Validity of Using Mobile Phone Surveys to Evaluate Community Health Worker Program in Mali

Xiaomeng Chen (chenxm@jhu.edu)
Johns Hopkins University Bloomberg School of Public Health https://orcid.org/0000-0001-9620-7361

Diwakar Mohan
Johns Hopkins University Bloomberg School of Public Health

Abdoulaye Maïga
Johns Hopkins University Bloomberg School of Public Health

Emily Frost
Johns Hopkins University Bloomberg School of Public Health

Djeneba Coulibaly
Johns Hopkins University Bloomberg School of Public Health

Luay Basil
Canadian Red Cross

Birahim Yaguemar Gueye
American Red Cross

Mariam Traore Guindo
National Institute for Research in Public Health

Assa Sidibé Keita
National Institute for Research in Public Health

Haoua Dembelé Keita
National Institute for Research in Public Health

Melissa A Marx
Johns Hopkins University Bloomberg School of Public Health

Research article

Keywords: Mali, implementation strength, validation, sensitivity, specificity, integrated community case management, family planning, community health worker, mobile phones

Posted Date: January 6th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-138977/v1

Abstract

Background

Assessing implementation strength through face-to-face interviews in hard-to-reach and unstable regions presents many challenges. Mobile phone-based interviews have become an alternative, but the validity of these data from this approach for Implementation Strength Assessment evaluation has not been sufficiently studied yet. The objective of this study was to assess the validity of mobile phone-based health provider interviews to measure the implementation strength of an integrated community case management (iCCM) and family planning program in Mali.

Methods

From July to August 2018, interviewers administered a structured questionnaire to community health workers (ASCs) implementing the iCCM and family planning program in six districts in Mali. Interviews were conducted, first by phone, then verified through in-person visits. Survey questions addressed background information, training, supervision, demand generation activities and supplies of essential drugs and contraceptive methods. Sensitivity and specificity of the phone responses were calculated using the in-person response as the gold standard. A threshold of 80% for sensitivity and specificity respectively was considered acceptable.

Results

Of 157 ASCs interviewed by phone, 115 (73.2%) were reached in-person. Most indicators (9/10 iCCM indicators, 6/6 family planning indicators) for training, supervision, and availability of supplies on the day of interview, and those related to patient utilization reached the 80% threshold for sensitivity, but few (2/10 iCCM indicators, 1/6 family planning indicators) reached 80% for specificity. In contrast, most indicators of supply stock-outs in the last 3 months reached the threshold for specificity (5/6 iCCM indicators, 3/3 family planning indicators) but few reached the threshold for sensitivity (1/6 iCCM indicators, 1/3 family planning indicators).

Conclusions

The validity of data collected by phone were adequate for indicators of training, supervision, and day-of-interview commodity availability. Phone-based surveys are useful as a low-cost option for data collection in the assessment of implementation strength on general activities in inaccessible and resource-limited regions with mobile network connectivity.

Background

Data availability and quality are essential for measurement and accountability of health initiatives and programs including the Sustainable Development Goals (SDGs) [1]. However, data incompleteness and lack of quality remain as challenges in low- and middle-income (LMICs), especially for programs that

operate in hard-to-reach areas due to logistical and/or security concerns. Evaluations of health programs and decision-making in LMICs largely rely on special surveys like the Demographic and Health Surveys (DHS) and facility assessments as main data sources [2]. Such surveys consume a considerable amount of human resources, time and financial resources, in part due to field-related costs. The UN Commission on Information and Accountability for Women's and Children's Health recommends all countries to use Information and Communication Technologies (ICTs) to obtain "better information for better results" [3]. Several pilot projects have found mobile phone-based data collection helpful, especially in remote and rural areas, enabling more frequent follow-ups and faster data collection [3].

According to the estimations from the Global System for Mobile Communications (GSMA) Intelligence, the percentage of the population in sub-Saharan Africa who subscribe to mobile services increased from 32% in 2012 to 44% in 2018, and is estimated to reach 50% in 2025 [4]. Given this prevalence of mobile phone ownership in Africa, communicating with survey respondents via their mobile phones can serve as an alternative method for data collection. This is especially relevant for collecting program and health system data from community health workers, as CHWs are often placed in areas that are hard-to-reach [5]. Using mobile phones makes it feasible to gather data in remote and high-risk environments where traditional face-to-face data collection is extremely difficulty or dangerous, allows for rapid response to newly identified needs, and could be cost-effective [6]. Some studies have explored the validity and reliability of mobile phone-based data collection in LMICs, and results show it can be a cost-effective way to collect data [7–12]. A study from Gujarat, India assessed the health data entry accuracy of three user interfaces on mobile phones, indicating that error rates per datum were 4.2% for electronic forms, 4.8% for short message service, and 0.45% for voice [7]. Studies have found reliable results in using mobile phones for the supervision of community health workers (CHWs) in an Integrated Community Case Management (iCCM) program in sub-Sahara African countries [8, 9]. The validity of collecting data via phone interviews has been tested in Malawi to assess the strength of program implementation for iCCM and family planning (FP) with reasonable sensitivity of at least 80% [10, 11].

Mali is a low-income country in sub-Saharan Africa with one of the highest rates of maternal and child mortality and morbidity. In 2015, the maternal mortality ratio was estimated to be 587 per 100,000 live births, and the probability of a child dying before reaching the age of five was 115 per 1,000 live births in Mali [13]. However, the data quality of Mali is imperfect. Of the nine regions in Mali, three were not covered for security reasons during the 2012-13 DHS [14] and the 2015 Malaria Indicator Survey (MIS) [15]. Beginning in 2016, the Canadian Red Cross (CRC), in partnership with the Mali Red Cross (MRC) and the Mali Ministry of Health (MoH), implemented a 4-year Maternal, Newborn and Child Health (MNCH) program in the Koulikoro and Sikasso regions, covering villages that were difficult to access purposefully in order to bring health services to hard-to-reach areas. This program trained CHWs and MoH staff to provide MNCH services, with the aim of reducing maternal and child mortality and morbidity. Community Health Workers, locally called ASCs (Agents de Santé Communautaires) were recruited from rural villages and trained to provide selected FP methods and essential care for children with malaria, diarrhea, acute respiratory infection and malnutrition. The evaluation of the above-mentioned program was conducted in six rural districts (Banamba, Dioila, Kolokani, Koulikoro, Nara, and Sikasso). In 2018, an evaluation of the

implementation strength and quality of care of the Red Cross program providing these services to sick children was performed. The evaluation used a streamlined protocol for implementation strength assessment (ISA) which was developed for rapid quantitative assessment of the strength of program roll-out by the Real Accountability: Data Analysis for Results (RADAR) program from the Institute of International Programs (IIP) at Johns Hopkins Bloomberg School of Public Health [16]. This protocol has been utilized in Ethiopia and Malawi to assess the intensity of FP and iCCM program implementation [10, 11, 17]. The present study is nested within the broader evaluation of the above described program in Mali. The objective of this study was to assess the validity of mobile phone-based ASCs interviews for collecting data on the implementation strength of the iCCM and FP program in Mali.

Methods

Study design

This study used the date from a cross-sectional survey among ASCs working in the Red Cross program, assessing the validity of iCCM and FP implementation strength indicators collected via mobile phone-based interviews. ASCs were first contacted and interviewed by mobile phone, during which data collectors administered the study survey. Within a week, the data collectors visited the ASCs to administer the same survey in-person. The interviews used identical implementation strength tools, collecting data on provider's program activities on iCCM and FP. Red Cross program staff provided the study team with a list of mediations, supplies, job aids, and registers that ASCs had been supplied with and should have in stock; these were all asked about (by mobile phone) or examined (in-person) as part of the survey. We compared the responses collected via mobile phone-based interviews with responses collected from in-person interviews among the same sample of ASCs, using the in-person data as the gold standard. To assess the validity of phone-based data, the phone-based samples were paired with corresponding in-person samples and only matched pairs were included in the analysis. We also collected socio-demographics information in both interviews for verification purpose.

Sample size

Participants in this study were the ASCs working with CRC-MRC and responsible for delivering iCCM in program areas throughout the program districts. The program implementers provided the sampling frame, including all ASCs from the 6 program districts who had received training on the program and were actively providing services in their communities. All the 441 ASCs on the list were included in the sampling frame and 300 ASCs were selected for in-person survey using a stratified random sample proportional to the numbers of ASCs by district for the overall study. Having confirmed that each ASC owns a mobile phone, random sampling proportional to the ASC number was used to select the 150 ASCs out the 300, and the 150 ASCs were assigned to receive both a mobile phone-based interview and an in-person interview.

In practice, out of the 300 ASCs, 63 in-person interviews were not possible due to rainy season (50), security reasons (5), defect phone network (2), closure/change of ASC site (4), ASC non available (2).

This study used responses from the 115 ASCs who were interviewed by both phone and in-field interviews. This sample size is adequate for sensitivity of 90% and specificity of 90%, with an indicator prevalence at 60% for in-person mode, with the Z statistics set to 1.96 corresponding to the 95% confidence interval [18]. The sampling process and the final sample are shown in Fig. 1. Supplemental table 1 shows the precision, according to Buderer's formula [18], of the estimate for sensitivity and specificity for differing values of indicator prevalence with type I error of 0.05.

Data collection

MRC provided a list of and phone contact numbers for most of ASCs, which was established during their recruitment and training. We completed the list for missing numbers and cross-checked the numbers with the contact numbers provided by the ASC supervisors, who were head nurses of the health facility catchments to which the villages covered by the ASCs belonged.

The selected ASCs were first contacted and informed of the study by their health facility-specific supervisors through mobile phone. The supervisors stressed that the results of the evaluation would only be used to improve health services in general and would not be used for individual assessment. The data collectors then contacted the ASCs by phone, introduced the process of the study including the future in-person interview, and obtained oral informed consent to participate. ASCs who provided oral informed consent were included in the study. With the oral consent from the ASCs, data collectors administered the interview to the ASCs. Each interview was required to be completed within three phone calls. If the interviewer did not complete the interview by the third call, no further calls were made, and data were excluded from the analysis. No more than a week later, the data collectors visited the ASCs to administer the in-person interview using identical assessment tools. Before the start of the in-person interview, the data collectors informed about the study again and obtained written informed consent from the ASCs. Any ASCs refusing to participate at any time were no longer contacted for this study, and all data collected up to that point was deleted. During in-person interview, the data collectors inspected the ASCs' records to obtain data regarding treatment supply and stockouts.

Data were collected from July through August in 2018 using an electronic data capture system running Open Data Kit (ODK) on portable Android tablets. Data were entered directly into ODK by the data collectors during the mobile phone-based and in-person interviews. In-person interview data were collected on the basis of the ASCs' statement by direct observation, from the register; phone-based data were based solely on the ASCs' statement. The survey consisted of three modules: module 1 was used to collect identification information of the ASCs, socio-demographic characteristics (age, sex, socio-cultural groups, marital status, length of living in the village site, and education level), training, supervision, and other work activities; module 2 focused on the stocks of drugs, equipment, and materials; module 3 collected data on the information recorded in the registers of sick children and FP kept by the ASC. In the training and supervision sections, questions focused on the reception of each activity by the ASC, the time and content of the activities. The availability of equipment for the provision of quality services, stocks of medicines for sick children and contraceptive supplies were requested in the equipment and

supplies section. Inventories of treatment supplies were measured by the “availability” of unexpired drugs or supplies on the day of the interview. For history of stock-outs, we asked whether the ASC had stock-outs of unexpired drugs or supplies lasting more than seven consecutive days in the last three months. A stock-out is defined as lasting more than a week. In terms of the register for sick child consultations, ASCs were asked about how the frequency with which they recorded the date of consultation, name of child, age of child, sex of child, signs/symptoms, diagnosis and treatment in the register of sick child. For the FP registry, the ASCs were asked about how often they recorded clients’ name, age and method of choice. Besides, we also collected socio-demographics information in both interviews for verification purpose. Both the iCCM and FP questionnaires used in this study were developed by the RADAR program and are available at https://www.radar-project.org/isaqoc [16].

Analysis

We analyzed provider interviews for iCCM (Supplemental Table 2) and FP (Supplemental Table 3) separately. To validate the indicators collected by the phone-based survey, we considered the in-person survey as the gold standard. Sensitivity and specificity for each indicator of interest was calculated as defined below.

Sensitivity of an indicator reported by phone survey was defined as the ability to correctly identify ASCs who reported having the activity or practice. Eq. (1) was used for sensitivity computation [19].

$$Sensitivity = \frac{Truepositives}{Truepositives + Falsenegatives}$$

1

True positives were counted when an ASC responded “yes” or in an otherwise positive manner (based on the question) to the same question in both the mobile phone and in-person survey. False negatives occurred when an ASC responded positively during the phone survey but negatively during the in-person survey.

Specificity of an indicator was defined as the ability to correctly identify ASCs who reported not having the activity or practice. Eq. (2) was used for specificity computation [19].

$$Specificity = \frac{Truenegatives}{Truenegatives + Falsepositives}$$

2

True negatives were counted when an ASC gave a negative response to the same question in both the mobile phone and in-person surveys. False positives occurred when an ASC responded positively in the mobile phone survey but negatively to the same question in the in-person survey.

Unpaired two-sample t-tests were used to assess the difference in distributions of continuous variables between ASCs who were reached by both (phone and in-person) interviews and those who were interviewed only via phone. Binominal test of difference in proportions was used for the assessment of categorical variables with these same populations. Both tests followed two-sided with α=0.05 to determine the statistical significance. Sensitivity analysis was conducted to assess the potential selection bias due to the failure to contact some participants for in-person interview. All the data analyses were conducted by Stata/SE 14.1.

Ethical consideration

This study was approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board (IRB#00008512) and the Faculté de Médecine, de Pharmacie et d'Odontostomatologie Review Board (N°2018/77/CE/FMPOS) in Mali. Both oral consent for phone interview and written consent for in-person interview were granted by each participant in this study before data collection.

Results

There were 157 ASCs reached by phone in the six program districts, of whom 115 (73.2%) were reached in the field for in-person interviews. The 115 ASCs that had both in-person and field-based interviews were used for the subsequent data analysis. These ASCs had a mean age of 29.6 ± 6.9 years and 39.1% were male. Most reported being of Bambara ethnic group (53.9%), married or engaged (82.6%), with primary education (53.0%) and living in the village for more than 3 years (51.3%). The in-person interviews were conducted after a median of 4 (range from 1 to 11 days) from when the phone-based interviews were conducted. Figure 2 shows the geographic distribution of the 115 interviews.

iCCM analysis

The results of implementation strength analysis for iCCM (Table 1) verified that all 115 ASCs surveyed both via phone and in-person reported having ever received training in iCCM, as expected by the program. According to in-person survey results, the majority (95.7%) of ASCs having been trained within the past two years, and most reported receiving a refresher training (87.0%) and receiving a supervision in the prior three months (82.6%). All three of these indicators (trained in the past two years, refresher training, and supervision) had high sensitivity (100.0%, 92.0%, 91.6%) but specificity varied (20.0%, 40.0%, 80.0%). A lower proportion of ASCs (72.2%) reported in-person that within the past three months they had received a supervision which included the observation of a sick child consultation, and this indicator showed a sensitivity of 72.3% and a specificity of 40.6%.

Table 1

Implementation strength assessment (ISA) indicators for integrated community case management (iCCM) reported by community health workers (ASCs) via phone versus in-person with sensitivity and specificity of the phone-based interview. Sensitivity and specificity that are considered adequate (> 80%) are bolded.

ISA indicator	Phone % (n/N)*	In-person % (n/N)*	Sensitivity % (95% CI)	Specificity % (95% CI)
Training				
Proportion of ASCs receiving training in the last 2 years	99.1 (114/115)	95.7 (110/115)	**100.0** (96.7, 100.0)	20.0 (0.5, 71.6)
Proportion of ASCs receiving refresher training	87.8 (101/115)	87.0 (100/115)	**92.0** (84.8, 96.5)	40.0 (16.3, 67.7)
Supervision				
Proportion of ASCs supervised in the last 3 months	79.1 (91/115)	82.6 (95/115)	**91.6** (84.1, 96.3)	**80.0** (56.3, 94.3)
Proportion of ASCs receiving supervision which included observation of a sick child consultation in the last 3 months	68.7 (79/115)	72.2 (83/115)	72.3 (61.4, 81.6)	40.6 (23.7, 59.4)
Treatment supply				
Proportion of ASCs with a supply of unexpired RDT on the day of the assessment	93.0 (107/115)	93.9 (108/115)	**97.2** (92.1, 99.4)	71.4 (29.0, 96.3)
Proportion of ASCs with a supply of unexpired Amoxicillin tablets/syrup on the day of the assessment	98.3 (113/115)	95.7 (110/115)	**100.0** (96.7, 100.0)	40.0 (5.3, 85.3)
Proportion of ASCs with a supply of unexpired ACT on the day of the assessment	98.3 (113/115)	93.0 (107/115)	**100.0** (96.6, 100.0)	25.0 (3.2, 65.1)

Note: ASC = Agente Sante Communitaire, i.e. community health workers; RDT = rapid diagnostic test for Malaria; ACT = artimisinin-based combination therapy; ORS = oral rehydration salts

*n is the frequency of each indicator; N is the total number of ASCs.

ISA indicator	Phone % (n/N)*	In-person % (n/N)*	Sensitivity % (95% CI)	Specificity % (95% CI)
Proportion of ASCs with a supply of unexpired ORS on the day of the assessment	74.8 (86/115)	53.0 (61/115)	**90.2** (79.8, 96.3)	42.6 (29.2, 56.8)
Proportion of ASCs with a supply of unexpired zinc on the day of the assessment	82.6 (95/115)	83.5 (96/115)	**93.8** (86.9, 97.7)	73.7 (48.8, 90.9)
Proportion of ASCs with a supply of unexpired Plumpy Nut on the day of the assessment	47.8 (55/115)	54.8 (63/115)	**81.0** (69.1, 89.8)	**92.3** (81.5, 97.9)
Treatment stock-outs				
Proportion of ASCs reporting RDT stock-out that lasted more than 1 consecutive week in the past 3 months	21.7 (25/115)	17.4 (20/115)	75.0 (50.9, 91.3)	**89.5** (81.5, 94.8)
Proportion of ASCs reporting Amoxicillin tablets/syrup stock-out that lasted more than 1 consecutive week in the past 3 months	7.0 (8/115)	24.3 (28/115)	10.7 (2.3, 28.2)	**94.3** (87.1, 98.1)
Proportion of ASCs reporting ACT stock-out that lasted more than 1 consecutive week in the past 3 months	10.4 (12/115)	4.3 (5/115)	40.0 (5.3, 85.3)	**90.9** (83.9, 95.6)
Proportion of ASCs reporting ORS stock-out that lasted more than 1 consecutive week in the past 3 months	16.5 (19/115)	20.9 (24/115)	50.0 (29.1, 70.9)	**92.3** (84.8, 96.9)
Proportion of ASCs reporting zinc stock-out that lasted more than 1 consecutive week in the past 3 months	17.4 (20/115)	15.7 (18/115)	72.2 (46.5, 90.3)	**92.8** (85.7, 97.0)
Proportion of ASCs reporting Plumpy Nut stock-out that lasted more than 1 consecutive week in the past 3 months	53.9 (62/115)	50.4 (58/115)	**86.2** (74.6, 93.9)	78.9 (66.1, 88.6)
Reporting				

Note: ASC = Agente Sante Communitaire, i.e. community health workers; RDT = rapid diagnostic test for Malaria; ACT = artimisinin-based combination therapy; ORS = oral rehydration salts

*n is the frequency of each indicator; N is the total number of ASCs.

ISA indicator	Phone % (n/N)*	In-person % (n/N)*	Sensitivity % (95% CI)	Specificity % (95% CI)
Proportion of ASCs with a completed patient registry for iCCM	98.3 (113/115)	88.7 (102/115)	**98.0** (93.1, 99.8)	0.0 (0.0, 24.7)
Note: ASC = Agente Sante Communitaire, i.e. community health workers; RDT = rapid diagnostic test for Malaria; ACT = artimisinin-based combination therapy; ORS = oral rehydration salts				
*n is the frequency of each indicator; N is the total number of ASCs.				

We assessed the availability of the following supplies and medications in both mobile phone-based and in-person surveys, based on a list of supplies and medications provided by Red Cross program staff: Rapid Diagnostic Test (RDT) for malaria, Amoxicillin, artemisinin-based combination therapy (ACT), oral rehydration salts (ORS), zinc and Ready to use therapeutic food (RUTF – locally called Plumpy Nut). Most ASCs reported in-person having most of the supplies on the day of assessment (83.5–95.7%), except for ORS (53.0%) and Plumpy Nut (54.8%). All six indicators had high sensitivity (81.0-100.0%). Specificity varied greatly (25.0-92.3%). Responses to the supply of Plumpy Nut had the highest sensitivity and specificity (81.0% and 92.3%, respectively).

Fewer than half of ASCs reported any stock-outs lasting more than seven consecutive days of RDT (17.4%), Amoxicillin (24.3%), ACT (4.3%), ORS (20.9%), and zinc (15.7%) in the prior 3 months, with specificity ranging from 89.5–94.3% and sensitivity from1 0.7–75.0%. About half of ASCs (50.4%) reported stock-outs of Plumpy Nut lasting more than seven consecutive days (86.2% sensitivity and 78.9% specificity). Most ASCs (88.7%) were able to produce completed patient registers for iCCM. Sensitivity for this indicator was 98.0%, but specificity was 0.0%.

FP analysis

All ASCs reported having received FP training. As shown in Table 2, most ASCs (79.1%) reported in-person having received a supervision on FP in the prior 3 months, with sensitivity of 83.5% and specificity of 75.0%. We assessed the availability of the following FP commodities, which ASCs were trained to provide as part of the Red Cross program: male condoms, oral contraceptive pills and injectables. Most ASCs reported having supplies of injectables (87.8%) and oral contraceptive pills (78.3%) on the day of assessment, but fewer ASCs reported having a supply of male condoms (59.1%). All three indicators had high sensitivity (92.2–96.0%) with specificity ranging from 57.1% for injectables to 87.2% for male condoms. Less than one third of the ASCs reported stock-outs in-person that lasted more than one week in the past three months for the three contraceptive methods; a higher proportion of ASCs reported stock-outs for male condoms (32.2%) than for oral contraceptives (13.9%) or injectables (3.5%). Male condoms stock-out had high sensitivity (91.9%) and high specificity (89.7%); both oral contraceptive pills and injectables stock-out indicators had similarly high specificity (82.8%, 96.4%) and lower sensitivity (43.8%,

50.0%). Both FP client eligibility and follow-up registers had the same validity. More than half of ASCs (61.7%) reported that they recorded patient information completely in FP registers, with high sensitivity (93.0%) and low specificity (34.1%).

Table 2

Implementation strength assessment (ISA) indicators for family planning (FP) reported by community health workers (ASCs) via phone versus in-person with sensitivity and specificity of the phone-based interview. Sensitivity and specificity that are considered adequate (> 80%) are bolded.

ISA indicator	Phone % (n/N)*	In-person % (n/N)*	Sensitivity % (95% CI)	Specificity % (95% CI)
Supervision				
Proportion of ASCs supervised in FP in the last 3 months	71.3 (82/115)	79.1 (91/115)	**83.5** (74.3, 90.5)	75.0 (53.3, 90.2)
Modern contraceptives supply				
Proportion of ASCs with a supply of unexpired male condoms on the day of the assessment	60.0 (69/115)	59.1 (68/115)	**92.6** (83.7, 97.6)	**87.2** (74.3, 95.2)
Proportion of ASCs with a supply of unexpired oral contraceptive pills on the day of the assessment	79.1 (91/115)	78.3 (90/115)	**92.2** (84.6, 96.8)	68.0 (46.5, 85.1)
Proportion of ASCs with a supply of unexpired injectables on the day of the assessment	89.6 (103/115)	87.8 (101/115)	**96.0** (90.2, 98.9))	57.1 (28.9, 82.3)
Modern contraceptives stock-outs				
Proportion of ASCs reporting male condoms stock-out that lasted more than 1 consecutive week in the past 3 months	36.5 (42/115)	32.2 (37/115)	**91.9** (78.1, 98.3)	**89.7** (80.8, 95.5)
Proportion of ASCs reporting oral contraceptive pills stock-out that lasted more than 1 consecutive week in the past 3 months	20.9 (24/115)	13.9 (16/115)	43.8 (19.8, 70.1)	**82.8** (73.9, 89.7)
Proportion of ASCs reporting injectables stock-out that lasted more than 1 consecutive week in the past 3 months	5.2 (6/115)	3.5 (4/115)	50.0 (6.8, 93.2)	**96.4** (91.0, 99.0)
Reporting				
Proportion of ASCs with a completed patient registry for FP client eligibility	82.6 (95/115)	61.7 (71/115)	**93.0** (84.3, 97.7)	34.1 (20.5, 49.9)

*n is the frequency of each indicator; N is the total number of ASCs.

ISA indicator	Phone % (n/N)*	In-person % (n/N)*	Sensitivity % (95% CI)	Specificity % (95% CI)
Proportion of ASCs with a completed patient registry for FP follow-up	82.6 (95/115)	61.7 (71/115)	**93.0** (84.3, 97.7)	34.1 (20.5, 49.9)
*n is the frequency of each indicator; N is the total number of ASCs.				

Sample comparison

The socio-demographics of ASCs interviewed by phone alone and both phone and in-person were compared (Supplemental Table 4). The ASCs who were interviewed only by phone, only in-person, and by both modes did not differ in sex, social-cultural group, marital status, education level, and whether living in the village for more than 3 years. ASCs reached only by phone were younger than those interviewed only in-person and fewer ASCs in the two hard-to-access districts (i.e. Nara and Koulikoro) received in-person interview.

Discussion

Mobile phone-based interview appeared to provide reasonably accurate data for many of the implementation indicators of iCCM and FP. Training, supervision, and most medication and contraceptive availability reports on the day of assessment were highly sensitive (sensitivity > 80%). However, the specificity was low for most indicators. Thus, the validity of phone-based survey is limited for specific information, such as reporting on supervision which includes the observation of case management, availability and long-term stock-outs of specific products and completeness of information in registers.

This study showed that ASCs can be contacted for mobile phone surveys in rural area even in rainy season, but the accessibility in-person was more difficult in remote areas in Mali; only 73.2% ASCs in total interviewed by phone were successfully reached for in-person interview. The proportions of ASCs interviewed by both approaches were considerably lower in Koulikoro and Nara than other districts – fewer than 40% of ASCs in these districts were able to be reached in the field. In Nara, the most remote among the six districts, the provider sites were spread across rural areas, which impeded the in-field interviews, and there was regional instability and terrorist violence that made it unsafe for interviewers to travel to the district. ASCs, especially in Koulikoro, were being inaccessible due to rainy season, from which Koulikoro suffered the most. The rainy season also influenced the interval between phone-based survey and in-person survey. This could be influential for the indicator of supply on the day of assessments and week-long stock-outs, even though data collectors were trained to go back in time using stock records, to assess stocks. In the meanwhile, our study was designed to bias towards remote and hard-to-reach areas, thus our results also indicated that it is probably difficult to get health services for

people living the areas that our data collectors failed to reach. It is essential to take these hard-to-reach areas into account when prioritizing equity in health care provision.

The observed difference in responses between the phone and in-person interviews may have many causes. A previous study exploring the validity of phone-based interviews reported that the respondents did not understand the questions being asked on phone clearly, especially for the exact definitions of certain documents or technical terms, and this confusion is more likely to occur among ASCs with less experience [11]. Due to the relatively small sample size for analysis, subgroup analysis for inspecting differences between groups did not have adequate power. Additionally, as was reported in a previous study conducted in Malawi, data collectors may have made errors in reporting or recording data during the phone interviews [10]. This could explain the poorer sensitivity and specificity calculated for detailed data of stock-outs and register completeness. To improve the validity overall and for mobile phone-based survey instruments, evaluators may conduct a pilot study prior to the formal phone-based evaluation in a certain cultural and geographic setting, so that the pilot results can inform a better wording of the instruments.

This study was the first to test the validity of mobile phone-based interviews of health care providers in Mali. Previous studies have been conducted in Malawi [10, 11], showing relatively high validity of implementation indicators. The accuracy of phone surveys shown in this study can inform future program monitoring and evaluation plans in hard-to-reach areas.

There are a few limitations of this study. In program evaluations, the gold standard is in-person interview with available records, as it was used in previous studies [10, 11]. Yet, records are not a perfect source of data. It is hard to determine how accurate the data in the registers are, but these registers were the best resource of reference available for the study purpose. Secondly, the study was powered to report on overall indicators and not for district-specific comparisons, so it was not possible to stratify further, for example to the district level. Due to both poor mobile phone connections (mobile phone coverage of 88%) and limited in-person accessibility, the sample used in this analysis suffered from potential selection bias. We were unable to reach some ASCs via phone due poor mobile networks, and some ASCs in-person who were located in areas rendered inaccessible. Furthermore, ASCs were not asked why there were discrepancies between the two interviews, if discrepancies were found, nor did the study verify the accuracy of program documents and paper-based registry. ASCs might be less concerned about the accuracy of data they reported in mobile phone-based interviews. Thus, the difference in responses might partially result from the inaccuracy of the data collected from phone-based interviews. Despite these limitations, we were able to evaluate the validity of phone-based versus in-person interview and gained knowledge about mobile phone use in low resource and hard to reach settings with considerable mobile phone usage and network reliability.

Using in-person interview as the gold standard, data collection via phone surveys was adequate for iCCM and FP indicators of general training, supervision, routine treatment and contraceptive supply on the day of interview, but had low validity for indicators of week-long stock-outs in the past 3 months and

completed registers in Mali. We recommend that local governments and organizations take advantage of mobile phone-based surveys to monitor and evaluate the implementation strength of programs, especially in areas in areas that are difficult to reach otherwise.

Conclusions

Mobile phone interviews can be used to monitor and evaluate the implementation of health programs in remote areas, given adequate mobile phone use, reliable network connection, and a reliable and up-to-date contact list of health workers. Further assessment of the validity of a mobile phone-based evaluation of program implementation strength in other low-resource settings are needed. And other mobile phone-based approaches to monitor and evaluate long-term stock-outs and completeness of patient registers are also needed.

Abbreviations

ACT
artemisinin-based combination therapy
ASCs
agents de santé communautaire (community health workers)
CHW
community health worker
CRC
Canadian Red Cross
DHS
Demographic and Health Survey
FP
family planning
GSMA
Global System for Mobile Communications
iCCM
integrated community case management
ICTs
Information and Communication Technologies
IIP
Institute for International Programs
ISA
implementation strength assessment
LMICs
low- and middle-income countries
MIS

Malaria Indicator Survey
MNCH
Maternal, Newborn and Child Health
MoH
Ministry of Health
MRC
Mali Red Cross
ODK
Open Data Kit
ORS
oral rehydration salts
RADAR
Real Accountability:Data Analysis for Results
RDT
Rapid Diagnostic Test
SDGs
The Sustainable Development Goals

Declarations

Ethics approval and consent to participate

This study was approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board (IRB#00008512) and the Faculté de Médecine, de Pharmacie et d'Odontostomatologie Review Board (N°2018/77/CE/FMPOS) in Mali. Both oral consent for phone interview and written consent for in-person interview were granted by each participant in this study before data collection. This procedure was approved by both ethical bodies.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by Global Affairs Canada through the Real Accountability: Data Analysis for Results (RADAR) program under the grant No. D001627-001. The funding body had no role in the design, data collection, analysis of the data or writing of the manuscript.

Authors' contributions

MM and DM conceived of the study, with key input from EF and AM. MM, DM and XC designed the quantitative analyses and instruments with input from AM and EF. EF and AM oversaw data collection, with assistance from DC, LB, BG, MG, GT, and DK. XC led data analysis and wrote the first draft. DM, EF, AM, and MM commented on all drafts. All authors contributed to interpretation of results and commented on drafts prior to publication. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to sincerely appreciate the valuable support from the following organizations: Mali Ministry of Health, Mali Red Cross, Canadian Red Cross, Mali National Institute of Public Health, and Institute for International Programs (IIP), and to everyone facilitating this evaluation in the field: regional health directors, head doctors of 6 health districts, technical directors of community health centers, community relays, and ASCs.

References

1. Kuruvilla S, Bustreo F, Kuo T, Mishra CK, Taylor K, Fogstad H, Gupta GR, Gilmore K, Temmerman M, Thomas J *et al*: The Global strategy for women's, children's and adolescents' health (2016-2030): a roadmap based on evidence and country experience. *Bulletin of the World Health Organization* 2016, 94(5):398-400.
2. Kuruvilla S, and D. Hogan.: Country data universal accountability. Monitoring priorities for the Global Strategy for Womens Childrens and Adolescents Health (2016-2030). In. Edited by Richard Cheeseman RTC. Geneva, Switzerland: World Health Organization.
3. Health ColaAfWsaCs: Keeping Promises, Measuring Results. Final Report of the Commission. In. Geneva, Switzerland: WHO; 2011.
4. GSMA: The Mobile Economy Sub-Saharan Africa 2019. In.: GSM Association; 2019.
5. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, Patel V, Haines A: The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. *PLoS medicine* 2013, 10(1):e1001363.

6. Dabalen A, Alvin Etang, Johannes Hoogeveen, Elvis Mushi, Youdi Schipper, and Johannes von Engelhardt: Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection. In: *Directions in Development.* Washington, DC: World Bank; 2016.
7. Patnaik S, Brunskill E, Thies W: 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): 17-19 April 2009 2009*; 2009: 74-84.
8. Tomlinson M, Solomon W, Singh Y, Doherty T, Chopra M, Ijumba P, Tsai AC, Jackson D: The use of mobile phones as a data collection tool: a report from a household survey in South Africa. *BMC medical informatics and decision making* 2009, 9:51.
9. Tumusiime DK, Agaba G, Kyomuhangi T, Finch J, Kabakyenga J, MacLeod S: Introduction of mobile phones for use by volunteer community health workers in support of integrated community case management in Bushenyi District, Uganda: development and implementation process. *BMC health services research* 2014, 14 Suppl 1:S2.
10. Hazel E, Amouzou A, Park L, Banda B, Chimuna T, Guenther T, Nsona H, Victora CG, Bryce J: Real-time assessments of the strength of program implementation for community case management of childhood illness: validation of a mobile phone-based method in Malawi. *The American journal of tropical medicine and hygiene* 2015, 92(3):660-665.
11. Pattnaik A, Mohan D, Chipokosa S, Wachepa S, Katengeza H, Misomali A, Marx MA: Testing the validity and feasibility of using a mobile phone-based method to assess the strength of implementation of family planning programs in Malawi. *BMC health services research* 2020, 20(1):221.
12. Kallander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten Asbroek AH, Conteh L, Kirkwood BR, Meek SR: Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries: a review. *J Med Internet Res* 2013, 15(1):e17.
13. Maternal, Newborn and Child Health in Mali [https://www.redcross.ca/how-we-help/international-programs/maternal-newborn-and-child-health/long-term-maternal-newborn-and-child-health-programs/maternal-newborn-and-child-health-in-africa/maternal-newborn-and-child-health-in-mali]
14. Cellule de Planification et de Statistique (CPS/SSDSPF), Institut National de la Statistique (INSTAT/MPATP), International I-Sel: Enquête Démographique et de Santé au Mali 2012-2013. In. Rockville, Maryland, USA: CPS, INSTAT, INFO-STAT et ICF International; 2014.
15. Programme National de Lutte contre le Paludisme (PNLP), Institut National de la Statistique (INSTAT), INFO-STAT, International INdIReSPIeI: Enquête sur les Indicateurs du Paludisme (EIPM) 2015. In: *MIS (Malaria Indicator Survey).* Rockville, Maryland, USA: INSTAT, INFO-STAT et ICF International; 2016.
16. Real Accountability: Data Analysis for Results (RADAR) [https://www.radar-project.org/]
17. Miller NP, Amouzou A, Tafesse M, Hazel E, Legesse H, Degefie T, Victora CG, Black RE, Bryce J: Integrated community case management of childhood illness in Ethiopia: implementation strength

and quality of care. *The American journal of tropical medicine and hygiene* 2014, 91(2):424-434.

18. Buderer NM: Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad Emerg Med* 1996, 3(9):895-900.
19. Lalkhen AG, McCluskey A: Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia Critical Care & Pain* 2008, 8(6):221-223.

Figures

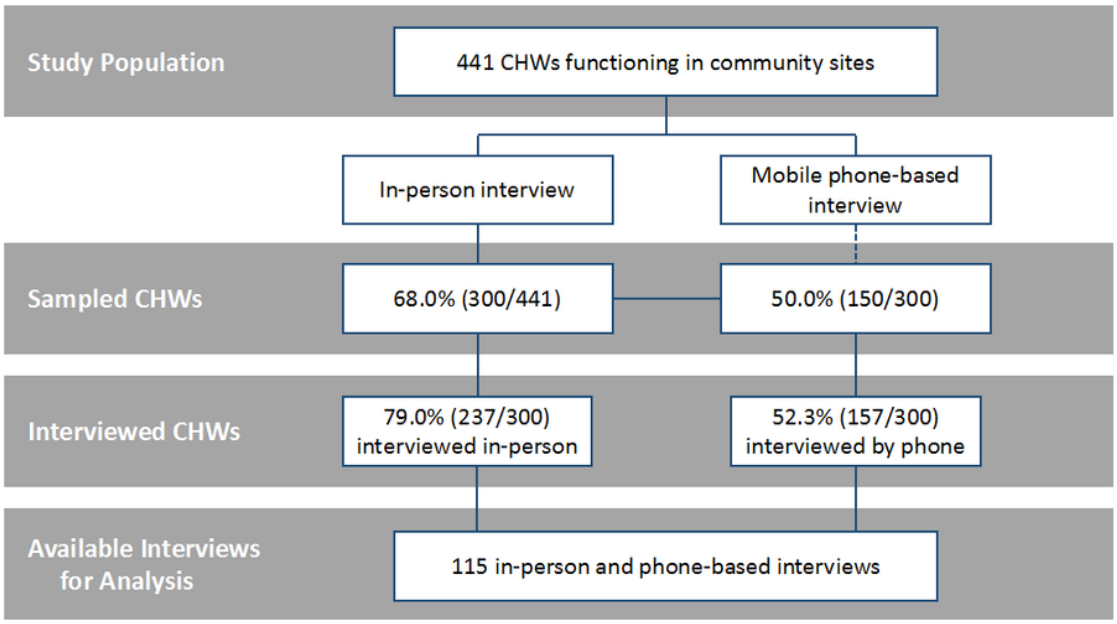

Figure 1

Community health workers (ASCs) population and sample selection process.

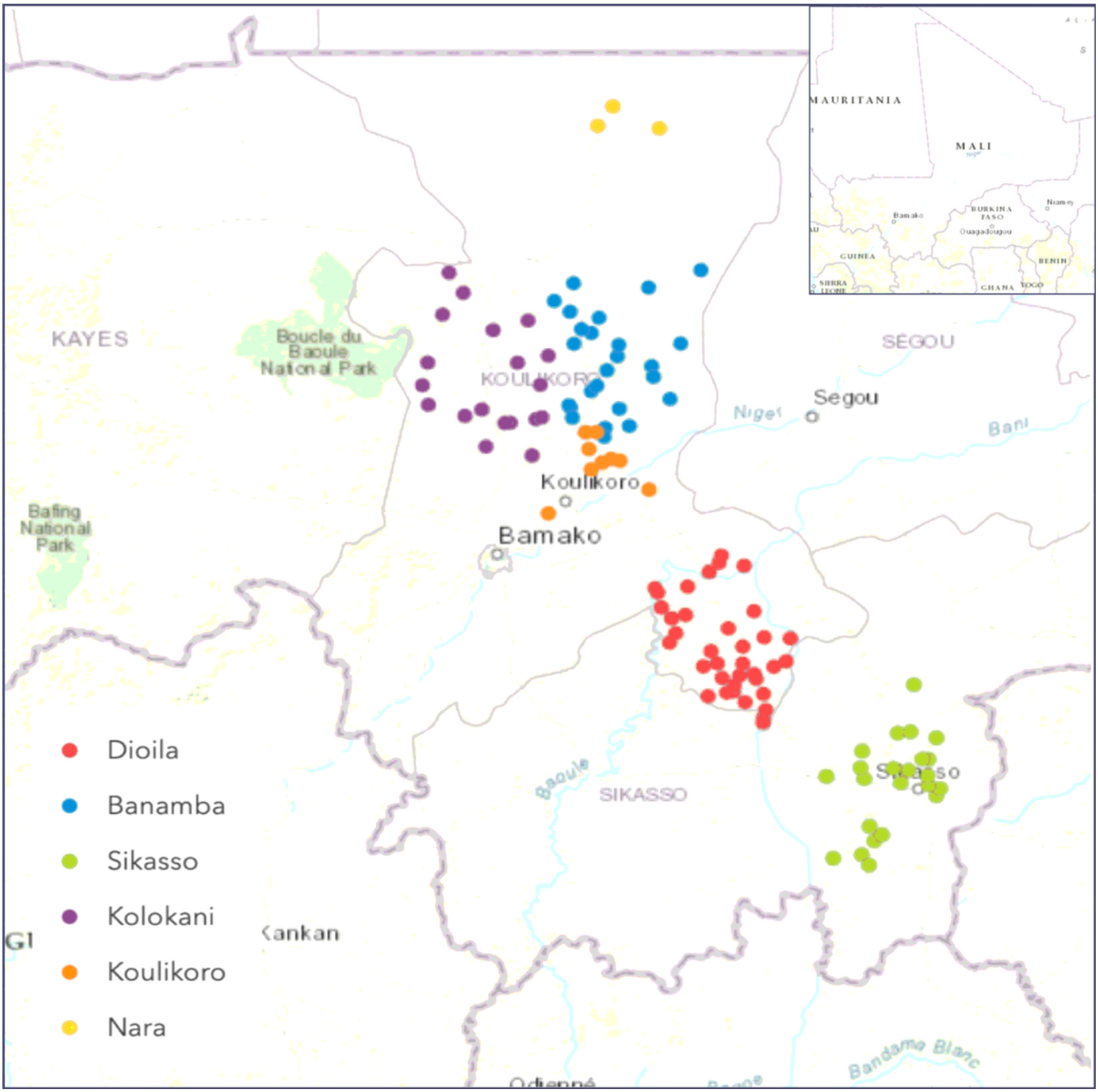

Figure 2

Geographic distribution of sample from 6 districts. Note: This map was created using ArcGIS® Online by Esri (https://maps.arcgis.com). The geographic data (latitude and longitude) were collected by Open Data Kit (ODK) on the portable Android tablets during in-person interviews. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its

authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaltablesmobilephoneinterview20201217.docx