2, 1-2 (2007), 63–81.
[25] Audun Jøsang, Roslan Ismail, and Colin Boyd. 2007. A survey of trust and reputation systems for online service provision. *Decision support systems* 43, 2 (2007), 618–644.
[26] Rebecca Koladycz, Gwendolyn Fernandez, Kate Gray, and Heidi Marriott. 2018. The Net Promoter Score (NPS) for insight into client experiences in sexual and reproductive health clinics. *Global Health: Science and Practice* 6, 3 (2018), 413–424.
[27] Donghun Lee, Woochang Hyun, Jeongwoo Ryu, Woo Jung Lee, Wonjong Rhee, and Bongwon Suh. 2015. An analysis of social features associated with room sales of Airbnb. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*. ACM, 219–222.
[28] Xiao Ma, Jeff Hancock, and Mor Naaman. 2016. Anonymity, intimacy and self-disclosure in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3857–3869.
[29] Mikhail I Melnik and James Alm. 2002. Does a seller's ecommerce reputation matter? Evidence from eBay auctions. *The journal of industrial economics* 50, 3 (2002), 337–349.
[30] Medic Mobile. 2018. Medic Mobile. Retrieved June 25, 2018 from https://medicmobile.org/.
[31] Anjali K Mohan, Edward Cutrell, and Balaji Parthasarathy. 2013. Instituting credibility, accountability and transparency in local service delivery?: helpline and Aasthi in Karnataka, India. In *Proceedings of the Sixth International Conference*

on Information and Communication Technologies and Development: Full Papers-Volume 1. ACM, 238–247.

[32] Maletsabisa Molapo, Melissa Densmore, and Limpho Morie. 2016. Apps and Skits: Enabling New Forms of Village-To-Clinic Feedback for Rural Health Education. In *Proceedings of the 7th Annual Symposium on Computing for Development*. ACM, 10.

[33] Mehrab Bin Morshed, Michaelanne Dye, Syed Ishtiaque Ahmed, and Neha Kumar. 2017. When the Internet Goes Down in Bangladesh.. In *CSCW*. 1591–1604.

[34] David Nemer, Ian Spangler, and Michaelanne Dye. 2018. Airbnb and the Costs of Emotional Labor in Havana, Cuba. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 245–248.

[35] Fabian Okeke, Lucas Nene, Anne Muthee, Stephen Odindo, Dianna Kane, Isaac Holeman, and Nicola Dell. 2019. Opportunities and challenges in connecting care recipients to the community health feedback loop. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*. ACM, 13.

[36] Somani Patnaik, Emma Brunskill, and William Thies. 2009. Evaluating the accuracy of data collection on mobile phones: A study of forms, SMS, and voice. In *2009 International Conference on Information and Communication Technologies and Development (ICTD)*. IEEE, 74–84.

[37] Trevor Perrier, Nicola Dell, Brian DeRenzi, Richard Anderson, John Kinuthia, Jennifer Unger, and Grace John-Stewart. 2015. Engaging pregnant women in Kenya with a hybrid computer-human SMS communication system. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1429–1438.

[38] Trevor Perrier, Brian DeRenzi, and Richard Anderson. 2015. USSD: the third universal app. In *Proceedings of the 2015 Annual Symposium on Computing for Development*. ACM, 13–21.

[39] Fahad Pervaiz, Trevor Perrier, Sompasong Phongphila, and Richard Anderson. 2015. User errors in SMS based reporting systems. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. ACM, 55.

[40] Rowan Phipps, Shrirang Mare, Peter Ney, Jennifer Webster, and Kurtis Heimerl. 2018. ThinSIM-based Attacks on Mobile Money Systems. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*. ACM, 23.

[41] Kiran Gopakumar Rajalekshmi. 2007. E-governance services through telecenters: The role of human intermediary and issues of trust. *Information Technologies & International Development* 4, 1 (2007), pp–19.

[42] Subhajyoti Ray. 2012. Reinforcing accountability in public services: an ICT enabled framework. *Transforming Government: People, Process and Policy* 6, 2 (2012), 135–148.

[43] Fred Reichheld. 2006. The ultimate question. *Harvard Business School Press, Boston, MA* (2006).

[44] Frederick F Reichheld and Rob Markey. 2011. *The Ultimate Question 2.0: How net promoter companies thrive in a customer-driven world*. Harvard business press.

[45] David Sando, Hannah Ratcliffe, Kathleen McDonald, Donna Spiegelman, Goodluck Lyatuu, Mary Mwanyika-Sando, Faida Emil, Mary Nell Wegner, Guerino Chalamilla, and Ana Langer. 2016. The prevalence of disrespect and abuse during facility-based childbirth in urban Tanzania. *BMC pregnancy and childbirth* 16, 1 (2016), 236.

[46] Kiley Sobel, Geza Kovacs, Galen McQuillen, Andrew Cross, Nirupama Chandrasekaran, Nathalie Henry Riche, Ed Cutrell, and Meredith Ringel Morris. 2017. EduFeed: A Social Feed to Engage Preliterate Children in Educational Activities. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 491–504.

[47] Vivek Srinivasan, Vibhore Vardhan, Snigdha Kar, Siddhartha Asthana, Rajendran Narayanan, Pushpendra Singh, Dipanjan Chakraborty, Amarjeet Singh, and Aaditeshwar Seth. 2013. Airavat: An automated system to increase transparency and accountability in social welfare schemes in India. In *Proceedings of the Sixth International Conference on Information and Communications Technologies and Development: Notes-Volume 2*. ACM, 151–154.

[48] Anselm Strauss and Juliet M Corbin. 1990. *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Publications, Inc.

[49] Africa's Talking. 2019. Africa's Talking. Retrieved March 26, 2019 from https://africastalking.com/.

[50] Nick Taylor, Keith Cheverst, Peter Wright, and Patrick Olivier. 2013. Leaving the wild: lessons from community technology handovers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1549–1558.

[51] Aditya Vashistha, Edward Cutrell, Gaetano Borriello, and William Thies. 2015. Sangeet swara: A community-moderated voice forum in rural india. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 417–426.

[52] Aditya Vashistha and William Thies. 2012. IVR Junction: Building Scalable and Distributed Voice Forums in the Developing World.. In *NSDR*.

[53] Michalis Vitos, Julia Altenbuchner, Matthias Stevens, Gillian Conquest, Jerome Lewis, and Muki Haklay. 2017. Supporting collaboration with non-literate forest communities in the congo-basin. In *Proceedings of the 2017 ACM Conference*

on Computer Supported Cooperative Work and Social Computing. ACM, 1576–1590.

[54] Galen Weld, Trevor Perrier, Jenny Aker, Joshua E. Blumenstock, Brian Dillon, Adalbertus Kamanzi, Editha Kokushubira, Jennifer Webster, and Richard J. Anderson. 2018. eKichabi: Information Access Through Basic Mobile Phones in Rural Tanzania. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 133, 12 pages. https://doi.org/10.1145/3173574.3173707

[55] Caroline Whidden, Kassoum Kayentao, Jenny X Liu, Scott Lee, Youssouf Keita, Djoumé Diakité, Alexander Keita, Samba Diarra, Jacqueline Edwards, Amanda Yembrick, et al. 2018. Improving Community Health Worker performance by using a personalised feedback dashboard for supervision: a randomised controlled trial. *Journal of global health* 8, 2 (2018).

[56] Susan P Wyche, Sarita Yardi Schoenebeck, and Andrea Forte. 2013. Facebook is a luxury: An exploratory study of social media use in rural Kenya. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 33–44.

Received April 2019; revised June 2019; accepted August 2019