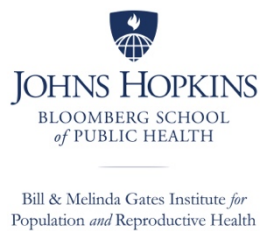Photo: Abigail Greenleaf, PMA2020/Burkina Faso

PMA2020 REMOTE DATA COLLECTION PHASE 2 TECHNICAL REPORT

Comparison of Remote Data Collection Modes to Monitor Family Planning Progress in Burkina Faso: Representativeness, Data Quality, and Cost

February 2019

PMA2020 Technical Report

Title: Comparison of remote data collection modes to monitor family planning progress in Burkina Faso: Representativeness, data quality, and cost

Authors: Abigail Greenleaf, Gadiaga Aliou, Shani Turke, Noelle Battle, Saifuddin Ahmed, Caroline Moreau, Georges Guiella, and Yoonjoung Choi

ACRONYMS

AAPOR	American Association for Public Opinion Research
CATI	Computer Assisted Telephone Interview
CI	Confidence interval
DHS	Demographic and Health Survey
EA	Enumeration area
FP	Family planning
FP2020	Family Planning 2020
FTF	Face-to-face
GNI	Gross National Income
INSD	l'Institut National de la Statistique et de la Démographie
IRB	Institutional Review Board
ISSP	Institut Supérieur des Sciences de la Population
IUD	Intrauterine device
IVR	Interactive Voice Recording
LMICs	Low- and middle-income countries
ODK	Open Data Kit
OR	Odds ratio
NC	Non-contact
PMA2020	Performance Monitoring and Accountability 2020
RE	Resident enumerator
SDP	Service delivery point
SSA	Sub-Saharan Africa
TFR	Total fertility rate
WASH	Water, sanitation and hygiene

ABSTRACT

The rapid expansion of mobile phone ownership provides a platform for low-cost, real-time data collection in low- and middle-income countries. The purpose of this study was to determine which remote data collection mode, Computer Assisted Telephone Interview (CATI) or CATI-to- Interactive Voice Recording (IVR) (referred to as Hybrid IVR), is better suited to monitor family planning progress in Burkina Faso, based on a comparison of sample representativeness, data quality, and cost. In November 2017, using a randomized cross-over design, we called women in Burkina Faso who provided a mobile phone number during Performance Monitoring and Accountability 2020 (PMA2020) data collection in December 2016. Callers who received CATI calls had a higher response rate and was completed by a more representative sample than Hybrid IVR. CATI also had higher reliability and less missing data. This study justifies the use of CATI over Hybrid IVR for population-level remote data collection in West Africa where literacy is low.

TABLE OF CONTENTS

1. INTRODUCTION

The projection that Sub-Saharan Africa (SSA) will be the main source of global population growth in the 21st century [1] has precipitated interest in developing rapid, cost-effective approaches to track family planning (FP) progress. Stakeholders would like to supplement the five-year periodicity of the Demographic and Health Survey (DHS) – the primary reference for reproductive health data in low- and middle-income countries (LMICs) for the past 30 years – with more timely data. In the past several years, organizations and governments have put great effort into collecting and using data more frequently to monitor progress in the 69 Family Planning 2020 (FP2020) countries, at both the global and country level. [2-5] Since 2013, Performance Monitoring and Accountability 2020 (PMA2020) has conducted face-to-face (FTF) surveys every six months to one year in 11 Family Planning 2020 (FP2020) priority countries to provide national or regional estimates of family planning indicators. [6]

However, 59 priority countries do not currently use PMA2020 to monitor family planning progress. Therefore, remote data collection has garnered attention in LMICs as a complement to FTF surveys. The rapid expansion of mobile phone ownership in SSA provides a platform for low-cost, real-time data collection. [7] Urbanization, increased mobile phone network coverage, and decreased cost of a mobile phone have contributed to the proliferation of mobile phone ownership throughout SSA. [8] Greater mobile phone ownership presents the opportunity to survey respondents remotely, whether via Computer Assisted Telephone Interview (CATI) or via Interactive Voice Response (IVR).

In high-income countries, telephone interviews offer certain improvements over FTF interviews, such as faster data collection and lower cost [9-13]. The body of literature on telephone interviews in high-income countries exists in contrast with the dearth of research on remote data collection in LMICs. [14] What literature does exist rarely investigates the representativeness of the sample [15], but rather addresses feasibility and data quality. A systematic review that compared remote and FTF data collection identified only four articles from LMICs that compared FTF with either IVR or CATI for health research. [16] No study has investigated quality of remote data collection modes in SSA specifically among women or for FP measures. [16]

Furthermore, there is limited information on cost of FTF surveys in SSA or LMICs and even less documenting the cost of surveys using remote data collection modes. Evidence is too disparate to draw conclusions; a recent study in Tanzania found that cost per question for a CATI survey was $0.42 (USD), which was not noticeably less expensive than the FTF alternative. [17] In Honduras, a 10-question survey cost $40 per FTF response and only $17 per IVR response. [18] In Ghana, a 2017 IVR 19-question survey cost $4.95 per complete survey, noticeably less than what a FTF survey would cost. [8]

The purpose of this study was to determine which remote data collection mode, CATI or CATI-to-IVR (referred to as Hybrid IVR), is most appropriate to monitor FP progress in Burkina Faso, based on comparison of sample representativeness, data quality, and cost. To do so, we used a

nationally representative survey in Burkina Faso as a baseline to collect demographic and FP data, then followed-up 11 months later by calling women who had mobile phones to administer a shortened questionnaire via both the CATI and Hybrid IVR modes.

2. STUDY POPULATION AND DATA

2.1. Study Setting: Burkina Faso

The study is implemented in Burkina Faso, a West African country with a population of 20 million people. [19] With a total fertility rate (TFR) of 5.7 children per woman, [20] Burkina Faso has one of the highest fertility rates in the world. Although the country is rapidly urbanizing, only 30% of the population is urban. [20] Literacy among female adults is low at approximately 30% but substantially higher among the young (53% among those aged 15-19 years), urban (57%), and wealthy population (63% among the wealthiest quintile). [21] Burkina Faso is sixth to last on the 2015 Human Development Index [22] and has Gross National Income (GNI) per capita of $610, [23] yet over 80% of households own a mobile phone. [24, 25]

According to the latest nationally representative survey data, 30% of women in union and 26% of all women use a modern contraceptive method. [26] Five methods – implants, intrauterine devices (IUDs), injectables, pills, and male condoms – make up 95% of current modern contraceptive use among women who are using modern contraceptives. Approximately 20% of Burkinabe women in union have unmet need for family planning. [26]

2.2. Data

This study used data from two sources: Performance Monitoring and Accountability 2020 (PMA2020) Burkina Faso 2016 (hereinafter referred to as the parent survey), and a cross-over randomized mobile phone survey among women who were identified and enrolled during the parent survey. The primary data source was the parent survey. The mobile phone data were only used to classify respondents' mobile phone survey participation and to assess reliability.

2.2.1. PMA2020 Burkina Faso 2016

PMA2020 is a survey program that monitors and reports FP progress annually. [27] Since 2013, PMA2020 has collected data in 11 countries, led by the Bill and Melinda Gates Institute for Population and Reproductive Health at the Johns Hopkins Bloomberg School of Public Health in collaboration with country research institutions and universities. PMA2020 employs female resident enumerators (REs) who collect the data by conducting FTF surveys with Open Data Kit (ODK) software on Android smartphones. A two-stage, cluster-sampling design is used to select a population-representative sample. Primary sampling units, called enumeration areas (EAs) are selected within sampling strata with a probability proportional to size. Once the EAs are selected, REs map and list all households and a sample of 35 households are randomly selected and interviewed for a household survey in each area. Among selected households, all women aged 15 to 49 years old are eligible to participate in the female survey. Detailed survey methods are available elsewhere. [3]

PMA2020 conducts population-based surveys, in addition to facility surveys, via three primary questionnaires: the household, the female and the service delivery point (SDP) questionnaire. For all surveys, PMA2020 uses a standardized questionnaire. The household questionnaire collects information on household members, assets and livestock, as well as water, sanitation, and hygiene (WASH) conditions of the household. Female surveys collect data on sociodemographic characteristics and FP and reproductive health – primarily contraceptive use, fertility preference, and sexual activity. [3] Typically, the median interview length for female surveys is less than 30 minutes. [28]

In Burkina Faso, PMA2020 partners with the *Institut Supérieur des Sciences de la Population* (ISSP)*.* PMA2020 Burkina Faso has collected six rounds of nationally-representative data since 2014. The surveys have been designed with urban-rural strata and rely on the sampling frame of EAs provided by the *l'Institut National de la Statistique et de la Démographie* (INSD). From November to December 2016, the fourth survey round was conducted, PMA2020 Burkina Faso 2016. Eighty-three EAs were selected (42 rural and 41 urban), with a target sample size of about 3,400 women. All female respondents were asked if they have mobile phones. All women who reported owning a mobile phone were invited to participate in a future survey but were not told about timing or mode of any follow up survey (i.e., FTF versus mobile phone).

2.2.2. Follow-up mobile phone survey
Study participants
The cross-over randomized follow-up phone survey was conducted in November 2017. Among the women who completed the PMA2020 Burkina Faso 2016 survey (n=3,252), we retained women who reported phone ownership (n=1,868), who consented to follow-up (42 women who owned phones did not consent, n=1,826), and who provided a phone number (n=1,766). The women who own phones have different sociodemographic characteristics compared to non-phone owners identified during the parent survey. [29] In particular, phone owners are more likely to be urban, have secondary education or higher, and be in the highest wealth quintile.

Figure 1. Study eligibility flow chart


```
┌─────────────────────────────────────────────────────┐
│         3,414 female surveys - PMA2020 R4 Data        │
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐     ┌──────────────────────────────┐
│         3,252 completed female surveys                │────▶│ 162 women did not complete the│
└─────────────────────────────────────────────────────┘     │ survey                        │
                          │                                   └──────────────────────────────┘
                          ▼
┌─────────────────────────────────────────────────────┐     ┌──────────────────────────────┐
│    1,868 Women with a phone number included           │────▶│ Answered "no" or "no answer" to│
└─────────────────────────────────────────────────────┘     │ phone ownership               │
                          │                                   │ (n = 1,384 )                  │
                          ▼                                   └──────────────────────────────┘
┌─────────────────────────────────────────────────────┐     ┌──────────────────────────────┐
│            1,826 women consented                      │────▶│ Owned phone but non-consented │
└─────────────────────────────────────────────────────┘     │ (n = 42)                      │
                          │                                   └──────────────────────────────┘
                          ▼
┌─────────────────────────────────────────────────────┐     ┌──────────────────────────────┐
│      1,768 women provided phone numbers               │────▶│ Reported phone ownership,     │
└─────────────────────────────────────────────────────┘     │ consented but failed to       │
                          │                                   │ provide a phone number        │
                          ▼                                   │ (n = 58). Reported phone      │
┌─────────────────────────────────────────────────────┐     │ number that was incorrect     │
│            1,766 women randomized                     │     │ (n= 2)                        │
└─────────────────────────────────────────────────────┘     └──────────────────────────────┘
              ╱                      ╲
             ▼                        ▼
┌──────────────────────────┐  ┌──────────────────────────┐
│ ARM A – Hybrid IVR 1st,  │  │ ARM B – CATI 1st, Hybrid  │
│ CATI 2nd                 │  │ IVR 2nd                   │
│ N = 882                  │  │ N = 884                   │
└──────────────────────────┘  └──────────────────────────┘
```

Design

The study objective is to compare the performance of two remote data collection modes, CATI vs. Hybrid IVR, in terms of sample representativeness, data quality, and cost. During IVR calls, respondents use their keypad to answer a pre-recorded question or prompt (e.g. "If yes, press 1. If no, press 2"). CATI calls involves a live interviewer who conducts the survey via a phone call. Hybrid IVR starts with a human operator placing the call, who confirmed the identity and consented the respondent, then explained the IVR survey process before transferring the respondent to the IVR survey. IVR was deemed too challenging for a population with low literacy (30%), so Hybrid IVR was chosen instead, in the hopes of improving data quality and response rates.

Because phone surveys have low response rates, a simple randomization of the women into two groups would not yield an analytic sample size with sufficient power. Thus, we employed a cross-over randomization design. The study was conducted in two rounds so that women could be called twice, with a different mode each round. All women were randomized into two arms, with an equal number of women per language (Moore, French, Gourmantchema, Dioula and Fulfulde)[1] per arm. The order of data collection mode varied by arm: Arm A (CATI first, Hybrid IVR second) or Arm B (Hybrid IVR first, CATI second) (see Figure 1). We assessed distributions of the study participants by background characteristics in each arm, and groups were comparable (see Appendix Table 1).

We calculated sample size by estimating the number of women needed to complete the survey to detect a difference in education between the two modes. We used education, a key

---

[1] Among the enrolled women (n=1766), 94% of women were interviewed for the parent FTF survey using one of the five most common languages: 44% in Moore, 28% in French, 4% in Gourmantchema, 17% in Dioula, and 1% in Fulfulde.

background characteristic, to assess sample representativeness and divided it into three groups – no education, primary education, and secondary education.

Among women that reported owning a mobile phone in the parent survey, 21% had secondary education or higher. We hypothesized that slightly more (25%) CATI respondents and 35% of Hybrid IVR respondents would have secondary education compared to phone owners in the parent study. Previous analyses of phone ownership and education show that educated women are more likely to own a cell phone than less educated women. [29]

We assumed that 70% of women would answer our call and among those that answered, a 80% response rate for CATI and 40% for Hybrid IVR (see Table 1). We further assumed that five percent of women would not consent during the first contact (non-consenters will not be contacted during the second round) and that 20% of women would be lost-to-follow-up in the crossover study. Therefore, we anticipated that 495 respondents would complete the CATI/Hybrid IVR sequence and about 247 would complete the Hybrid IVR/CATI sequence.

With a power of 0.80 and an alpha of 0.05, we needed 231 women per group to detect an absolute difference of 10 percentage points, corresponding to hypothesized $p1=0.25$ and $p2=0.35$. Thus, the expected number of women who completed CATI (n=865: 495+370) and Hybrid IVR (n=432: 247+185) would be adequate for the study.

Table 1. Sample size calculation

	Arm A Hybrid IVR 1st, CATI 2nd (n=882)	Arm B CATI 1st, Hybrid IVR 2nd (n=884)
Number of women contacted for the first cell phone survey (30% non-response)	618	618
Response rate for the first mode	40%	80%
Number of women completing the first MPS	247	495
Number of women contacted for the second MPS (5% refusal, 20% attrition between the two rounds; thus 75% of women who were contacted for first mode will be contacted for second mode)	463	463
Response rate for the second mode	80%	40%
Number of women completing the second MPS	370	185

Note: Different response rates are assumed (40% for Hybrid IVR and 80% for CATI).

Questionnaire
The questionnaires for the CATI and Hybrid IVR surveys were a subset of the FTF survey, with the phrasing of questions adhering to the FTF question as closely as possible. There were 17 key remote survey questions: five introductory questions to identify the respondent, four demographic questions, five questions about the respondent's awareness of modern contraceptive methods, and three questions on contraceptive use. Introduction questions were identical prior to the consent administration for Hybrid IVR and CATI. After consent was

administered, women receiving the CATI were asked if the interviewer could call the participant back should the call be dropped. Women receiving the Hybrid IVR survey heard the interviewer explain what to expect during an IVR survey following consent, and then asked to press one on their keypads. If the respondent was unable to press one, she was unable to participate in the IVR survey, and the interview ended. If the woman pressed one, the respondent's call was transferred, heard recorded instructions about repeating or skipping a question, then was asked to answer an IVR practice question about which country she currently lived in. From this point, the questionnaires for both modes were identical until the last question. The last question in the CATI survey asked the respondent which province she lived in and the last IVR question asked the respondent to enter her age. Appendix 2 presents the mobile phone survey questionnaires.

IVR user testing and pilots
The automated nature of an IVR survey affords the respondent no opportunity to clarify a survey question. Therefore, in July 2017 we conducted IVR "user testing," which consisted of qualitative research that included aspects of cognitive interviewing. The goal of the user testing was to identify any words or questions that were potentially offensive to the participant or any questions that were too difficult or unclear for the respondent to understand. Language specialists, who have previously worked with PMA2020 during FTF surveys, conducted one-on-one interviews with three to four participants per each of the five study languages. Over half of the Gourmantchema language questions had to be re-recorded due to sub-par translation, but the other languages had minimal edits.

In September 2017, working with Viamo, an international social enterprise experienced in IVR, [8] we called randomly generated phone numbers (random digit dial) to complete 100 IVR pilot interviews. We used the information from this pilot to identify any questions with high hang-up or refusal rates and estimated the length of the survey. We also embedded two A/B tests [30] within the survey to find out if two encouraging messages strategically placed in the survey would increase response rates compared to no encouraging messages. Unexpectedly, the response rate was higher for those who did not receive the encouraging message, thus we decided not to include the encouraging message for the follow-up survey (Appendix Table 4A).

The second A/B test examined language order. Respondents were randomized to hear the language options (the first question, which asked respondents the language they wanted to hear the survey) either from most widely spoken language (Order 1: Moore, French, Dioula, Gourmantchema, Fulfulde) or from least widely spoken to most widely spoken (Order 2: Fulfulde, Gourmantchema, Dioula, Moore, French). Approximately 1.5 times more people chose Moore when offered it first compared to when offered it last. A larger effect is seen when Fulfulde was offered first, with 2.5 times more people choosing Fulfulde when it was offered first compared to last (see Appendix Table 4B).

The month before data collection, October 2017, the call center implementing partner Kantar International used an internal list of valid phone numbers to call respondents and complete 10

CATI questionnaires to ensure the script and questionnaire were appropriate. No major changes were made to the script after this pilot.

Implementation

Supervisor and interviewer trainings were conducted in October 2017, and data collection took place in November 2017. Three call center managers and two supervisors oversaw data collection. The managers were in charge of quality control and call center management, whereas the supervisors ensured all materials – such as phones, tablets, and internet – were prepared each day and work stations ready for interviewers.

We trained 20 interviewers and retained 15 for data collection. Interviewer training was conducted for four days, and interviewers conducted practice interviews for two days. We calculated the percent of consenting phone owners (n=1,766) that spoke each of the five languages during the FTF survey then hired interviewers according to the language distribution. Each interviewer was assigned approximately 120 women to call per round. Interviewers were assigned a mode (Hybrid IVR or CATI) for both rounds. Most interviewers were not married (87%), were currently enrolled or had finished university (73%), and were nulliparous (66%). Just under half (46%) of the interviewers had previous survey experience but none of the previous survey experience was with PMA2020.

Viamo developed and designed the CATI and Hybrid IVR software in collaboration with PMA2020. The software ran on a web browser on an Android Tablet and required an internet connection and an E1 telephone line. An E1 telephone line is more reliable and provides better audio quality than a fixed telephone line. Each interview required two phone lines: one for the operator and one for the respondent, connected through the software. The software automatically dialed phone numbers and recorded each attempt. The questionnaire was displayed on the tablet screen, with additional questions appearing when the previous question was answered. The data were stored on Viamo's cloud server and downloaded daily by PMA2020 staff. The questionnaire appeared in French in the software and interviewers translated to the other four languages in real-time, after having developed and refined consistent translations in the local language during the training.

The first and second rounds of data collection took place November 5-17 and November 19-30, 2017, respectively. Participants were called between 8 a.m. and 8 p.m. local time, with interviewers working in three shifts: morning, mid-day, and evening. Women were called up to six times per round, thus a maximum of 12 calls during the study. Respondents were called until they completed or refused the survey or had reached the maximum number of calls. The calls alternated between morning and afternoon: for example, if the first call to a respondent was in the morning, the second call was in the afternoon. At least one of the six calls was on the weekend. Respondents could not specify an exact call-back time unless the participant wanted to be called back within 15 minutes. If a participant refused to consent, she was not called back during that round or the next round. The interviewers used women's first names as recorded during the parent survey to ask for the respondent when talking to the person who picked up the phone, who was sometimes the respondent, but other times someone else. If the person

who answered the call was not the respondent, the interviewer asked that the phone be passed to the respondent, if she was available.

The phone follow-up survey was approved by both the Johns Hopkins University's Institutional Review Board (IRB) and the national ethics committee for health research in Burkina Faso, the *Comité d'éthique pour la recherche en santé*.

2.2.3. Cost data

PMA2020 routinely collects detailed cost data for FTF surveys, and we extended this diligence to the mobile phone survey. For this analysis, we compared cost data for CATI and Hybrid IVR during the mobile phone surveys conducted in 2017. Detailed cost data were collected for each of the two data collection modes throughout the preparation, implementation and analysis phases of the research through the maintenance of project budgets and expense reports from ISSP and Viamo. Key cost categories for comparison include procurement costs, preparatory and pilot costs, training costs, and data collection costs.

3. MEASURES

Interview outcome

To assess key survey implementation outcomes, such as refusal and completion rates, participants have to be classified by call disposition codes in each mode. Using the American Association for Public Opinion Research (AAPOR)'s classifications in conjunction with considerations specific to our study, we defined eight interview call disposition categories (Table 2). Detailed classification was critical to understand and compare performance outcomes between CATI and Hybrid IVR. First, we classified non-contacts into two categories: enrolled women who did not pick up any of the 12 calls (labeled "did not pick-up") and enrolled women where someone picked up the phone, but the desired respondent was not found (labeled "picked up but woman not found"). Women who were identified but refused to participate were classified as refusals. We further distinguished two refusal categories according to whether refusal happened before or at the time of consent. In Hybrid IVR, women were asked to press one on their keypad after consent but before being transferred to the IVR survey. If the woman was unable to press one on her keypad, the survey ended, and she was classified as an "IVR break-off." All other consented women who answered less than 50% of the 17 questions and did not answer the last question were classified as "break-offs". Break-offs post-consent indicates an interest in participating, but the woman was unable to complete 50% of the questions either due to technology difficulties or timing of the call. Finally, a woman is classified as a complete interview if she consented and answered the last question of the survey. Women who did not answer the last question of the survey but consented to the survey and answered more than 50% of questions were labeled as partially completed.

Table 2. Individual Disposition Code Definitions

AAPOR Code	Title	PMA Analysis Group	Definition
(2.20)	Non-contact – NC (didn't pick up)	Eligible	No phone calls were picked up over the 12 attempts
(2.36)	Non-contact – NC (Other) (someone answered the call, but interviewer never spoke with the woman)	Eligible	A phone call was picked up, but the respondent was either unknown or never available to speak to the interviewer
(2.12)	Refusal pre-consent – R	Contact	The respondent refused the study before consent
(2.111)	Refusal – R	Contact	The respondent refused the study during consent
(2.121)	Break-off IVR – R/IVR	Consent	Consented but unable to push 1 on phone (applicable for only hybrid IVR)
(2.12)	Break-off – R	Consent	The respondent consented but answered less than 50%* of the questions and not the final survey question
(1.2)	Partial – P	Consent	The respondent consented and answered more than 50%* of the questions but not the final survey question
(1.1)	Complete – I	Completion	The respondent consented and answered the last survey question

50% of the questions throughout the interview, not necessarily the first half of the questionnaire.

Women's background characteristics

Sociodemographic variables were collected during both the FTF parent survey and the mobile phone surveys. However, for analysis we only used the demographic data collected during the parent survey. The variables include: age (categorized in seven 5-year groups, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49); current union status (in union – i.e. currently married or living with a partner – versus not in union); highest school ever attended (none, primary, or secondary and higher); parity (ever given birth versus never given birth); and residential area (urban versus rural).

In addition, we assessed two household characteristics for which data were collected only during the parent survey: household wealth (wealthiest quintile, middle three quintiles, and poorest quintile) and having electricity (yes or no). Finally, interview language of the FTF parental survey (French, Moore, Dioula, Fulfulde, Gourmantchema and Other) was included in the descriptive analyses.

Mobile phone survey implementation

Finally, we examined practical measures of the two modes: interview time and number of contact days. Interview time is a daily variable that was created by adding up all calls within a day for a respondent, if multiple calls were made. Respondents were to be contacted only once a day, but multiple calls in a day could occur due to bad reception, the call dropping, or other technical difficulties once an initial call of the day was connected. The number of contact days ranges from one to six in each mode. Contact days is woman-level, per mode.

4. ANALYSIS

4.1 Descriptive analysis of the mobile phone survey outcomes

AAPOR offers four essential outcome rates to gauge the validity of a study: response, cooperation, refusal, and contact. The standardized outcomes rates are useful to compare our survey results with the existing literature. For each rate, AAPOR offers multiple definitions. For our study, we selected definitions that align with the study objectives.

Generally, the response rate is the number of interviews – complete or partial – over all attempted respondents. For this study, we use the AAPOR's response rate five, which includes only complete interviews in the numerator (see Table 3). The cooperation rate is similar to the response rate in that the numerator is comprised of complete interviews, but only contacted participants are included in the denominator. AAPOR's cooperation rate three is the minimum cooperation rate. The contact rate measures the proportion of all cases in which some responsible member of the unit was reached by the survey. Thus, contact rate looks similar to the cooperation rate, but the contact rate denominator includes phone numbers where someone answered but the target respondent was not found. The refusal rate is the proportion of all cases in which the respondent refuses the interview, among all attempted respondents. We used AAPOR's refusal rate three. We added a fifth measure, pertinent to our study: IVR break-off. The denominator for this rate includes only women in the Hybrid IVR arm that consented to the survey. The numerator is the number of women who were deemed able by the interviewer to continue the survey (e.g. was able to press one on the keypad).

Table 3. Survey Outcome Rates[+]

Response rate (AAPOR response rate 5): $\dfrac{I}{I+P+R+O+NC}$
Cooperation rate (AAPOR cooperation rate 3): $\dfrac{I}{I+P+R+O}$
Refusal rate (AAPOR refusal rate 5): $\dfrac{R}{I+P+R+O+NC}$
Contact rate (AAPOR contact rate 3): $\dfrac{I+P+R+O}{I+P+R+O+NC}$
IVR Break-off rate*: $\dfrac{R-IVR}{R/IVR+P+I}$ *among those randomized to IVR arm

[+]*disposition codes from Table 2 used in survey outcome rate equations*

4.2 Sample representativeness by mode

To compare sample representativeness between CATI and Hybrid IVR, we assessed differential distribution of background characteristics at four stages of the survey – enrollment, contact, consent, and completion – by mode. Data for each mode were pooled across the two cross-

over arms, and the respondent is the unit of analysis, pooled by mode. All analyses were unweighted, since we randomized all eligible women into two arms, and our study compares outcomes of interest between the two data collection modes.

We first compared demographic characteristics of women at each survey stage against the characteristics at the previous stage, using the chi-square test. The first group is all enrolled women. The second stage is contacting the respondent. Contact is defined as the interviewer confirming the person on the phone is the enrolled participant. There were two causes of a respondent not being contacted – either because the calls were not answered (NC in Table 2) – or someone answered but the study participant from the parent survey was not found (NC - O in Table 2). Women who were identified by the interviewer as enrolled participants were considered contacts regardless of survey interview outcome. We compared characteristics of women who were contacted to characteristics of all study participants enrolled. Third, consent was obtained after a respondent was contacted. The fourth stage, completion, consists of women who answered the last survey question (I in Table 2), and all other women were considered non-completers.

We conducted multivariable logistic regressions to generate odds of completion by sociodemographic characteristics, among all enrolled participants. We did not include marital status or parity in the multivariable model because these characteristics were not related to consent or completion in the bivariate analyses. Electricity was also excluded because of high correlation with wealth quintiles.

4.3 Data quality by mode

To compare data quality between the two modes, we assessed the amount of missing data and reliability. Again, data are pooled from the two cross-over arms by mode. First, we calculated the percent of answers missing, and the unit of analysis is the response.

Second, we assessed reliability in responses between the FTF parent survey and each of the two mobile phone survey modes by calculating the kappa statistic, focusing on responses for basic demographic characteristics that are unexpected to change over 11 months for a majority of women: marital status, ever attending school, and ever given birth. In addition, responses were compared for awareness of five selected contraceptive methods. Kappa ranges from -1 to +1 and accounts for some agreement being caused by chance. It is used to calculate inter-rater reliability, and in this study, inter-mode reliability. [31] We interpreted the kappa values using the standard of Landis and Koch. A kappa around 0.80 will be almost perfect agreement, between 0.60 and 0.80 substantial agreement, and between 0.40 and 0.60 moderate agreement. [32] Using STATA, a statistical software, we calculated kappa and reported p-values to determine whether the agreement was due to chance.

Finally, we assessed reliability of reported age, which commonly has measurement errors and bias in quantitative surveys. We compared reported age during the FTF survey to age reported during CATI. Because the mobile phone surveys were conducted 11 months after the FTF survey, women were expected to have reported an additional year of age. During the Hybrid

IVR survey, women provided age twice: once to the interviewer at the beginning of the survey, and once at the end of the survey as they recorded their age by pressing the two numbers on the keypad that corresponded to their age. We additionally compared the two values.

4.4 Cost by mode

Survey implementation costs were compared as total cost per mode and cost per completed interview. Itemized costs were summarized by category of cost (procurement, preparatory and pilot, training, and data collection) and by mode. Many itemized costs were equivalent between the two modes. Assumptions were made to adjust the duration of training and data collection for each mode as if they had been conducted as stand-alone activities in order to determine a more accurate cost per mode. Specifically, we assumed each of the two modes used four supervisors and 15 interviewers, training was three days for supervisors and four days for interviewers, and that it would take 10 days to call all 1,766 women, including up to six call-backs per mode. Most costs were incurred in local currency, the West African CFA franc. Costs in CFA were converted to the US dollar (USD) using the 2017 average exchange rate of 1 USD: 582 CFA. We did not include the cost of software development in the fieldwork cost comparisons. As a joint software solution was created to meet the needs for the study for both CATI and Hybrid IVR and therefore the costs for software specific for each mode were not estimated,[2] the assumption was made that the cost of building the software was the same for both modes.

Descriptive analyses were conducted to compare costs by data collection mode on a total and cost-per-interview basis.

5. RESULTS

5.1. Individual call disposition outcomes and key survey rates by mode

For Hybrid IVR, 45% of the 1,700 women[3] contacted never spoke to an interviewer and were classified as non-contacts. Just over a quarter of all women (27%) did not answer any of the six Hybrid IVR phone calls (Table 4). A slightly lower percentage (18%) answered the call but the respondent was unreachable (the person who answered the phone call did not know the woman or the woman was not available when the phone call was answered). Among the 55% who were contacted by the interviewer, 10% refused to participate, with the majority of refusals (9%) taking place before the interviewer read the consent script. Almost 3% of women consented to the survey but were unable to press one on their keypad, thus were not transferred by the interviewer to the IVR survey. Almost a quarter of women (23%) started the survey but broke off before completion, and only 19% of the 1,700 women completed the survey.

---

[2] We paid Viamo $25,000 to create the software that allowed a live interviewer to start the survey then transfer the respondent to IVR (i.e. Hybrid IVR). The cost of building the CATI-only additions to the software was a few thousand dollars on top of the cost of building the Hybrid IVR software.

[3] Women in Arm A (CATI first, Hybrid IVR second) who refused the survey or were not identified by the person who answered the survey call were not called during the second round. Women who did not answer any of the six first-round CATI calls were contacted, however, in the second round.

Table 4. Call disposition codes by mode for both data collection rounds, defining complete as answering the last question of the survey

		Hybrid IVR		CATI	
		N	%	N	%
NC	Non-contact (didn't pick up)	464	27.3	452	27.1
O	Non-contact (picked but woman not found)	310	18.2	244	14.6
PR	Refusal pre-consent	160	9.4	124	7.4
R	Refusal	19	1.1	12	0.7
IVRBP	IVR Break-off (consented but unable to push 1 on phone)	46	2.7	N/A	N/A
BP	Partial + breakoff	382	22.5	38	2.3
I	Complete (answered the last question)	319	18.8	798	47.8
	Total	1700		1668	

Among the 1,668 CATI respondents,[4] a similar percentage were non-contacts (42%) compared to Hybrid IVR respondents: 27% of women did not answer and 15% of respondents' phone numbers were answered by someone other than the respondent and the female respondent could not be located. Fifty-eight percent of respondents were contacted by the interviewer. Overall refusal was slightly lower than Hybrid IVR (8%), but again the majority of refusals took place before consent (7%). Only 2% of respondents consented and did not complete the survey; break-off was low. In total, 48% of women completed the CATI survey.

Table 5 presents response, cooperation, refusal and contact rates. The response rate (19%) and cooperation rate (34%) were substantially lower in Hybrid IVR than CATI (48% and 83%, respectively), due to the small number of complete Hybrid IVR surveys, as presented in Table 4. The refusal rates were comparable between the two modes (19% for Hybrid IVR, 14% for CATI) as was the contact rate (55% for Hybrid IVR, 58% for CATI). The IVR transfer rate, which is the percent of women who were successfully transferred to Hybrid IVR by an interviewer, was 93%.

---

[4] The same rules around second-round calls explained above for Arm A apply to women in Arm B. However, refusal rate was higher among women randomized to Arm B (Hybrid IVR first, CATI second) thus fewer women were contacted for CATI across the two rounds.

Table 5. Key survey rates by mode

Outcome Rates	Hybrid IVR	CATI
Response Rate		
I / (I+BP+R+PR+NC+O)	18.8%	47.8%
Cooperation Rate		
I / (I+BP+R+PR)	34.4%	82.7%
Refusal Rate		
(R+PR) / (I+BP+R+PR)	19.3%	14.0%
Contact Rates		
(I+BP+R+PR) / (I+BP+R+PR+NC+O)	54.5%	57.6%
IVR Transfer Rate		
(BP + I) / (IVRBP + BP + I)	93.8%	N/A

5.2. Sample representativeness at contact, consent, and completion by mode

5.2.1. Characteristics of study participants

We expected and found that the distribution of sociodemographic characteristics was similar among Hybrid IVR and CATI respondents, because this was a cross-over randomized study (see Appendix Table 1). Among all eligible women (n= 1,766), the average age for both modes was 28 years. Two-thirds of women lived in rural areas, and as many were currently in a union. About 40% had attended secondary school or higher. A third of women reported current use of a modern contraceptive. Forty-three percent of women completed the FTF survey in Moore, followed by 28% of women in French, and 17% in Dioula.

Only 8% of women belonged to the lowest wealth quintile households in the country and 54% were in the highest quintile. Slightly less than two-thirds of women lived in households with electricity, and 94% of women lived in a household with a mobile phone.

5.2.2. Characteristics of respondents at contact, consent and completion by mode

In both survey modes, respondents who were from urban areas, over 20 years old, and educated accounted for a greater proportion of the sample after contact and consent. However, while attrition between consent and completion is not associated with background characteristics in the CATI arm, those who completed Hybrid IVR have statistically significant different background characteristics from those who consented, introducing further selection error in Hybrid IVR compared to CATI. Changes in the sample characteristics are described in detail below, by mode.

<u>Hybrid IVR</u>

Compared to the 1,700 women called for the Hybrid IVR study, the 926 women who were successfully contacted by the interviewer had a statistically significant different age distribution (see Table 6A): women aged 15-19 and 20-24 were less represented in the contacted group compared to the eligible study participants. Urban women were less likely to be contacted than rural women. There was no difference in marital status between the two groups. Slightly more women reported using contraception in the contacted group (36% versus 32%), but the difference was not statistically significant. The dominant language groups – Dioula, French and Moore – comprised a greater percentage of the language distribution in the contacted group than in the study population.

Table 6. Comparison of background characteristics of study participants at contact, consent, and completion by mode

6A. Hybrid IVR

	% distribution among total study participants	% distribution among study participants who were identified		% distribution among study participants who consented		% distribution among study participants who completed	
Total	n=1700	n=926		n=740		n=319	
Age			**				**
15-19	18.8	13.3		12.8		15.7	
20-24	19.8	16.1		16.9		20.7	
25-29	19.6	21.1		22.3		21.9	
30-34	15.1	19.1		18.7		20.4	
35-39	11.6	13.3		12.3		9.4	
40-44	9.2	9.9		10.1		7.8	
45-49	5.9	7.2		7.0		4.1	
Urban/rural			**		*		*
Rural	32.9	24.4		26.2		20.7	
Urban	67.1	75.6		73.8		79.3	
Marital status							**
Currently not in union	35.1	33.2		31.8		39.2	
Currently in union	64.9	66.8		68.2		60.8	
Highest school attended			**				**
Never	41.3	37.4		37.3		27.3	
Primary	20.1	21.1		21.4		17.6	
Secondary or higher	38.6	41.6		41.4		55.2	
HH Wealth (quintile)			**				
Lowest	8.3	5.2		5.8		3	**
Lower	9.2	6.8		7.0		5	
Middle	12.2	10.3		10.4		7	
Higher	16.6	16.4		16.7		14	
Highest	53.7	61.3		60.0		70	
Parity			**				**
Yes	70.4	74.6		74.7		67.6	
No	29.6	25.4		25.3		32.3	

	% distribution among total study population	% distribution among study participants who were contacted		% distribution among study participants who consented		% distribution among study participants who completed	
HH Electricity			**				**
Yes	62.8	67.9		66.6		73.4	
No	37.2	32.1		33.4		26.6	
Language			**		**		**
Dioula	16.8	19.1		16.4		14.7	
French	28.5	29.8		29.9		37.3	
Fulfulde	1.2	0.5		0.7		0.3	
Gourmantchema	3.9	3.4		3.8		1.6	
Moore	43.9	44.3		46.2		44.2	
Other	5.7	2.9		3.1		1.9	

Comparison is with the previous column, using chi-squared test: i.e., all participants vs. contacted; contacted vs. consented; and consented vs. completed
* p-value 0.01-0.05, ** p-value less than 0.01

6B. CATI

	% distribution among total study population	% distribution among study participants who were contacted		% distribution among study participants who consented		% distribution among study participants who completed	
Total	n=1668	n=972		n=833		n=798	
Age			**				
15-19	19.4	13.0		13.1		13.0	
20-24	20.1	18.4		19.0		18.5	
25-29	19.1	21.8		22.1		22.3	
30-34	15.4	17.8		17.9		18.1	
35-39	11.5	11.8		10.7		10.7	
40-44	8.8	10.3		10.3		10.5	
45-49	5.8	6.9		7.0		6.9	
Urban/rural			**		**		
Rural	32.3	25.7		27.2		27.3	
Urban	67.7	74.3		72.8		72.7	
Marital status							
Currently not in union	35.1	34.2		34.3		34.2	
Currently in union	64.9	65.8		65.7		65.8	
Highest school attended							
Never	40.8	37.2**		38.3		38.2	
Primary	20.0	21.7		21.4		21.8	

Secondary or higher	39.3	41.1		40.3		40.0	
HH Wealth (quintile)			**				
Lowest	8.0	5.5		6.0	*	6.3	
Lower	8.9	7.0		7.3		7.4	
Middle	12.3	10.8		11.2		11.4	
Higher	17.0	16.5		17.4		16.9	
Highest	53.9	60.4		58.1		58.0	
Parity			**				
Yes	70.0	73.1		72.3		72.2	
No	30.0	26.9		27.7		27.8	
HH Electricity			**				
Yes	63.2	68.6		67.1		67.0	
No	36.8	31.4		32.9		33.0	
Language			**				
Dioula	17.2	17.6		17.3		17.5	
French	28.7	31.0		29.9		29.2	
Fulfulde	1.1	0.7		0.7		0.8	
Gourmantchema	3.8	3.0		2.9		2.9	
Moore	43.4	44.2		45.5		46.1	
Other	5.8	3.5		3.7		3.5	

Comparison is with the previous column, using chi-squared test: i.e., all participants vs. contacted; contacted vs. consented; and consented vs. completed
* p-value 0.01-0.05, ** p-value less than 0.01

Seven hundred forty women consented to the Hybrid IVR survey. Age distribution, marital status, educational attainment, modern contraceptive use, parity, and household electricity were similar across the contacted and consented groups. Women who consented were slightly less likely, but still statistically significant, to be urban than the women contacted and had a different language distribution than contacted women.

Finally, only half of the women who consented ultimately completed the Hybrid IVR survey (n=319). There were a number of differences in the sociodemographic make-up of women who completed the survey compared to the women who consented. A greater percentage of women who completed the survey were younger, from rural areas, and not currently in a union. A smaller percentage of women had ever given birth among those who completed the survey than among those who consented. In total, 55% of those who completed the survey had attended secondary school or higher education, compared to only 41% of those who consented. All differences presented above were statistically significant. A greater percentage of those who completed the survey had electricity compared to those who consented.

CATI
The differences in sociodemographic distribution patterns among enrolled and contacted CATI participants (n=972) were similar to observation under Hybrid IVR. Compared to those in the study population, women who were contacted were younger and more rural (see Table 6B).

A total of 139 women who were contacted refused to participate in the study. Among the 833 who consented, there were very few differences in the distribution of characteristics compared to women who were contacted. In total, 96% of women who consented completed the CATI survey (n=798). There were no statistically significant differences in the distribution of sociodemographic characteristics of those who consented compared to women who completed the survey in its entirety.

Odds of completing the surveys: multivariable logistic regression by mode
To estimate the odds of completing by sociodemographic characteristics, we conducted multivariable logistic regression. In Hybrid IVR, the odds of completion were greater among women aged 25-35 years, compared to women aged 15-19 years (Table 7). Having a secondary or higher education was associated with 2.3 increased odds of completing the survey (95% confidence interval (CI): 1.6 – 3.2) compared to women with no education. Women in the highest wealth quintile had higher odds of completion compared to women in all other wealth quintiles (odds ratio (OR) 1.6, 95% CI: 1.1 – 2.3).

In CATI, all age groups had greater odds of completion than women aged 15-19 years. Having primary education and secondary education both increased the odds of completing the survey, but with a lesser effect than Hybrid IVR (primary education OR: 1.4, 95% CI: 1.1 – 1.9; secondary education OR: 1.4, 95% CI: 1.1 – 1.8).

Table 7. Adjusted odds ratio of completion by background characteristics by mode among all enrolled women: multivariable logistic regression analyses

	Hybrid IVR (n=1,700)	CATI (n=1,668)
Age group		
15-19 (reference)		
20-24	1.4 (0.90 – 2.1)	**1.7 (1.2 – 2.3)**
25-29	**1.7 (1.1 – 2.6)**	**2.8 (2.0 – 4.0)**
30-35	**2.6 (1.7 – 4.0)**	**3.0 (2.1 – 4.3)**
35-39	1.2 (0.7 – 2.0)	**1.9 (1.3 – 2.8)**
40-44	1.5 (0.88 – 2.6)	**3.4 (2.2 – 5.2)**
45-49	1.1 (0.60 – 2.3)	**3.4 (2.1 – 5.5)**
Residential area		
Rural (reference)		
Urban	1.1 (0.7 – 1.6)	1.3 (1.0 – 1.8)
Highest school attended		
No education (reference)		
Primary	1.2 (0.8 – 1.8)	**1.4 (1.1 – 1.9)**
Secondary or more	**2.3 (1.6 – 3.2)**	**1.4 (1.1 – 1.8)**
HH Wealth (quintiles)		
Lowest (reference all other groups)	0.6 (0.32 – 1.2)	0.74 (0.50 – 1.1)
Highest (reference all other groups)	**1.6 (1.1 – 2.3)**	1.0 (0.76 – 1.3)

5.2.3. Survey implementation para data by mode

<u>Average interview time among women who completed the survey</u>

CATI completed interviews were, on average, 6 minutes shorter than the Hybrid IVR completed interviews (see Table 8). The mean length of a completed CATI was 6 minutes and 34 seconds (n=798, range 3.5 minutes to 15 minutes, 48 seconds). The mean length of a completed Hybrid IVR call was 12 minutes and 54 seconds (n=319). The minimum Hybrid IVR length was 8 minutes, 12 seconds and the maximum length was 21 minutes and 48 seconds.

Table 8. Interview time by mode among women who completed survey

	Hybrid-IVR (n=319)	CATI (n=798)
Interview time (minutes) (adding up interview time of calls, as long as interviewers talked to the women)	12 minutes, 54 seconds (SD:2 minutes, 31 seconds)	6 minutes, 34 seconds (SD: 1 minute, 54 seconds)

SD: Standard deviation

5.3 Data quality by mode

Missing values among women who listened to the survey question

Missing values were common during the Hybrid IVR survey. Only 68% of women answered the residential area question, but responsiveness slightly increased throughout the demographic questions, with 71% answering about school attendance and 73% answering about marital status (see Table 9). Over 70% of women answered the five ever-use questions on contraceptive use, and 71% answered the current use question. Only 64% of women who reported using a current method specified which method she was using. Women who were not currently using a method were asked if they were currently pregnant. Ninety seven percent of women who heard the question answered the current pregnancy question.

Among women who listened to a question during the CATI, the only missing values were associated with the last question, which was only asked of women who were not using a modern form of contraception; only 3% of women did not answer the question about being pregnant.

Table 9. Comparison of complete responses among women who listened to the question by mode

	Hybrid IVR	CATI
Background characteristics questions		
Residential area	376/557 (68%)	813/813 (100%)
Ever attended school	380/537 (71%)	805/805 (100%)
Marital status	386/525 (73%)	807/807 (100%)
Family planning questions		
Ever heard of implants	355 /510 (70%)	808/808 (100%)
Ever heard of injectables	354/ 499 (71%)	805/805 (100%)
Ever heard of pills	350/ 494 (71%)	807/807 (100%)
Ever heard of condoms	353/ 489 (72%)	805/805 (100%)
Ever heard of IUD	352 / 481 (73%)	805/805 (100%)
Current use	338 / 477 (71%)	806/806 (100%)
Method, if using currently	221/345 (64%)	398/398 (100%)
Are you pregnant, if not using	109/112 (97%)	364/376 (97%)

Reliability of key survey questions between FTF, CATI and Hybrid IVR

The highest kappa for FTF and Hybrid IVR (0.68) was for the question about ever attending school (see Table 10a). Similarly, kappa for marital status was 0.66. All five kappa values for the 'ever heard' questions were close to 0, meaning there was no correlation between the FTF and Hybrid IVR responses.

All kappa values for responses between FTF and CATI surveys were higher than the kappa values comparing Hybrid IVR and FTF responses. 'Ever given birth' had the highest kappa value (0.87) and education the second highest (0.84). These measures can be considered reliable.

Table 10a. Comparison of responses between FTF and each of the two modes: percent agreement and Kappa statistics

	Hybrid IVR		CATI	
	% agreement	Kappa	% agreement	Kappa
Marital status	84%	0.66**	87%	0.70**
Education	87%	0.68**		0.84**
Ever given birth	79%	0.54**	95%	0.87**
Ever heard of implants	79%	-0.03	89%	0.12**
Ever heard of injectables	76%	0.03	88%	0.12**
Ever heard of pills	87%	0.07	93%	0.22**
Ever heard of condoms	85%	0.03**	92%	0.23**
Ever heard of IUD	69%	0.19**	67%	0.23**

*p-value 0.01-0.05
**p-value less than 0.01

Marital status was slightly less reliable (0.70) and the 'ever heard' questions were unreliable with kappa values ranging from 0.12 to 0.23. Among women who answered an 'ever heard' question during the IVR survey, between 11% (pills) and 22% (injectables) changed their answer from 'yes' in the FTF survey to 'no' during the IVR survey (see Table 10b). Fewer women changed their answers from the FTF to CATI survey: between 4% (pills) and 17% (IUD) changed from 'yes' to 'no'. In both surveys, a sizable percentage of women who responded to the question changed their answer from 'no' during FTF to 'yes' during the remote mode. In the Hybrid IVR survey, the number of women that changed a 'no' to 'yes' answer was quite small (sample size of approximately 10) but the results nonetheless mirrored the patterns seen in CATI: over 60% of women who answered 'no' to an FTF contraceptive knowledge question answered 'yes' during the follow-up survey.

Comparing reported age in CATI and the FTF survey, overall, 75% of women reported an age during CATI +/- 2 years compared to the age in FTF (see Figure 2a). Among women with no education who answered the CATI age question, 61% reported an age +/- 2 years of the FTF age. The reliability of age increased with education, with 78% of women with primary education and 85% of women with secondary education reporting an age during CATI +/- 2 years of age.

In Hybrid IVR, 60% of women provided the same age as that recorded by the interviewer and by using her keypad (see Figure 2b). There was no discernable pattern for the 40% that did not report the same age at the two time points. The mean difference for the 296 women that answered both age questions during the Hybrid IVR survey was 5.4 years, with a standard deviation of 18.0 years. Difference in age values ranged from -71 to 50 years. As education increased, the reliability of age increased (see Figure 2c).

Figure 2a. Comparison of age responses between face to face parental survey and CATI follow-up survey (n=1,677)
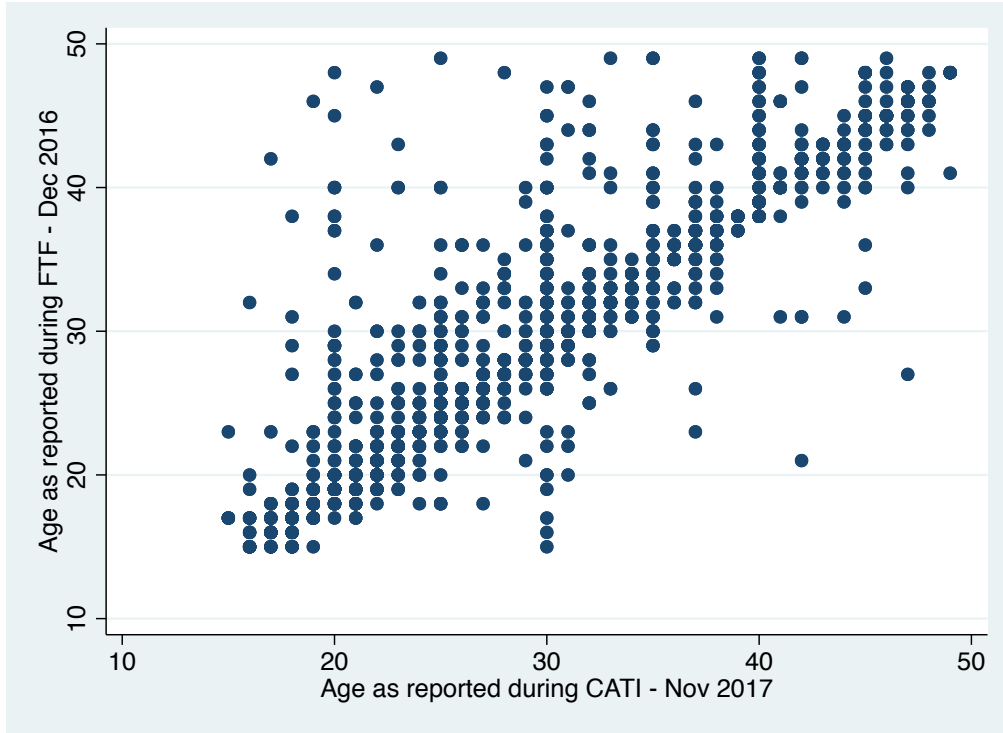
Figure 2b. Comparison of two age responses in Hybrid IVR: Age as reported to interviewer during introduction, then age as entered via keypad at the end of the survey (n=296). All women.
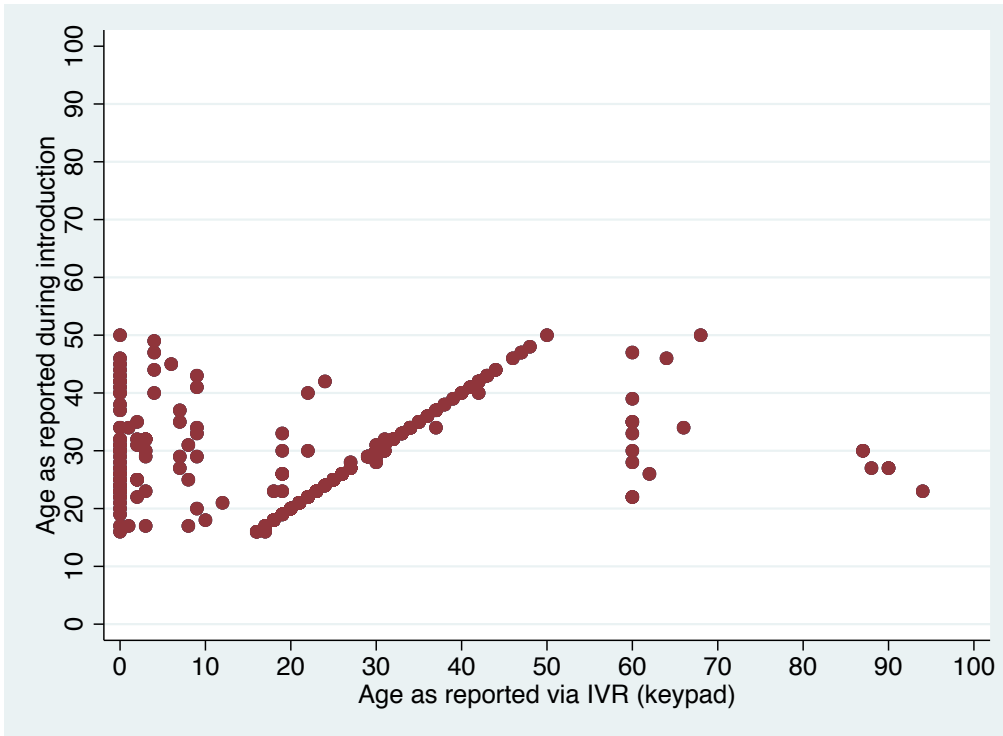
Figure 2c. Comparison of two age responses in Hybrid IVR by education level: Age as reported to interviewer during introduction, then age as entered via keypad
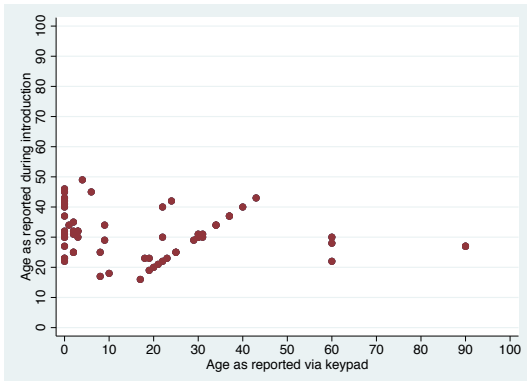
Among women with no education (n= 63)
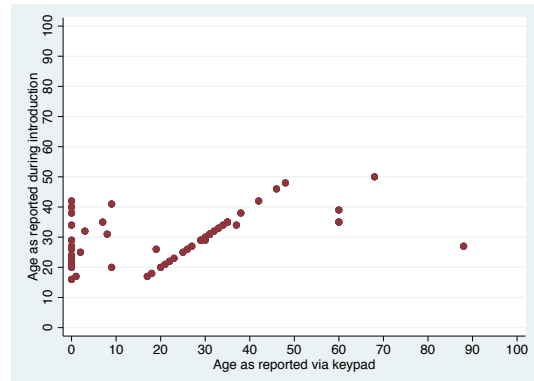
Among women with primary education (n=54)
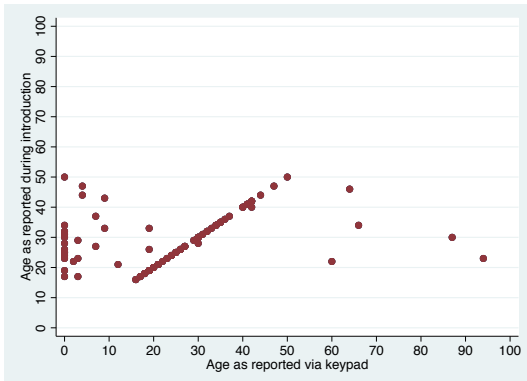
Among women with secondary education, 1st cycle (n = 105)
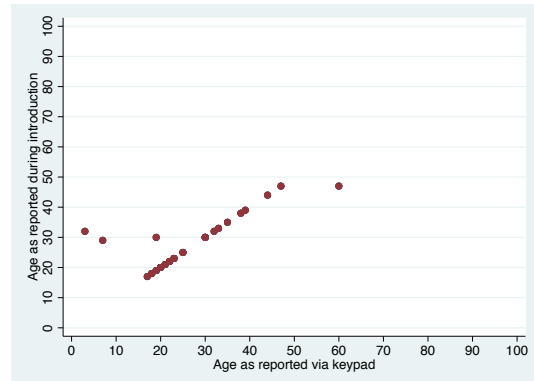
Among women with secondary education, 2nd cycle (n = 35)
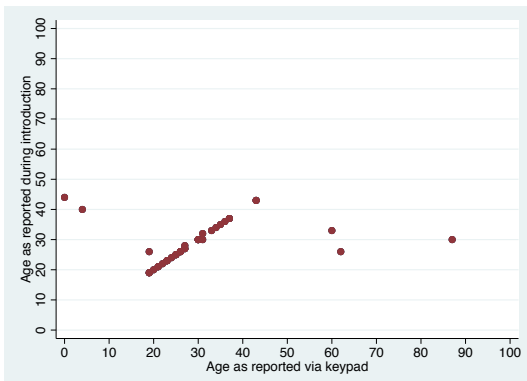
Among women with tertiary education (n= 39)

Table 10b. Contraceptive knowledge: % change in answers from FTF to Hybrid IVR or CATI survey, by direction of change

	FTF	Hybrid IVR		CATI	
	% yes during FTF survey (n = 1,766)	% yes- FTF to no[+] – IVR	% no FTF to yes[*]- IVR	% yes FTF to no[+] – CATI	% no FTF to yes[*] – CATI
Ever heard of implants	95%	19% (n = 57)	90% (n = 9)	8% (n = 64)	71% (n = 35)
Ever heard of injectables	93%	22% (n = 67)	71% (n = 12)	9% (n = 67)	73% (n = 40)
Ever heard of pills	94%	11% (n= 33)	75% (n = 12)	4% (n = 33)	72% (n = 36)
Ever heard of condoms	95%	15% (n = 47)	30% (n = 10)	5% (n = 40)	68% (n = 41)
Ever heard of IUD	59%	18% (n = 40)	64% (n = 90)	17% (n = 87)	62% (n = 291)

+ among women who responded YES during the FTF survey
*among women who responded NO during the FTF survey

5.3 Cost

Overall costs were less for CATI than for Hybrid IVR. CATI overall costs were $41,235; Hybrid IVR overall costs were $45,793 (see Table 11). Cost per completed response is significantly higher for Hybrid IVR than CATI given the much lower completion rate: 319 complete Hybrid IVR surveys versus 798 complete CATI surveys. The cost per completed response is $143.55 for Hybrid IVR and $51.67 for CATI.

Table 11. Fieldwork Costs in US Dollars

	CATI	Hybrid IVR
Procurement *Laptops, tablets, cell phones, headphones, headphone spliters, t-shirts*	15,312	15,312
Preparatory Work *Pre-testing, piloting, translation, compensation to pilot participants*	302	4,168
Training costs *Per diems, facility costs, meals, printing, staff salary*	5,232	5,232
Data collection *field staff, travel, infrastructure*	18,390	18,390
Airtime *Cost of calls, respondent compensation*	1,999	2,691
TOTAL	**41,235**	**45,793**
Number of complete surveys	798	319
Cost per completed survey	51.67	143.55

Costs were further analyzed by category. Procurement costs of $15,312 were assumed to be equivalent for the two survey modes. These costs included laptops for call center managers, tablets, mobile phones, and headphones for data collectors and t-shirts for the call center team. Hybrid IVR had higher preparatory costs (CATI: $302 versus Hybrid IVR: $4,168), driven by the need to utilize a professional studio to record the survey in five languages for the IVR portion of the survey. Training costs ($5,232) were also the same for Hybrid IVR and CATI. This assumed the same duration of training would be needed for either mode if training had been done independently. Training costs included per diems for trainers, supervisors, and interviewers as well as facility costs, internet, and catering during the trainings.

Data collection costs, apart from airtime for calls during data collection, were equivalent between the two modes ($18,390). These included field staff pay, infrastructure costs for the call center, and costs for technical support from Viamo. Airtime costs were higher for Hybrid IVR ($2,691) than CATI ($1,999) even though there were half as many completed Hybrid IVR calls as CATI calls. The increased cost is due to Hybrid IVR interviewers' detailed explanation about how to navigate an IVR survey, and the IVR questionnaire took respondents longer to answer than with CATI. The higher cost of airtime and preparatory work and the minimal number of complete surveys (Hybrid IVR: 309 complete surveys) led to the higher cost per completed survey for Hybrid IVR as compared to CATI.

6. DISCUSSION

This study describes the population that completed remote surveys and the differences in sample representativeness and data quality between Hybrid IVR and CATI completers. The response and cooperation rates for CATI (48% and 83%, respectively) were 2.5 times the Hybrid IVR rates. Both modes saw high non-contact rates: 27% of eligible women did not answer interviewers' phone calls, and approximately 15% of phone numbers called were answered by someone other than the respondent and the respondent subsequently was not found. The refusal and contact rates were comparable across the two modes as expected, considering randomization and comparable questions up to consent in both modes. For both modes, respondents who were from urban areas, over 20 years old, and educated accounted for a greater proportion of the sample after contact and consent, implying lower contact and consent rates among their counterparts.

Similar to cooperation and response rates, completion rates were higher for CATI: 48% in CATI and 18% in Hybrid IVR. While attrition between consent and completion is not associated with background characteristics in CATI, those who completed Hybrid IVR have statistically significant differences in background characteristics from those who consented, introducing further sample distortion in the Hybrid IVR arm specifically by education, compared to CATI. Women who were younger, from an urban area, had a secondary or higher education, and spoke French were more likely to complete the Hybrid IVR survey compared to their counterparts. Interestingly, report of modern contraceptive use during the PMA 2016 parent survey was the same among Hybrid IVR and CATI completers (34% and 33%, respectively), and did not significantly differ between the overall eligible participants (32%) and women who

completed either survey. The multivariable logistic regression for Hybrid IVR showed the odds of completing the Hybrid IVR survey were 2.4 higher for women with a secondary education than for women with no formal education. Women responding to CATI had reduced odds of completion if they were younger than 20 years or belonged to the lowest wealth quintile households. Education was also statistically significantly associated with completion in CATI, but to a lesser effect than among Hybrid IVR participants.

All data quality metrics were lower for Hybrid IVR than for CATI. Most questions saw a substantial amount of missing values for Hybrid IVR, whereas missing values were rare for CATI. For all analyzed questions, reliability was higher for CATI than for Hybrid IVR. Reliability was higher for both modes among demographic questions than the contraceptive awareness questions. Neither of the two modes (CATI or Hybrid IVR) is particularly reliable across all questions, but less than optimal consistency can be expected, considering the reporting inconsistencies in longitudinal studies. [33] Finally, interview time was six minutes longer on average for Hybrid IVR than for CATI. The brevity of our survey (20 questions) meant that the CATI survey was often finished before the Hybrid IVR interviewer had completed the explanation of how the IVR survey functions.

The increased cost for the Hybrid IVR design (for recording the IVR survey in multiple languages and the additional airtime to complete the survey), combined with the lower completion rates also make the Hybrid IVR design unfavorable from a cost perspective.

Comparing our results to the existing research, we find both similarities and differences. A study from Lebanon [34] that followed up FTF respondents with CATI had higher kappa scores for their study's questions, which is not surprising since the population in Lebanon has a higher education level than women in Burkina Faso. The study in Lebanon did not look at the representativeness of the CATI sample, but a World Bank study conducted in Honduras and Peru enrolled respondents in FTF and followed up with them via IVR and CATI. [18] Regardless of mode, attrition was highest among less educated, less affluent, older, and rural participants. [16] Similar to our results, reliability was higher between CATI and FTF answers than IVR and FTF answers.

This study helps identify at what point the participant profile distorts in a remote data collection survey (i.e. contact, consent, or completion). This study also identifies which mode produces more reliable data among women of reproductive age. We conclude that CATI is the superior choice, due to the ease of implementation, sample representativeness, data quality, and cost.

Our study identifies the profile of those who complete remote data collection surveys, comparing those who complete compared to those who do not complete, among those enrolled during a FTF survey 11 months earlier. Few, if any, nationally representative FTF studies have followed up with respondents with a remote data collection mode in an LMIC. Furthermore, while the profile of non-respondents is often inferred from census data or other sources, our study design allowed us to construct the profile of respondents who did not

complete the mobile phone survey by using the parent survey data. Therefore, the extent of non-response error is quantifiable, as is reliability of the two remote modes, by comparing select measures to those in the FTF survey. Another strength of our study is the comparison of two remote data collection mode costs. There is a paucity of information about data collection cost in SSA, thus this study's systematic collection and comparison of costs between the two modes fills a gap in the literature.

There are limitations to this study. The interval between the FTF survey and the phone follow-up was lengthy; 11 months. Research shows that the shorter the period between first contact and follow-up, the lower the attrition. Therefore, following up with respondents at an earlier time would increase the contact rate, perhaps resulting in a more representative study population. Generalization of results to other populations is somewhat limited, considering that we surveyed only women, and literacy in Burkina Faso is low among women. Nevertheless, the CATI results are more reliable than the Hybrid IVR results. Furthermore, CATI is more feasible since CATI can be more easily implemented in a country with high language fractionalization and Hybrid IVR, or even simple IVR, is best implemented in a country that has approximately five or fewer main languages. The assumptions in the cost analysis are delineated in the methods section of this report; however, the ability to generalize these costs results to other settings is limited. Costs to set up and operate a call center may vary greatly across countries. Burkina Faso is a country without existing call center infrastructure, but many other SSA countries such as Nigeria and Kenya have call center operations that could be utilized for similar studies. Also, the software we used was quite expensive due to the necessary infrastructure (particularly the internet) as well as using two phone lines (one for interviewer, one for respondent), making the cost for both remote data collection modes more expensive than using other survey software, such as Open Data Kit, modified for phone surveys.

In conclusion, Hybrid IVR was more expensive per complete survey and had a lower response rate, lower reliability and more missing data than CATI. In addition, the characteristics of women who completed CATI compared to those who completed Hybrid IVR better align with the FTF sample. Therefore, for future remote data collection studies among women in Burkina Faso and similar settings, we recommend the use of CATI.

APPENDIX

Appendix Table 1. Characteristics of study participants: overall and by study arm

	% of study participants	% of Arm A: Hybrid IVR 1st, CATI 2nd	% of Arm B: CATI 1st, Hybrid IVR 2nd
Total	n=1766	n=882	n=884
Age			
Mean	28.5	28.5	28.6
15-19	19	20	19
20-24	20	20	19
25-29	19	19	20
30-34	15	14	16
35-39	12	11	12
40-44	9	10	8
45-49	6	6	6
Urban/rural			
Urban	67	67	66
Rural	33	33	34
Marital status			
Currently not in union	37	36	33
Currently in union	63	64	66
Highest school attended			
Never	42	41	43
Primary	20	20	20
Secondary or higher	38	39	37
HH Wealth (quintile)			
Lowest	8.3	8.3	8.9
Lower	9.1	9.9	8.8
Middle	12.2	12.4	12.3
Higher	16.8	15.3	18.1
Highest	53.8	54.2	51.8
Parity			
Yes	72	72	72
No	28	28	28
HH Electricity			
Yes	62	62	62
No	38	38	38

Appendix Table 2. Distribution of call outcomes by mode for both data collection rounds, using AAPOR standard definitions

Code		Hybrid IVR (n=1700)		CATI (n=1668)	
		N	%	N	%
	Eligible, non-interview				
NC (2.20)	Non-contact (didn't pick up any call)	464	27.3	452	27.1
O (2.36)	Non-contact (someone picked up call but woman not found)	310	18.2	244	14.6
R (2.12)	Refusal pre-consent	160	9.4	124	7.4
R (2.111)	Refusal at consent	19	1.2	12	0.7
R (2.121)	IVR Break-off (consented but unable to push 1 on phone)	46	2.7	N/A	N/A
R (2.12)	Break-off (consented but less than 50% of relevant questions)	324	19.0	24	1.4
	Interview				
P (1.2)	Partial (50-80% of relevant questions answered)	43	2.5	8	0.5
I (1.1)	Complete (more than 80% of relevant questions answered)	334	19.7	804	48.2

Appendix 3. Number of contacts to a complete interview, by round

The percent of women that completed the interview by contact was fairly comparable by mode. During the first round of calls (i.e., CATI in Arm A and Hybrid IVR in Arm B), 44% of women who completed were only called once before completing (Table 5B). 21% completed on the second contact, 16% on the third. The percent of women who completed decreased each call, with only 4% of women who completed answering the survey on the sixth contact.

The trend was the same during the second round of calls, but slightly more women who completed did so on the first (46%) and second (25%) contact (Table 5C). A slightly lower percent completed on the third called during the second round compared to the first (12% compared to 16%) but the trend for four or more calls to completion was the same as the first round.

Appendix Table 3a. Total number of contacts by call outcome, both modes of data collection, Round 1

	Complete	Partial	Break-off	Refusal @ consent	Refusal pre-consent	Non-contact Picked up	Non-contact didn't pick-up	Total
1	255	83	14	9	45	110	0	516
	43.6	36.7	45.2	45.0	31.3	35.8	0	29.2
2	122	45	8	2	24	66	0	267
	20.9	19.9	25.8	10.0	16.7	21.5	0	15.1
3	92	39	3	3	9	34	0	180
	15.7	17.3	9.7	15.0	6.3	11.1	0	10.2
4	54	18	4	0	18	13	0	107
	9.23	8.0	12.9	0	12.5	4.2	0	6.1
5	37	21	1	0	8	12	3	82
	6.3	9.3	3.2	0	5.6	3.9	0.7	4.6
6 +	25	20	1	6	40	70	450	606
	4.3	8.9	3.2	30.0	27.8	22.8	99.3	34.3
Total	585	226	31	20	144	307	453	1,766

Appendix Table 3b. Total number of days called by call outcome, both modes of data collection, Round 2

	Complete	Partial	Break-off	Refusal @ consent	Refusal pre-consent	Non-contact Picked up	Non-contact didn't pick-up	Total
1	245 46.0%	92 47.4	7 46.7	1 9.1	65 46.4	90 36.4	23 5.0	523 32.6
2	133 25.0%	44 22.7	2 13.3	4 36.4	23 16.4	44 17.8	0 0	250 15.6
3	65 12.2%	24 12.4	4 26.7	3 27.3	8 5.7	23 9.3	0 0	132 8.2
4	36 6.8%	14 7.2	2 13.3	2 18.2	8 5.7	23 9.3	0 0	85 5.3
5	31 5.8	10 5.2	0 0	1 9.1	7 5.0	16 6.5	1 0.2	66 4.1
6 +	22 4.1%	10 5.2	0 0	0 0	29 20.7	46 18.6	439 94.8	544 34.0
Total	532	194	15	11	140	247	463	1,602

Appendix 4. IVR Pilot – A/B Test Results, September 2017

In September 2017, Viamo, PMA2020 and ISSP randomly called 9,514 phone numbers in Burkina Faso. 2,293 picked up our phone call, and 107 women completed our survey. We embedded two A/B tests in this survey. Respondents in Group A were randomized to hear no encouragement messages. Respondents in Group B were randomized to hear two encouragement messages during the survey, one message after all demographic questions were asked that said "Thank you for providing us this information. You have already completed more than half of the survey! Please continue to the end." The second encouragement message was five questions later, after listening to the contraceptive knowledge questions and said "Let's go. It continues. Thank you for your participation!" Respondents in Group A (no encouragement) were more likely to complete the survey than women that received encouragement.

Appendix Table 4a. Survey completion by encouragement group

Encouragement Group	# of women that completed	# of women that consented	% of women who completed that consented the survey
Group A – No encouragement	54	60	90
Group B – Encouragement	53	68	78

During the same survey, we also randomized respondents (all persons who listened to the first question of the survey are included in this analysis) to hear language order in one of two ways: Order 1 offers languages from most widely to least widely spoken (Moore, French, Dioula, Gourmantchema, Fulfulde) whereas Order 2 offers languages from least to most spoke (Fulfulde, Gourmantchema, Dioula, Moore French). We found that the order of questions impacts the rate at which the respondent chooses a language, with more respondents choosing a language when it is offered as the first or second option, compared to if it is offered as the 4th or 5th. Dioula was consistently offered third and the same number of women chose Dioula in Order 1 and Order 2.

Appendix Table 4b. Language choice among all respondents that listened to the language question, by group

	Moore	French	Dioula	Gourmantchema	Fulfulde
Order 1 # of persons who chose language	840	234	132	32	38
Order 2	581	134	136	74	92

# of persons who chose language					
% of total 1	65.8%	18.3 %	10.3 %	2.5 %	3.0 %
% of total 2	57.1%	13.2 %	13.4 %	7.3 %	9.0%
Radio 1 : 2	1.45	1.75	0.97	0.43	0.41

Order 1: Moore, French, Dioula, Gourma, Fulfulde
Order 2: Fulfulde, Gourma, Dioula, Moore, French

Appendix 5a. Burkina Faso Remote Data Collection Follow-up Survey – CATI

NO	QUESTIONS AND FILTERS	CODING CATEGORIES	SKIP
	Section 0 – Identification		
0a	Please specify your interviewer number (1-15) *(not asked to respondent)*	Number ☐	
0b	Is this the first contact with the participant? *(not asked to respondent)*	Yes ....1 No 0	
1	What language would you like to speak?	Language other than 5 available…………………………….....1 same language as interviewer………..2 Language other than interviewer's langue but 1 of 5 survey languages….3	
2	Hello, my name is [Interviewer name]. I am calling on behalf of Institute Superieur de Science de la population. Is this [respondent name]?	Yes ....1 No 0	If yes, go to Q6 If no, go to Q3
3	Do you know who [respondent name] is?	Yes ....1 No (Do not call back)..................0	If yes, go to Q4 If no, go to Q7a
4	Is [respondent name] available to speak to me now?	Yes ....1 No 0	If yes, go to Q5 If no, go to Q7b
5	Hello, my name is [Interviewer name]. I am calling on behalf of Institute Superieur de Science de la population. Is this [respondent name]?	Yes ....1 No (Do not call back)..................0	Go to Q6
6a	We called last week and we are calling again to ask the same questions. We want to learn the best way to interview women by telephone. This is why we are calling you back.	(Message)	Second round calls only
6	Your responses to all questions in this survey are strictly confidential and anonymous. Can you please tell me how old you are?	Number ☐	If 15-49, go to Q7 If not 15-49, go to 7c
7a	Thank you for your time but this call was to [participant's name]. Goodbye	(Message)	
7b	Thank you for your time but this call was to [participant's name]. I will call back another time. Goodbye	(Message)	
7c	Thank you for your time but you are not in the age range of people who are part of this study. Goodbye	(Message)	
	Informed Consent		

7	You are eligible for this study. We are calling you to participate in a survey about women's health. Your participation is voluntary and responses to all questions are strictly confidential and anonymous. This survey takes 10-15 minutes and you can end the call at any time. The information you provide will be used to improve the health of women in Burkina Faso. I will send you 500 cfa of phone credit to compensate your time. Do you agree to participate?	Yes ....1 No 0	If yes, go to Q11 If no, go to Q8
8	What is the respondent's hesitation?		If "other" go to Q9 If 1-3, go to Q10
9	Please note participants' hesitation:	Other (TEXT ENTRY BELOW)....1 No confidence in the study ....2 General hesitation ....3 Too occupied 4	Go to Q10
10	What is the outcome of your consent discussion?	Consented (continue survey).......1 Not consented (end survey/do not call back) ....2 Call back.....................................3	Go to Q11
11	If we get interrupted during this call, would I be able to call you back to finish the survey?	Yes ....1 No 0	Go to Q12

Section 1 – Respondent's Background, Marital Status, HH characteristics			
12	Do you live in a city or village? For example, a city like Ouaga, Bobo, Koudougou, Fada or a rural village?	Urban .... 0 Rural... 1	
13	What is the highest level of school you attended?	Never Attended .... 0 Primary..................................... 1 Secondary (1st cycle)................ 2 Secondary (2nd cycle)............... 3 Tertiary .... 4 No response............................ -99	
14	Are you currently married or living together with a man as if married?	Yes……………………………1 No.. 0 No response............................ -99	

	Section 2 – Reproduction, Pregnancy & Fertility Preferences		
15	Have you ever given birth?	Yes .... 1 No... 0	
	Section 3 – Contraception		
16	Now I'd like to ask you some questions about family planning methods. I'll ask if you've heard them before. The methods are implants, injectables, the pill, condoms, IUD or IUD.	(Message)	
17	The first method: implants. Have you ever heard of the contraceptive implant? PROBE: Women can have one or several small rods placed in their upper arm by a doctor or nurse, which can prevent pregnancy for one or more years.	Yes .... 1 No... 0 No response .... -99	
18	The second method: injectables. Have you ever heard of injectables? PROBE: Women can have an injection by a health provider that stops them from becoming pregnant for one or more months.	Yes .... 1 No... 0 No response .... -99	
19	The third method: pills. Have you ever heard of the (birth control) pill? PROBE: Women can take a pill every day to avoid becoming pregnant.	Yes .... 1 No... 0 No response……….…..……..-99	
20	The fourth method: condoms. Have you ever heard of condoms? PROBE: Men can put a rubber sheath on their penis before sexual intercourse.	Yes .... 1 No... 0 No response……….…..……..-99	
21	The fifth method: IUD. Have you ever heard of the IUD? PROBE: Women can have a loop or coil placed inside them by a doctor or a nurse.	Yes .... 1 No... 0 No response……….…..……..-99	
22	Are you or your partner currently doing something or using any method to delay or avoid getting pregnant?	Yes .... 1 No... 0	If yes, go to Q24 If no, go to Q23a
23a	Are you pregnant?	Yes .... 1 No... 0 No Response………..……… -99	If yes, go to Q26a If no/NR, go to Q23b

23b	Would you like to become pregnant?	Yes .... 1 No... 0 No response……….....……..-99	Go to Q26a
24	Which method are you using?	Implant .... 3 IUD .... 4 Injectables .... 5 Pill .... 7 Male Condom .... 9 Other modern .... -55 No response .... -99	Go to Q25
25	Do you use a second method?	Yes .... 1 No... 0 No response……….....……..-99	Go to Q26
26	Which region of Burkina Faso do you live in?	Boucle du mouhoun……………….1 Cascades…………….……………. 2 Centre……………….……………. 3 Centre-est.……………….………. 4 Centre-nord.…………….………. 5 Centre-ouest.………………….…. 6 Centre-sud.…………….…………. 7 Est.……………….………………. 8 Hauts-bassins.……………….…. 9 Nord……………….……………. 10 Plateau Central ……………….. 11 No repsonse (active)……….…. -77	
27a	Thank you very much, the survey is now over! We will call you back within the next 2 weeks to re-interview you.	(Message)	Round 1 calls only
27b	Thank you very much, the survey is now over! You will receive your communication credit in the next few days. Have a nice day.	(Message)	Round 2 calls only

Appendix 5b. Burkina Faso Remote Data Collection Follow-up Survey – Hybrid IVR

NO	QUESTIONS AND FILTERS	CODING CATEGORIES	SKIP
	Section 0 – Identification		
0a	Please specify your interviewer number (1-15) *(not asked to respondent)*	Number [ ]	
0b	Is this the first contact with the participant? *(not asked to respondent)*	Yes ....1 No 0	
1	What language would you like to speak?	Language other than 5 available .1 Same language as interviewer ....2 Language other than interviewer's language but not 1 of 5 survey languages 3	
2	Hello, my name is [Interviewer name]. I am calling on behalf of Institute Superieur de Science de la population. Is this [respondent name]?	Yes ....1 No 0	If yes, go to Q6 If no, go to Q3
3	Do you know who [respondent name] is?	Yes ....1 No (Do not call back) 0	If yes, go to Q4 If no, go to Q7a
4	Is [respondent name] available to speak to me now?	Yes ....1 No 0	If yes, go to Q5 If no, go to Q7b
5	Hello, my name is [Interviewer name]. I am calling on behalf of Institute Superieur de Science de la population. Is this [respondent name]?	Yes ....1 No (Do not call back) 0	Go to Q6
6a	We called last week and we are calling again to ask the same questions. We want to learn the best way to interview women by telephone. This is why we are calling you back.	(Message)	Second round calls only
6	Your responses to all questions in this survey are strictly confidential and anonymous. Can you please tell me how old you are?	Number [ ]	If 15-49, go to Q7 If not 15-49, go to 7c
7a	Thank you for your time but this call was to [participant's name]. Goodbye	(Message)	
7b	Thank you for your time but this call was to [participant's name]. I will call back another time. Goodbye	(Message)	
7c	Thank you for your time but you are not in the age range of people who are part of this study. Goodbye	(Message)	
	Informed Consent		
7	You are eligible for this study.	Yes ....1 No 0	If yes, go to Q11

	We are calling you to participate in a survey about women's health. Your participation is voluntary and responses to all questions are strictly confidential and anonymous. This survey takes 10-15 minutes and you can end the call at any time. The information you provide will be used to improve the health of women in Burkina Faso. I will send you 500 cfa of phone credit to compensate your time. Do you agree to participate?		If no, go to Q8
8	What is the respondent's hesitation?		If "other" go to Q9 If 1-3, go to Q10
9	Please note participants' hesitation:	Other (TEXT ENTRY BELOW)....1 No confidence in the study ....2 General hesitation ....3 Too occupied 4	Go to Q10
10	What is the outcome of your consent discussion?	Consented (continue survey).......1 Not consented (end survey/do not call back) ....2 Call back.....................................3	If consented, go to Q11 If not consented, Q7a If call back, Q7b
11	Now that you have consented, I'm going to introduce you to the survey, which is automated. That means that in this survey, the procedure to answer questions is as follows: once you hear the question, you press the number corresponding to the answer you want to choose. For example, you will hear a recorded voice ask are you a man or a woman? If you are a man, press 1, if you are a woman press 2. [Ask] Which number would you press? [should be 2]. If you don't want to respond to a question, you can press 8 to skip that question. You can hang-up at any time to end the survey. I will be disconnected from the line once the survey begins. Do you have any questions for me? Ok, please press 1 to begin the survey. This is when I will leave the call and you will answer the automated survey questions. Have a good day!	Participant incapable of pressing (End survey) ....0 Participant began IVR survey 1	If pressed 1, go to Q12a If incapable, end survey

	Section 1 – Respondent's Background, Marital Status, HH characteristics			
12a	Hello, we can now start the study.	(Message)		
12b	If you don't want to answer, you can skip the question by pressing 8. To repeat the question, press 9	(Message)		
12c	Which country do you live in? For Ivory Coast, press 1 For Burkina Faso, press 2 For Niger, press 3 To repeat the question, press 9.	Ivory Coast ....1 Burkina..2 Niger ....3 Hung up .... -66 Listened but no response .... -77 No response (active - skip) .... -88 Don't know .... -99	If **not** 2, then play 12e If 2, play 12d Repeat question once if respondent does not get the question correct on the first try. If gets wrong both tries, just advance without playing 12d	
12d	Well done! You chose the right answer. You understand the questions very well. We can continue to the survey.	(Message)		Play only if gets 12c correct
12e	The response you chose is incorrect. Please try again.	(Message)		Play only if 12c incorrect
12	Do you live in a city or village? For example, a city like Ouaga, Bobo, Koudougou, Fada or a rural village? For a city, press 1 For a town, press 2 If you do not know, press 3	Urban ....0 Rural...1 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99		
	Have you ever been to school ? If yes, press 1 If no, press 2	Yes ....1 No...0 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99		
13	What is the highest level of school you attended? For primary, press 1 For secondary or higher, press 3 If you do not know, press 3	Primary...0 Secondary + ...1 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99		
14	Are you currently married? If yes, press 1 If no, press 2	Yes ....1 No...0 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99		

Section 2 – Reproduction, Pregnancy & Fertility Preferences

15	Have you ever given birth? If yes, press 1 If no, press 2	Yes ....1 No..................................0 Hung up.................................. -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	

Section 3 – Contraception

16	Now I'd like to ask you some questions about family planning methods. I'll ask if you've heard them before. The methods are implants, injectables, the pill, condoms, IUD or IUD.	(Message)	
17	The first method: implants. Women can have one or several small rods placed in their upper arm by a doctor or nurse, which can prevent pregnancy for one or more years. Have you ever heard of the contraceptive implant? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No..................................0 Hung up.................................. -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	
18	The second method: injectables. Women can have an injection by a health provider that stops them from becoming pregnant for one or more months. Have you ever heard of injectables? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No..................................0 Hung up.................................. -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	
19	The third method: pills. Women can take a pill every day to avoid becoming pregnant. Have you ever heard of the (birth control) pill? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No..................................0 Hung up.................................. -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	
20	The fourth method: condoms. Men can put a rubber sheath on their penis before sexual intercourse. Have you ever heard of condoms? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No..................................0 Hung up.................................. -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	

21	The fifth method: IUD. Women can have a loop or coil placed inside them by a doctor or a nurse. Have you ever heard of the IUD? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No...0 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	
22	Are you or your partner currently doing something or using any method to delay or avoid getting pregnant? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No...0 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	If yes, go to Q24 If no, go to Q23a
23a	Are you pregnant? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No...0 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	If yes, go to Q26a If no/NR, go to Q23b
23b	Would you like to become pregnant? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No...0 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	Go to Q26a
24	Which method are you using? For implants, press 1 For injectables, press 2 For pills, press 3 For condoms, press 4 For IUD, press 5 For another method, press 6 To repeat the question, press star	Implant ....3 IUD ....4 Injectables ....5 Pill .... 7 Male Condom ....9 Other modern .... -55 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	Go to Q25
25	Do you use a second method? If yes, press 1 If no, press 2 If you are not sure, press 3	Yes ....1 No...0 Hung up.................................... -66 Listened but no response .... -77 No response (active - skip) ...... -88 Don't know .... -99	Go to Q26a

26a	Use your keypad to type 60. Press two digits on your keyboard: 6 and 0.	Hung up.................................. -66 Listened but no response .... -77 No response (active - skip) ...... -88	
26b	Use your keypad to indicate your age. Press the two digits of your age.	Hung up.................................. -66 Listened but no response .... -77 No response (active - skip) ...... -88	
27a	Thank you very much, the survey is now over! We will call you back within the next 2 weeks to re-interview you.	(Message)	Round 1 calls only
27b	Thank you very much, the survey is now over! You will receive your communication credit in the next few days. Have a nice day.	(Message)	Round 2 calls only

REFERENCES

[1] Bongaarts J. Development: Slow down population growth. Nature. 2016;530:409-12.

[2] Family Planning 2020. 2016.

[3] Zimmerman L, Olson H, Group PMAPI, Tsui A, Radloff S. PMA2020: Rapid Turn-Around Survey Data to Monitor Family Planning Service and Practice in Ten Countries. Stud Fam Plann. 2017;48:293-303.

[4] Track20 - Monitoring progress in family planning

[5] United Nations DoEaSA, Population Division. World Contraceptive Use 2018. 2018.

[6] PMA2020 Methodology. Baltimore: Performance Monitoring and Accountability 2020 (PMA2020); 2016.

[7] Moura EC, Moreira CR, Bernal R, Ribeiro J, Malta DC, Neto OM. A feasibility study of cell phone and landline phone interviews for monitoring of risk and protection factors for chronic diseases in Brazil. Cad Saúde Pública. 2011;27:10.

[8] Greenleaf AR, Vogel L. Interactive Voice Response Technology for Data Collection in Sub-Saharan Africa. In: Viamo, editor. Brief Toronto, Canada: Viamo; 2018. p. 6.

[9] Corkrey R, Parkinson L. Interactive Voice Response Review of Studies 1989–2000. Behavior Research Methods, Instruments & Computers. 2002;34:342-53.

[10] Kempf AM, Remington PL. New Challenges for Telephone Survey Research in the Twenty-First Century. Annual Review of Public Health. 2007;28:14.

[11] Cannell C, Miller PV, Oksenberg L. Research on Interviewing Techniques. Sociological Methodology. 1981;12:39.

[12] Groves RM. Three Eras of Survey Research. Public Opinion Quarterly. 2011;75:861-71.

[13] Jackle A, Roberts C, Lynn P. Assessing the Effect of Data Collection Mode on Measurement. International Statistical Review. 2010;78:18.

[14] Toninelli D, Pinter R, de Pedraza P. Mobile Research Methods: Opportunities and challenges of mobile research methodologies. London: Ubiquity Press; 2015.

[15] Gibson DG, Pereira A, Farrenkopf BA, Labrique AB, Pariyo GW, Hyder AA. Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review. J Med Internet Res. 2017;19:e139.

[16] Greenleaf AR, Gibson DG, Khattar C, Labrique AB, Pariyo GW. Building the Evidence Base for Remote Data Collection in Low- and Middle-Income Countries: Comparing Reliability and Accuracy Across Survey Modalities. J Med Internet Res. 2017;19:e140.

[17] Hoogeveen J, Croke K, Dabalen A, Demombybes G, Giugale M. Collecting high frequency panel data in Africa using mobile phone interviews. Canadian Journal of Development Studies. 2014;35:186-207.

[18] Ballivian A, Azevedo J, Durbin W, et al. Listening to LAC: Using Mobile Phones for High Frequency Data Collection. Washington, DC: The World Bank; 2013.

[19] World Population Data Sheet In: Bureau PR, editor. Washington, D.C.2018. p. 19.

[20] 2016 World Population Data Sheet. Washington, DC: Population Reference Bureau; 2016.

[21] Desire Lohe K, Sinare T, Seroussi M. Enquête Démographique et de Santé Burkina Faso 1993. Calverton, Maryland: Institut National de la Statistique et de la Démographie and Macro International.; 1994.

[22] Human Development Index (HDI). In: Programme UND, editor.2015.

[23] Bank TW. Burkina Faso - Open Data. 2018.

[24] PMA2020 - Burkina Faso- Round 3. In: PMA2020, editor. Key Indicator Briefs - Family Planning Baltimore, MD: Johns Hopkins University; 2016. p. 2.

[25] Burkina Faso. In: AfroBarometer, editor.2016.

[26] PMA2020 - Burkina Faso- Round 5. In: PMA2020, editor. Key Indicator Briefs - Family Planning Baltimore, MD: Johns Hopkins Univeristy 2018. p. 2.

[27] About PMA2020. In: PMA2020, editor. Baltimore, MD2016.

[28] Jiang S. Use Item-Level Timestamp Paradata for Multi-Country Comparison. Comparative Survey Design and Implementation Workshop. Limerick, Ireland2018.

[29] Greenleaf AR, Ahmed S, Moreau C, Guiella G, Choi Y. Cell Phone Ownership and Modern Contraceptive use in Burkina Faso: Implications for Research and Interventions using Mobile Technology Contraception. 2018.

[30] Kohavi R, Thomke S. The Surprising Power of Online Experiments. Harvard Business Review. Boston, MA.2017.

[31] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012;22:7.

[32] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:16.

[33] Anglewicz P, Adams J, Obare F, Kohler HP, Watkins S. The Malawi Diffusion and Ideational Change Project 2004-06: Data collection, data quality, and analysis of attrition. Demogr Res. 2009;20:503.

[34] Mahfoud Z, Ghandour L, Ghandour B, Mokdad AH, Sibai AM. Cell Phone and Face-to-face Interview Responses in Population-based Surveys: How Do They Compare? Field Methods. 2014;27:39-54.