# Cell Phone Surveys among Women of Reproductive Age in Burkina Faso: Identifying Sources of Error

By

Abigail R Greenleaf, MPH

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

December, 2018

## Dissertation Committee

<u>Committee of Final Thesis Readers</u>

Saifuddin Ahmed, PhD
Professor and Co-Thesis Advisor
Department of Population, Family and Reproductive Health
Johns Hopkins Bloomberg School of Public Health

Dustin Gibson, PhD
Assistant Scientist
Department of International Health
Johns Hopkins Bloomberg School of Public Health

Caroline Moreau, MD
Associate Professor and Co-Thesis Advisor
Department of Population, Family and Reproductive Health
Johns Hopkins Bloomberg School of Public Health

Doug Storey, PhD
Associate Professor & Dissertation Defense Committee Chair
Department of Health, Behavior and Society
Johns Hopkins Bloomberg School of Public Health

<u>Alternate Committee Members</u>

Tamar Mendelson, PhD
Associate Professor
Department of Mental Health
Johns Hopkins Bloomberg School of Public Health

George Pariyo, PhD
Senior Scientist
Department of International Health
Johns Hopkins Bloomberg School of Public Health

Amy Tsui, PhD
Professor
Department of Population, Family and Reproductive Health
Johns Hopkins Bloomberg School of Public Health

## Abstract

**Background:** Demographers use reproductive health data to contextualize the state of a country's contraceptive behaviors, assess trends in reproductive health indicators and generate population projections. The rapid growth of cell phone ownership in low and middle income countries (LMIC) provides an opportunity to frequently collect data at low cost by calling respondents remotely. However, little is known about the validity of reproductive health estimates derived from cell phone surveys in LMIC – either collected through computer assisted telephone interviews (CATI) or interactive voice response (IVR). This dissertation identifies sources of frame and non-response error in a cell phone survey among women of reproductive age in Burkina Faso; a country that despite having one of the highest total fertility rates in the world is rapidly increasing contraceptive prevalence: from 15% in 2010 to 26% in 2018. [1]

This dissertation had three aims: 1) examine cell phone survey frame error by identifying differences in sociodemographic characteristics by women's cell phone ownership status and evaluate the implications of cell phone ownership on estimates of modern contraceptive use, 2) evaluate non-response error among cell phone owners by sociodemographic characteristics and mode of interview related to participating in a follow-up cell phone survey and evaluating the implications of non-response error on contraceptive use estimates and 3) compare estimates of modern contraceptive use from a face-to-face (FTF) survey and a CATI cell phone survey with random digit dial (RDD) sampling.

**Methods:** The study used four datasets collected between 2016 and 2018 in Burkina Faso to address the three study aims. In Aim 1, we used data from the Performance Monitoring and Accountability 2020 (PMA2020) Survey Round 4, a nationally representative population-based FTF survey of women of reproductive age conducted in 2016. For Aim 2, we conducted a follow-up cell phone survey of cell phone owners identified in Round 4. These women were randomized to receive either a CATI or IVR survey that was introduced with a human operator. Finally, in Aim 3 we used data from a national probability FTF survey (PMA2020 Round 5) and a CATI phone survey selected through RDD.

Aim 1 used bivariate and multivariable logistic regression to asses of the sociodemographic characteristics associated with cell phone ownership. Aim 2 used the same regressions but to assess characteristics associated with cell phone non-response. In Aim 3 we assessed the equivalence of modern contraceptive use estimates between the FTF and RDD surveys with four percent margin of error and further explored differences in contraceptive use by survey mode by conducting a multivariable logistic regression. Analyses in Aims 1 and 3 were adjusted for complex survey design and survey weighting.

**Results:** About 47% of women in Burkina Faso owned a cell phone based on PMA2020 Round 4 survey data, with greater ownership among women with secondary education or higher (Odds Ratio (OR) = 4.3, 95% confidence interval 2.9 – 6.2), women who resided in urban areas, women from wealthier households (highest wealth quintile), and women who were over the age of 19.

Phone owners reported higher modern contraceptive use than non-cell phone owners (29% versus 16%, p-value <0.001).

When examining non-response patterns in the second aim, we found the survey response rate was higher among women randomized to CATI (50%) versus Hybrid IVR (19%). Cell phone owners who consented to our follow-up cell phone survey were more likely to reside in an urban area, were more educated and were more likely to be over the age of 20 years old compared to women who did not consent to participate (were not contacted or refused participation). Women who completed the Hybrid IVR survey were more likely to have secondary education than women who completed the CATI survey. However, we found no difference in contraceptive use between CATI completers and non-completers (28% versus 30%, p-value = 0.692), nor between Hybrid IVR completers and non-completers (36% versus 35%, p-value = 0.708).

Finally, the results of Aim 3 showed that the estimate of modern contraceptive use generated by the CATI survey with RDD sampling was not equivalent but higher to the FTF PMA2020 Round 5 (R5) survey estimate (RDD: 40%, FTF: 26%), even after applying post-stratification weights (RDD weighted modern contraceptive prevalence estimate: 39%). The odds of modern contraceptive use remained significantly higher among RDD respondents versus FTF respondents, even after adjustments for additional covariates.

**Conclusions:** This dissertation identified sources of error for both follow-up and RDD cell phone surveys in Burkina Faso suggesting that over-estimation of modern contraceptive use in phone surveys stems from frame bias rather than non-response bias and potentially measurement bias. As cell phone ownership increases, frame bias may be reduced. We recommend tracking cell phone ownership in population based surveys to identify when phone survey frame no longer introduces error in reproductive health estimates.

# Table of Contents

# List of Tables and Figures

## List of Tables

## List of Figures

# Chapter 1. Introduction

## 1.1 *Overview*

Over the next 80 years the global population will increase by 50%, driven by population growth in low and middle income countries (LMIC). These population projections have prompted renewed interest in family planning (FP) programs [2] for accelerating voluntary contraceptive use, which is a proximate determinate of fertility. [3] By preventing unintended pregnancy, including high risk pregnancies that occur too early, too soon, too late or too many, contraceptive use reduces maternal mortality risks, improves child survival through birth spacing, and improves the nutritional status of both mothers and children. [4, 5] In addition, by enabling a shift in the age structure of a population toward a favorable ratio of working population, family planning can spur economic development through the demographic dividend. [6]

Currently 214 million women of reproductive age from LMIC who want to avoid a pregnancy are not using a modern contraceptive method, increasing their risk of an unwanted pregnancy. [7] The 2012 London Summit on Family Planning convened stakeholders to discuss the opportunity to revitalize family planning (FP) programs in LMIC to reduce high rates of unmet need and subsequently reduce unwanted fertility. To track progress as part of the Family Planning 2020 (FP2020) initiative, [8] there is an interest in developing rapid, cost-effective approaches for data collection. The Demographic and Health Surveys (DHS) have served as the main source of reproductive health data in LMIC for the past 30 years. However, the five year periodicity and high cost of these face-to-face (FTF) surveys prevent tracking the anticipated rapid changes in FP behaviors related to FP2020 programmatic efforts.

The proliferation of cell phone ownership in Sub-Saharan Africa (SSA) provides a platform for low-cost, rapid turnaround data collection. [9] Urbanization, increased network coverage and decreased cell phone cost have contributed to exponential increases in the number of cell phone owners throughout the continent, providing an opportunity to reach a large population remotely by phone for data collection.

A transition from FTF to cell phone surveys, however, raises concerns of survey quality, including representativeness and measurement errors. Cell phone survey administration, whether through Computer Assisted Telephone Interview (CATI) or Interactive Voice Response (IVR), impacts all five components of survey errors: frame, non-response, specification, measurement, and data processing errors. Three errors are of particular concern for cell phone surveys in LMIC: frame, non-response and measurement. Cell phone ownership, which constitutes the sample frame of cell phone surveys, is not ubiquitous in LMIC but rather is skewed towards male, young, and educated populations, introducing frame error. Non-response error occurs when people who are sampled (cell phone owners) but not interviewed differ in a non-negligible way from those who are successfully interviewed. [10] These two errors introduce selectivity biases in cellphone surveys. Finally, measurement error occurs when incorrect information is recorded.

There is a dearth of research on cell phone surveys in LMIC. [11] A systematic review identified only four articles from LMIC that compared a FTF survey with either IVR or CATI phone surveys for health research [12] and none focused on FP measures.

The goal of this dissertation is to assess two sources of survey error, frame and non-response error, when using cell phone surveys to estimate modern contraceptive use. We chose to explore this question in Burkina Faso, a country experiencing rapid expansion of cell phone ownership, offering new opportunities to monitor its ambitious program to increase family planning uptake nationwide. Building on the Performance Monitoring and Accountability 2020 (PMA2020) platform that tracks annual modern contraceptive prevalence rates using FTF surveys, we designed and analyzed a study tracking the same indicators via cell phone survey, to assess frame and response errors (compared to a reference FTF survey) and their implications for estimates of modern contraceptive use rates. We further assess non-response error by cell phone survey mode (Hybrid IVR and CATI). We also conducted concurrent FTF and cell phone CATI surveys, the latter employing random digit dial (RDD) sampling, to test the equivalence of modern contraceptive estimates by survey mode. The study results provide information on the validity of sexual and reproductive health (SRH) estimates collected via cell phone surveys informing the opportunity to improve cost-effectiveness of family planning monitoring by substituting FTF surveys with phone surveys in a country like Burkina Faso.

## 1.2 *Specific Aims*

The specific aims of this dissertation were organized to address sources of total survey error. [13] The first aim assessed frame error and bias, by comparing women's sociodemographic characteristics by cell phone ownership and assessing the difference in modern contraceptive use estimates between cell phone owners and non-owners. The second aim looked at non-response error among women who owned cell phones by comparing women who completed a follow-up cell phone survey to those who did not. This was examined for two cell phone survey modes: CATI and Hybrid IVR. We also compared modern contraceptive prevalence rates between survey completers and non-completers to identify non-response bias. Finally we tested for the equivalence in estimates of modern contraceptive use between a CATI RDD cell phone survey and a nationally representative FTF survey (PMA2020).

The key aims of the study are follows:

Aim 1: Identify sociodemographic factors related to cell phone ownership among a nationally representative sample of women who participated in the 2016 PMA2020 survey in Burkina Faso and assess the relationship between cell phone ownership and modern contraceptive use.

*Hypothesis 1a: Women who own cell phones are more likely to be younger, educated, wealthier and live in an urban area compared to women who do not report cell phone ownership.*

*Hypothesis 1b: There is a statistically significant difference in modern contraceptive use between women who own cell phones and those who do not.*

Aim 2: Among cell phone owners identified in a nationally representative sample of women surveyed in a 2016 PMA2020 FTF survey in Burkina Faso, evaluate the sociodemographic factors associated with non-response to a follow-up cell phone survey and assess the implications of non-response on estimates of modern contraceptive use.

*Hypothesis 2a: Among women who own a cell phone, survey completion in a follow-up cell phone survey is higher for urban and young women.*

*Hypothesis 2b: Among women who own a cell phone, prevalence of modern contraceptive use is different among women who complete a follow-up cell phone survey and those who do not, and* *this difference is greater among women randomized to a Hybrid IVR follow-up survey than those randomized to receive a CATI follow-up survey.*

Aim 3: Compare the national estimate of modern contraceptive use in Burkina Faso between an RDD cell phone survey and a nationally representative FTF survey.

*Hypothesis 3a: The modern contraceptive prevalence rate estimated from an RDD phone survey is equivalent to the FTF sample estimate within +/- 4% margin-of-error after applying post-stratification weights.*

*Hypothesis 3b: The odds of contraceptive use are different in the RDD sample than in the FTF sample, after adjusting for women's sociodemographic characteristics and applying post-stratification weights.*

# Chapter 2. Background

## 2.1 *Overview*

Over the next 80 years, the global population will increase by 50% - from 7.2 to 11.2 billion. Population growth is projected to be greatest in Sub-Saharan Africa (SSA), where the United Nations predicts a quadrupling of the population by 2100, increasing from 1 to 3.9 billion. [2] Voluntary family planning programs are a key component of slowing population growth by reducing global fertility rates, [3, 14-16] contributing to a decline from 4.7 births per woman in the 1970s to 2.6 in the 2000s. [4] Globally, use of modern contraceptives avoids 308 million unplanned pregnancies annually. [7]

To monitor these trends, great effort has gone into producing comparable reproductive health indicators in low and middle-income countries (LMIC) [17, 18], starting with the World Fertility Surveys in the 1970s, followed by Demographic and Health Surveys in the 1980s. [19] These reproductive health indicators such as contraceptive prevalence rate, method mix, and unmet need for family planning contextualize the state of a country's contraceptive behaviors and family planning programs. [20, 21] Demographers use these estimates to compare progress within and between countries, forecast demographic transitions and generate population projections. They have successfully used these indicators to describe the demographic transition, characterized by a decrease in infant mortality which prompts a desire among the population to reduce the number of births, triggering an increased need for family planning. [22]

Reproductive health data are also essential for practical purposes – specifically for evidence-based public health planning. It is important for governments to track population changes so that infrastructure such as schools, hospitals and roads are properly planned, health commodities accurately forecasted, and economic growth appropriately calculated. [23]

This dissertation uses data from Burkina Faso, a West African country. West Africa has the highest fertility rate in the world (5.5 children per woman) [24] and by 2050 the majority of West African countries will triple their population. [25] West Africa's total fertility rate (TFR) is double Asia's and three times Europe's TFR. Modern contraceptive use is low in West Africa with only nine percent of married women using modern contraception versus 25 percent in East Africa. [24] Yet almost 30 percent of married women of reproductive age report the desire to delay or avoid a pregnancy but are not using contraception. [24] Investing in family planning (FP) would accelerate development of the region. Specifically, meeting unmet need in West Africa would avoid 500,000 child deaths, and the cumulative cost savings for maternal and child health care would be $182 million over the next 10 years. [24]

In recent years, two major conferences renewed interest in FP. The Ouagadougou Partnership, held in 2011, aimed to accelerate FP use in eight West African countries. [24] A year later, the 2012 London Summit on FP launched the Family Planning 2020 (FP2020) initiative, pledging new financial investments in family planning in LMIC. Specifically, stakeholders including priority-country governments, civil society representatives and multilateral institutions, private sector partners, foundations and donor governments agreed to provide 120 million additional women access to FP by 2020 in 69 priority countries, which have a gross national income (GNI) per capita of less than US $2,500. [8] To reach this ambitious target, stakeholders expressed a need for timely family planning data. The Demographic and Health Surveys (DHS), which have been the main source of reproductive health data in LMIC for the past 30 years, and other similar surveys, such as UNICEF's Multiple Indicator Cluster Surveys, are conducted at five-year intervals making them unsuitable for tracking rapid change. Demand from governments, multilateral institutions and civil society justifies improved FP data collection approaches that provide data in shorter intervals. Specifically, rapid turn-around FP data is needed for tracking the United Nation's Sustainable Development Goals (SDG) 3 (Good Health & Wellbeing) and 5 (Gender Equity), which are recommended to be measured annually. [26]

## 2.2 ***Current innovations in data collection platforms in SSA***

Since 2013, Performance Monitoring and Accountability 2020 (PMA2020) has collected face-to-face (FTF) surveys every six months to one year in 11 priority countries to provide national or regional estimates of family planning indicators. [27] To improve data quality and speed-up data collection and processing, PMA2020 interviewers record data on smart phones and upload data to a cloud server in real-time. [28] In most countries, PMA2020 hires local female interviewers rather than a mobile team of interviewers to conduct the survey. [29]

Although the PMA2020 platform has improved both the efficiency and quality of data collection, the approach is still costly and prohibits PMA2020's expansion to the remaining 58 FP2020 priority countries. The next application of technology for survey data collection in LMIC is contacting respondents on their personal cell phones, although there is minimal research on the feasibility and validity of this approach in such settings. [30]. Cell phone surveys may offer the opportunity to collect data faster and at a lower cost than FTF surveys. [31-33] Survey researchers in SSA  are becoming increasingly interested in this approach due to the rapid expansion of cell phone ownership over the past eight years. [34] Cell phone ownership is growing due to urbanization, increased network coverage and decreased cell phone cost. [35-37] Currently, there are over 690 million cell phone subscriptions for just over a billion people in SSA. [25] Over 1 billion subscriptions are anticipated by 2021. [36] By 2021, mobile phone penetration, which is the percent of unique users within a specific population, [38] is expected to be at least 50 percent in SSA. [39] Greater ownership presents the opportunity to communicate with respondents via their cell phones for myriad purposes: [40] to conduct research [30, 41] to collect surveillance data, [42] to improve medical adherence [43] and to send appointment reminders or conduct behavioral change interventions. [44]

There are three common approaches to collecting data via phones: Computer Assisted Telephone Interview (CATI), Interactive Voice Response (IVR), or Short Message Service (SMS). In a CATI approach, live interviewers administer a questionnaire verbally. CATI interviewers work in a call center using the main languages of a country. In contrast, IVR is conducted without an interviewer and requires that the respondents use their keypad to answer a pre-recorded question or prompt (e.g. “If yes, press 1. If no, press 2”). A modified version of this approach is a Hybrid IVR survey. In a Hybrid IVR survey, a live interviewer opens the call, confirms respondent eligibility, administers consent and explains how to respond to an IVR survey before transferring the respondent to the IVR survey. Finally, a message-based SMS approach sends and receives information via text messages. [45]

For the purposes of this dissertation, we only examined Hybrid IVR and CATI; SMS was not considered as it is not suitable for low literacy populations, [45] such as is the case in Burkina Faso. These two approaches, CATI and IVR, have strengths and weaknesses that are summarized in Table 2-1. A major advantage of CATI is that it allows trained interviewers to interact with the respondents, which increases response rates [32] and improves the respondent's understanding of questions. The major disadvantage is cost and potential social desirability bias when responding to sensitive questions. On the other hand, IVR uses pre-recorded questions, which can reduce interviewer bias and is less costly than a live interviewer. However, in the absence of direct communication, response rates and question comprehension are a source of concern. In addition, IVR requires a minimum level of literacy to press buttons on a keypad. In LMIC, CATI response rates for population-based surveys range from 30-98%, whereas IVR survey response rates are usually below 30%. These rates drop to single digit levels when respondents are recruited via random digit dialing (RDD). [46]

*Table 2-1. Strengths & weaknesses of IVR and CATI data collection*

| | **Strengths** | **Weaknesses** |
|---|---|---|
| IVR | • Recorded questions reduce interviewer bias<br>• Inexpensive<br>• Software automates data collection so human oversight of data collection is not needed | • Respondent may not be familiar with pre-recorded calls which decreases response rate<br>• Keypad-based response menu can be difficult to navigate or impossible to use if low literacy<br>• Difficult to judge data quality (i.e. respondents understanding, or if respondents was distracted) |

| | | |
|---|---|---|
| CATI | • Live operator, familiar interaction<br>• Good for low literacy settings<br>• Operators can clarify questions<br>• Interviewer builds relationships with respondents thus reducing refusals and attrition | • Resource intensive (interviewers, supervisors)<br>• Interviewer bias<br>• Higher cost<br>• Takes longer to complete surveys |

## 2.3 ***Changing survey platforms: theoretical issues moving from FTF to Cell Phone Surveys***

A transition from FTF to cell phone data collection raises issues of bias, data quality and comparability of indicators for trends analysis. All stages of survey design influence survey quality. While definitions vary, survey quality generally means "data fit for use". [47] In order to operationalize survey quality, survey researchers devised the concept of Total Survey Error (TSE). [10] TSE is often divided into measurement and representation errors. Representation errors are frame and non-response errors. Measurement errors are specification, measurement and data processing errors. [47] A more detailed description of the different sources of errors and how they apply to current FTF surveys versus telephone surveys is summarized in Table 2-2.

*Table 2-2. Types of Total Survey Error and applications to FTF & Cell Phone surveys in LMIC*

| Type of Total Survey Error* | Brief Explanation | Application to FTF Surveys+ | Application to Cell Phone Surveys |
|---|---|---|---|
| Frame error | A sampling frame lists all elements of the population from which samples are drawn. Errors in constructing, maintaining and using sample frame for selecting the survey sample lead to frame errors. Three types of frame error are population specification, selection and coverage. | Sampling frame (census) may have poor quality if the census was conducted several years earlier or coverage rate was poor. | In countries where phone ownership is universal, a phone survey is able to reach the target population. However, under-coverage in phone surveys in LMIC could be substantial due to inadequate levels of cell phone ownership or network coverage. For these reasons, members of the population who do not have a phone have a zero probability of being selected in phone surveys. Moreover, the sampling frame may not be publicly available. |
| Non-response error | People who are sampled but not interviewed differ in a non-negligible way from those who are successfully interviewed. | The causes of non-response in FTF may be linked to 1) failure to contact sampled respondents; 2) refusal to participate in survey; (e.g., due to confidentiality concerns); 3) refusal to answer questions (item non-response) | The causes of non-response in telephone surveys are different than FTF surveys. Non-response in telephone surveys stems from 1) failure to contact sampled respondents; 2) refusal to participate 3) health or language problems. |
| Specification error | Concept implied by survey question does not reflect concept it aims to measure. | Interviewer can read respondent verbal and non-verbal cues to gauge what is understood or not in FTF surveys. | Could be an egregious problem if subject matter experts (particularly cultural) are not engaged to pre-test questions in multiple languages. During CATI the interviewer has less information than FTF. IVR provides no opportunity to check respondent's understanding or comprehension. |
| Measurement error | Incorrect information is recorded for a reason other than mis-conceptualization of question (i.e. specification error). Source of error can be respondents, interviewers or survey questions. | With less supervision in the field, there is more opportunity to stray from protocol. However, FTF interaction allows interviewer to better gauge respondent comprehension than via phone interview which is particularly important in a country with a plethora of languages. | Processing survey questions and answers differ between FTF and cell phone surveys. Furthermore, it is harder to control the respondent's environment during the interview. It is difficult to ascertain whether a respondent is isolated from potential influencers (husband, family). |
| Processing error | Errors in editing, data entry, coding and assignment of survey weights. | Interviewer behavior, which is mostly unsupervised in the field can impact processing error. | Since CATI and IVR are relative new technology in SSA, there is possibility of data error entry in the data collection software. |

* Definitions adapted from: Biemer P. Overview of Design Issues: Total Survey Error. In: Marsden PV WJ, editor. Handbook of Survey Research. 2 ed. Bingley, United Kingdom: Emerald Group; 2010. p. 886.
+assuming clustered random probability surveys of households in Burkina Faso

Frame error is a type of representation error and arises in the process of "constructing, maintaining and using the sample frame(s) for selecting the survey sample." [10] More specifically, frame error occurs when there is a difference between the target population and the frame population. Under-coverage frame error occurs when members of the target population are not included in the frame. The frame for a FTF survey is usually provided by the national statistical agency, based on the most recent census. This is in contrast to cell phone surveys in LMIC that have an unknown (phone number directories are difficult to obtain or do not exist) or limited frame (i.e. individuals who do not own a phone are excluded). Under-coverage error is likely to occur since cell phone ownership is not ubiquitous in LMIC. In SSA, cell phone ownership is skewed towards educated, younger and majority-language-speaking men (usually French or English). [35] Network coverage is also unequally distributed, with greater cell phone coverage in urban compared to rural areas. When conducting population-based surveys, identifying frame error or bias is of prime importance as to avoid an incorrect (biased) estimate. [48, 49] Frame bias occurs when the outcome of interest is differential for those included in the frame compared to those not included in the frame.

The next type of representation error is non-response error. Non-response occurs when people who are sampled but not interviewed differ in a non-negligible way from those who are successfully interviewed. [10] Respondents may not answer calls from unknown phone numbers, or even if they do, they may be unwilling to participate in a telephone survey. Traditionally, response rates are used to evaluate non-response error and thus gauge survey quality, but high non-response does not necessarily result in bias if non-response is randomly distributed in the target population (random non-response). [50] However, if the propensity to respond is correlated with the outcome of interest, this will result in non-response bias. [10, 51] Response rates are typically higher in FTF than phone surveys due to social expectations that apply to FTF surveys that make it harder to refuse an interview when directly interacting with an interviewer rather than interacting remotely (CATI) or not at all (IVR). [52] Furthermore, during a FTF survey the interviewer connects both verbally and non-verbally with the respondent, in contrast with a CATI survey, where there is only a verbal communication channel. [53]

We now consider the first of three types of measurement error. Specification error occurs when a survey instrument fails to measure the theoretical constructs guiding the research. [54] Specification error, which is a lack of construct validity, can be avoided by pre-testing surveys to ensure that questions are measuring the intended construct. If the question fails to represent the underlying construct of interest, responses will produce specification error. [55]

Measurement error occurs when questions are properly specified, but incorrect information is recorded. Measurement error can caused by respondents or interviewers. Broadly, mode of data collection affects cognitive processes related to information retrieval, [56] judgements about the appropriate response, and answer choice. [57] Specifically, the effect of mode on survey responses is organized into three groups: 1) media related factors 2) factors influencing information transmission (cognitive processes) and 3) interviewer impact. [58] Media related factors depend on how knowledgeable and familiar respondents are with the survey medium (e.g. cell phone, computer). This factor also considers the impact of the media on control of the interview (interviewer or respondent), and the social conventions about silences. The second factor, information transmission, assesses channel capacity, which is the ability to use verbal, nonverbal and paralinguistic communication. This includes how information is presented, whether visually or auditory or both, and regulation of communication flow between interviewer and respondent. The final factor explores the impact of the interviewer on the question-answer process. For example, an IVR survey has limited interviewer effects on the data compared to an FTF survey or CATI. [59] In addition, there is less opportunity to build rapport on the phone compared to FTF, but greater opportunity to preserve confidentiality and potentially elicit more honest responses to sensitive questions. Biased reporting due to social desirability is a common challenge in FTF household surveys, which can partially be addressed by cell phone surveys, especially with IVR.

Generally, self-administered surveys or remote data collection yield higher report of socially undesirable behaviors than interviewer-administered surveys. [60, 61] However, the impact of mode of data collection on the reporting of sensitive behaviors in LMIC is discrepant. [23] A meta-analysis of 15 studies (mostly comparing FTF and audio computer-assisted self-interview (ACASI)) found that non-FTF methods did not consistently produce a significant increase in the report of four sensitive HIV risk behaviors. [62] Family planning research in Zimbabwe comparing ACASI to FTF found that level of education improved reliability of ACASI responses and that more highly educated women preferred ACASI to FTF. [62] Measurement error also varies by mode of cell phone survey. Studies in the United States comparing CATI & IVR sought to examine differences in report of sensitive behaviors such as substance abuse, [63-65] youth risk behaviors, [66] and child maltreatment. [67] In two of the studies, IVR elicited higher report of sensitive behaviors for alcohol and non-heteronormative sexual behaviors, but the third study had comparable estimates from the IVR and CATI modes. [63-65] Other factors contributing to measurement error in cell phone surveys include the quality of phone line, amount of multi-tasking or distractions and the extent to which by-standers are privy to the interview. [68]

Finally, processing errors occur at the end of the survey's life cycle. Processing errors can take place during data entry, editing, coding and assignment of survey weights.

## 2.4 *Research on sources of Total Survey Error in cell phone surveys*

While studies comparing cell phone and FTF surveys are rare in LMIC and even more scarce in SSA, there is a robust literature in Western countries investigating frame and non-response errors in cell phone surveys. In the next section, we will first review existing phone survey research in LMIC, then broaden the perspective to other regions of the world, as some findings are generalizable to other country settings.

A 2017 systematic review of the literature was conducted to identify health research from LMIC that compared at least one mode of remote data collection to another mode. [12] Ten articles were identified, covering seven countries, across four regions (Asia, Latin America, Europe and the Middle East); but none were conducted in SSA. One article, published by the World Bank, [32] directly compared IVR, CATI and FTF surveys whereas all other articles only compared one cell phone data collection mode to a FTF survey. [32] Four of the ten articles comparing CATI to FTF or IVR to FTF are pertinent to the present research. While none focused on FP, these comparisons inform both the feasibility of conducting cell phone surveys in LMIC as well as the quality of data collected remotely. These articles made four comparisons of FTF to CATI and one comparison of IVR to FTF. Of these five comparisons, two compared responses in independent samples [33, 69] and three compared responses among the same population interviewed by the two modalities, but none collected sexual and reproductive health (SRH) indicators. [32, 41] All comparisons generally showed concordance of results between modes. The most comprehensive project was a two-country study (Honduras and Peru) conducted by the World Bank. This study found that the sociodemographic characteristics of CATI respondents were very similar to the characteristics of FTF respondents. Compared to the FTF survey, discordant survey responses from the panel in Honduras ranged from -2.1% to 0% for CATI. Furthermore, CATI had the lowest discordance of responses with the FTF survey compared to SMS and IVR. [32] Two other studies that compared FTF and CATI, both in Brazil, concluded that the telephone survey was a rapid alternative to FTF surveys to provide non-communicable disease prevalence estimates. [33, 69] However, one article found that groups with higher telephone ownership or coverage were more likely to report better health. [33] A study in Lebanon concerning non-communicable disease included sensitive questions, whereas none of the other articles did. Report of past-year alcohol consumption, a sensitive topic in Lebanon, was slightly higher via CATI compared to FTF [41] but mode-specific estimates were comparable when data were stratified by age, gender, and education. [41]

Three articles, which were not included in the systematic review because they were stand-alone cell phone surveys, employed RDD sampling and attempted to assess frame and non-response errors. These studies compared survey respondents to a recent census or to the DHS surveys to quantify the differences between respondents and the general population. The first of the three studies was conducted by the Center for Global Development in four countries: Afghanistan, Ethiopia, Mozambique and Zimbabwe in 2014. [70] The project aimed to answer three research questions: 1) can phone survey platforms reach a nationally representative sample of a country's population via RDD sampling; 2) to what extent does linguistic fractionalization affect the ability to produce a representative sample; and 3) how effectively does monetary compensation impact survey completion. [70] The authors found that the countries with higher mobile penetration (Ethiopia and Zimbabwe) resulted in lower sample distortion and more comparable estimates to recent FTF surveys compared to the countries with lower phone penetration (Afghanistan and Mozambique). In Zimbabwe and Ethiopia, the majority languages (Shona and Amharic) were chosen by respondents more often than would be expected based on population make-up. Survey drop-out rates were higher for people who chose to take the survey in a minority language. Finally, respondents that were randomized to receive compensation were more likely to complete the survey compared to respondents who were offered no compensation. The impact of an incentive on response rate varied by country.

The second relevant study was conducted in the Ivory Coast in 2013. [71] The study, operated by a call center in France, used RDD sampling to assess HIV risk behaviors in the general population and found that RDD respondents were more likely reside in urban areas and be male than the DHS reference population. However, when sample distortion was accounted for by conducting stratified analysis by sex, age group, level of education and place of residence, HIV-related behaviors were similar in the RDD survey compared to DHS estimates. The authors concluded that CATI was feasible to assess self-reported HIV prevalence in Ivory Coast. The third and most recent survey was conducted in Ghana in 2017 and examined the combination of frame, non-response and measurement errors related to IVR RDD sampling. [72] The study found that fewer women, rural and older residents completed the survey compared to recent household surveys (DHS and census). Two-thirds of RDD respondents were male and more than half were 15-24 years of age whereas according to the 2017 census, 48% of the population was male and 30% were aged 15-24.

Broadening the research perspective to developed country settings, we see a number of studies have examined the combination of frame and non-response error in phone surveys. The seminal studies in the US exploring landline phone non-response took place in the 1980s and 1990s, using a variety of administrative data to indirectly construct the profile of non-respondents. [51] These studies consistently reported differential participation by education, age, minority status and civic engagement. [73] The increase in both cell phone ownership and internet connection combined with increasing distrust of phone surveys have caused RDD response rates in landline based phone surveys to plummet to single digits in the United States, estimated around nine percent by Pew Research Center. [74, 75] There is a dearth of research regarding cell phone non-response in the US, as noted by both 2010 [73] and 2017 [74] American Association for Public Opinion reports on cell phone surveys in the US. Non-response has also been extensively studied in Europe. A recent health survey of 15,635 people using a mobile and landline frame in France yielded much higher response rates than in the US: the landline survey response rate was 47.3%

and cell phone survey response rate was 37.7%. [76] In Europe, individuals who only own a cell phone are younger, better educated, earn higher incomes and have greater technology competencies than individuals who have landline phones. [77]

In conclusion, research comparing data quality between phone and FTF surveys is scarce in LMIC, limiting the ability to inform survey implementation in LMIC. The five sources of total survey error are not explicitly assessed in current research literature and no studies investigate these sources of errors among female populations. Existent research from the US, Europe and a few LMICs provide some insights which need further investigation to inform phone survey development for sexual and reproductive health research and monitoring in SSA.

# Chapter 3. Research Design, Data & Methods

## 3.1 *Proposed research*

This dissertation used the total survey error (TSE) framework to assess frame error and response error when using cell phone surveys compared to a reference FTF nationally representative population-based survey and its implications for modern contraceptive estimates in Burkina Faso. We present an overview of the country, the state of women's reproductive health in Burkina Faso and information about relevant infrastructure for cell phone surveys. We then describe the research design and methods for addressing the dissertation objectives.

## 3.2 *Overview of study setting: Burkina Faso*

The study was implemented in Burkina Faso, a West African country with a population of approximately of 20.3 million, which is projected to grow to 28.9 million by 2030 and to 48.5 million by 2050. [25] Burkina Faso is sixth to last on the 2015 Human Development Index. [78] Although the country is rapidly urbanizing, only 30% of the population is urban. [79] Of the urban population, 46% lives in Ouagadougou. [80]

*Figure 3-1. Burkina Faso highlighted in map of Africa*

## 3.3 *Modern contraceptive use in Burkina Faso*

The Total Fertility Rate (TFR) – the average number of births expected per woman during her reproductive life – in Burkina Faso is 5.7 - one of the highest TFRs in the world. [79] Although recent contraceptive uptake in Burkina Faso is encouraging – with an increase in modern contraceptive prevalence among women in union from 15% in 2010 to 26% in 2018 [1] – 20% of women in union still have an unmet need for family planning. [81] Burkina Faso launched its National Plan to revive family planning programs in 2013 which was followed by the current plan which aims to accelerate progress; setting its national target to 32% contraceptive prevalence among married women by 2020. [82] Notable investments will be needed to accelerate the progress to reach the national goal, which would require an increase in contraceptive use by two percent annually.

Burkina Faso first collected Demographic and Health Survey (DHS) data from 1992-93. At the time, although 62 percent of women of reproductive age had heard of a modern contraceptive method, only four percent reported modern method use. [83] In the most recent round of PMA2020 (R5), among women in a union, 30% reported use of a modern contraceptive method, and among all women 26% used a modern contraceptive method. [84] Including traditional methods minimally increased the percent of women using a method: 32% for women in a union and 28% for all women. [84] Five methods make up 98.8% of current modern contraceptive use: implants, IUDs, injectables, pills and male condoms. [85] See Table 3.1 for details.

*Table 3-1. Distribution of modern methods among women of reproductive age using a modern method, Burkina Faso Round 5 PMA2020*

| **Modern Method (n = 1000)** | **% using method** |
|---|---|
| Implant | 42.7% |
| Injectable | 23.8% |
| Pill | 13.5% |
| Male condom | 14.1% |
| IUD | 4.7% |
| Emergency | 0.3% |
| Female sterilization | 0.6% |
| Other modern | 0.3% |

## 3.4 ***Infrastructure related to cell phone surveys in Burkina Faso***

Infrastructure plays a key role in cell phone surveys due to the reliance on a cell phone network connection for data collection. Over 80 percent of households in Burkina Faso own a phone, [85, 86] which is comparable to the percent of households that had landlines when the United States made the transition from FTF to telephone surveys in the 1980s. [74] Network coverage and electricity greatly influence non-response error. There are currently three mobile network operators in Burkina Faso: Telemob, Telecel, and Orange. However, extent of network coverage (i.e. geographic coverage of cell phone service) in Burkina Faso is unknown. PMA2020 Round 4 (R4) data reported that 68% of urban households have electricity compared to 25% of rural households. Important to note, however, is that cell phones can be charged through means other than household electricity, including shops that charge cell phones for a small fee, or at a family member or neighbor’s house that has state electricity or a generator. Although populations living without electricity are less likely to have a charged cell phone due to reduced convenience, it cannot be assumed that those living in an area without electricity do not have a charged cell phone.

Literacy levels are also an important consideration since numeric literacy is required to participate in an IVR survey. It is challenging or impossible for those with low literacy to identify the number that corresponds with their answer during an IVR survey. According to the 2014 Malaria Indicator Survey, approximately 30% of women in Burkina Faso are literate. The youngest generation surveyed, 15-19 year olds, have the highest literacy rate (53%) and the oldest generation (45-49) has the lowest (10%). [87] Furthermore, literacy is much higher in urban (57%) compared to rural areas (18%) and differs by household economic status, with the highest quintile at 63% literacy compared to the lowest quintile's 12%. [87]

## 3.5 ***Study Design***

This was a sequential study that addressed three aims with relevant datasets: Aim 1 was addressed by a secondary analysis of PMA2020 R4 exploring characteristics of female cell phone owners in Burkina Faso. In Aim 2, we estimated response rates and evaluated non-response error and bias using a follow-up sample of cell phone owners who participated in the 4$^{th}$ round of the PMA2020 survey, conducted about 11 months earlier. Finally, Aim 3 used two national cross-sectional studies, one using a nationally representative FTF cross-sectional survey (PMA2020 R5), the other a CATI cellphone survey conducted using RDD. We first present an overview of PMA2020, the parent study for the project, and then describe the methodology used for each research aim.

## 3.6 ***Overview of PMA2020 platform***

This dissertation used the PMA2020 platform, launched in 2013 to monitor reproductive and water and sanitation indicators in support of the FP2020 goals. Funded by the Gates Foundation, PMA2020 is managed by the Bill and Melinda Gates Institute for Population and Reproductive Health at Johns Hopkins University in collaboration with country partner institutions. The team based at Johns Hopkins provides technical leadership and support to implementing partners in the host country. The first round of data collection took place in Ghana, and over the past five years, PMA2020 has conducted almost 500,000 interviews over 56 rounds of data collection in 11 countries in Africa and Asia. [28, 29] PMA2020 has trained a network of 2,700 female resident interviewers who carry out FTF surveys in their communities, using cell phones equipped with Open Data Kit software.

### 3.6.1 PMA Methodology

*Sampling*

PMA2020 FTF surveys use a two-stage stratified cluster survey design, typically with urban-rural strata, but in some countries regions or districts are the strata. [88] The survey draws a sample of enumeration areas (EA), which are geographic areas defined by the national census, using probability proportional to size sampling. [27] All households within selected EAs are listed by resident interviewers and then, a random sample of 35 households per EA are selected in each round of data collection. For the selected households, interviewers complete a household roster and questionnaire and then invite women ages 15 to 49 to participate in a female survey. Specifically, females who are usual members (*de jure)* or slept in the selected households the night before data collection (*de facto)* are interviewed. The sample size is determined based on the desired precision of the modern contraception prevalence rate, usually with a margin error of ± 2-3% points nationally, and ± 3-5% points for urban and rural strata. [89]

*Data collection*

PMA2020 interviewers and supervisors are trained over a two-week period on the topical content of surveys and well as survey implementation, including household listing, interviewing, and data collection using smart phones. After the initial training, interviewers and supervisors complete a three-day refresher training before each round of data collection. Interviewers conduct FTF interviews with the head of household and all eligible females. Responses are recorded on smart phones and sent to a secured cloud server. Johns Hopkins staff and in-country managers monitor data collection in real time to track progress and assess data quality. Data collection for each round typically takes two months.

*Measures*

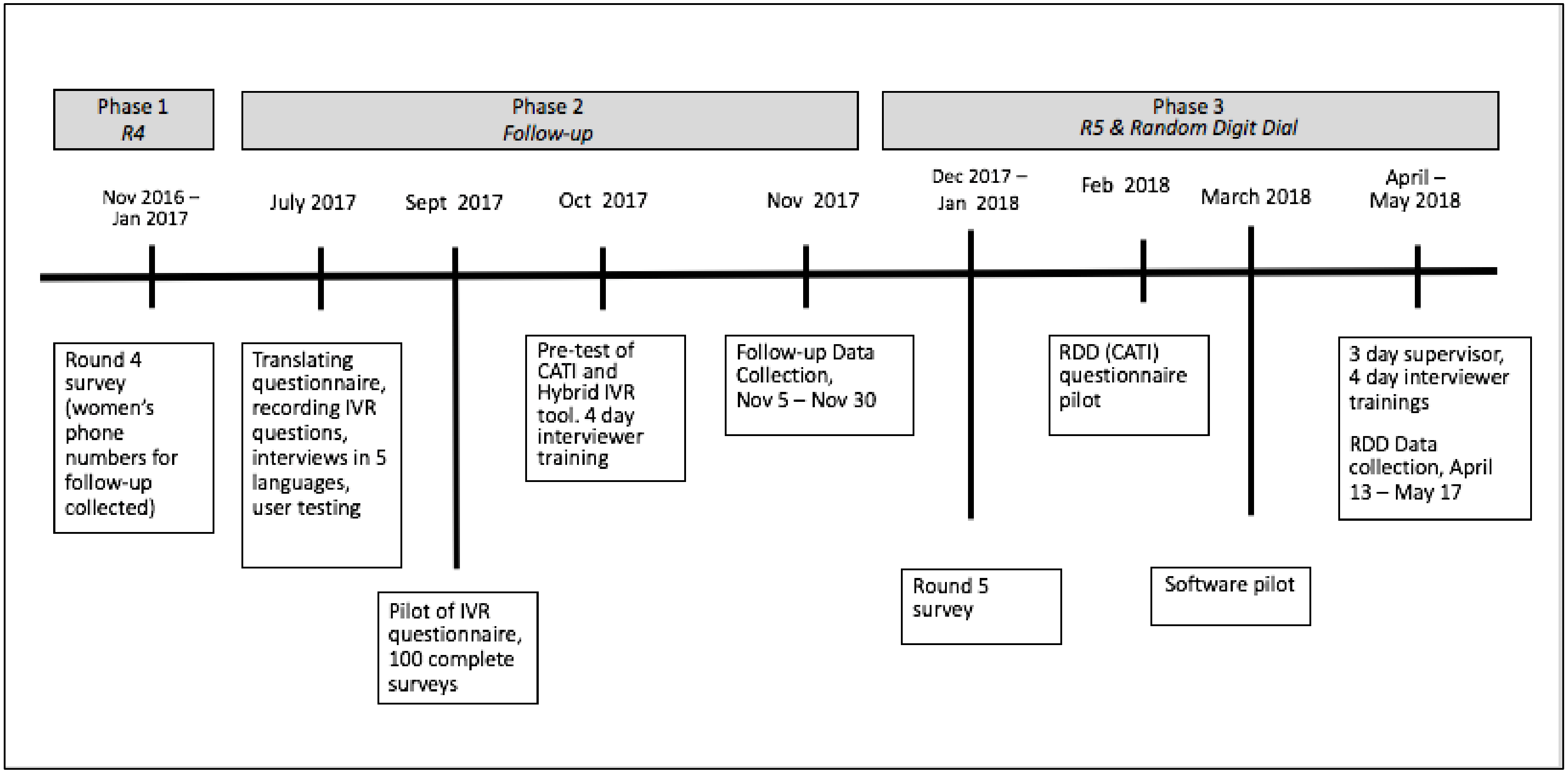
PMA2020 conducts three types of surveys: household, female and service delivery point (SDP). For both the female and household survey, PMA2020 uses a standardized questionnaire, based on the standard DHS questionnaires. [90] The household questionnaire asks the head of household to list household members, count assets and livestock, and about water, sanitation, and hygiene conditions. The two objectives of the female survey are to record women's sociodemographic characteristics and measures of reproductive health. Measures include dates of women's first and two most recent births, age at first sex and at first marriage, and age and parity at first contraceptive use. Contraceptive questions include current use of contraception, contraceptive use 12 months preceding the interview among current non-users, reasons for not using or stopping a method of contraception, intention to use contraception in the future among non-users and autonomy and influences related to contraceptive decision-making. [89] Details on the SDP questionnaire, which is not pertinent to this dissertation, are available elsewhere. [89]

### 3.6.2 PMA2020 in Burkina Faso

In Burkina Faso, PMA2020 partners with the *Institute Supérieur des Sciences de la Population* (ISSP). The PMA2020 surveys were launched in Burkina Faso in 2014. ISSP has since collected five rounds of nationally-representative data. In Burkina Faso, the regions serve as sampling domains and EA are selected within these domains using probability proportional to size. Rounds 1 and 2 were conducted in 53 EA. Starting in the third round, PMA2020 sampled 83 EAs; 41 rural and 42 urban. Response rates are generally high, both for household surveys (at least 97.9% since Round 3) and female surveys (97.7% in Round 5).

Both the Johns Hopkins University and ISSP Institution Review Boards reviewed the PMA2020 protocol. The PMA2020 core survey has exempt status (IRB #00000287, exempt as "public health practice", defined by DHHS regulations 45 CFR 46.102), as determined by JHU's Institutional Review Board. Approval was granted in July 2014 by the ethical committee associated with ISSP at the University of Ouagadougou, *Comité d'éthique pour la recherche en santé* (IRB #2014-7-81).

*Figure 3-2. Key activities and dates for Burkina Faso remote data collection project*



## 3.7 ***Aim 1***

This aim assessed sociodemographic factors related to cell phone ownership among women of reproductive age in Burkina Faso.

### 3.7.1 Study Design

Analysis was based on PMA2020 R4, a cross-sectional national probability survey. The study was conducted in all 13 regions in Burkina Faso between November 2016 and January 2017. The study received IRB approval from the Johns Hopkins Bloomberg School of Public Health and from the ethical committee in Burkina Faso, *Comité d'éthique pour la recherche en santé*.

### 3.7.2 Study Sample

Using the same sampling and data collection methodology as described in Section 3.6.1, PMA2020 R4 included a total of 2,751 households and 3,215 women aged 15 to 49 years. The response rates were 97.9% for the household interview and 95.4% for the female interview. [84]

### 3.7.3 Sample Size

The sample size was restricted to women who completed the R4 survey (3,215). The outcome (owning a cell phone) among the reference group (15 to 19 year old women) was 40%. In order to detect an odds ratio of owning a cell phone of 2 according to women's age group (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49) with a power of 80% and alpha of 0.05, we needed a sample size of at least 107 women per group. Thus, the study had adequate sample size (seven age groups, 107 needed per group; 749 total).

### 3.7.4 Sample Weights

All PMA2020 FTF data in this dissertation are weighted to adjust for complex survey design and non-proportional sampling scheme with sampling weights. Specifically, we used design weights to adjust for unequal probabilities of selection at each sampling stage (EAs and households). The weight corresponds to the inverse probability of EA selection and of household selection. We also applied non-response weights to adjust for non-participation, at the household and woman level. [91]

### 3.7.5 Variables of Interest

In Aim 1, the outcome was self-reported cell phone ownership, a dichotomous variable. During the Burkina Faso R4 female survey, interviewers asked women "How many phone numbers do you have?" Women were considered cell phone owners if they reported one or more phone numbers. When examining the impact of frame bias, the outcome of interest was modern contraceptive use. We defined modern contraceptive use as a dichotomous variable (yes vs. no) using data from two questions in the PMA2020 survey. The first question asked respondents if she or her partner was using anything to prevent or delay pregnancy. If the respondent said yes, she was asked to report what method of contraception she was using. If she was using a modern method, she was classified as a modern contraception user. Modern methods, as defined by the World Health Organization, include pills, implants, injectables, intrauterine device, condoms, female and male sterilization, lactational amenorrhea method (LAM), emergency contraception, and standard days method. [92] Long-acting reversible contraceptive methods include implants and IUDs, methods that provide protection years at a time but can be reversed. [93] We also constructed a 5-category indicator of method mix, distinguishing the following contraceptive methods: implant, IUD, injectables, pills, and condoms.

### 3.7.6 Independent Variables

We chose independent variables based on the literature conducted in LMIC that showed education and wealth are consistently correlated with cell phone ownership [94, 95] as well as urban-rural residence. [96] All considered, independent variables included women's sociodemographic characteristics such as current union status (in union – i.e., currently married or living with a partner vs. not in union), residential area (urban vs. rural), highest level of school ever attended (none, primary, or secondary and higher), household wealth (asset score quintiles: lowest quintile vs. all other quintiles; highest quintile vs. all other quintiles), and having electricity (yes vs. no). Age was categorized into seven groups (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49) when comparing phone owners and non-owners in descriptive and regression analyses. Household wealth in PMA2020 surveys is a summary measure of household assets, based on Principal Component Analysis (PCA) method, which is also used in the DHS surveys. [97] Although electricity is conventionally used to calculate household wealth index, in this analysis, electricity was also a separate co-variate because cell phone use requires access to electricity.

### 3.7.4 Analytical Plan

We first examined patterns of missingness of outcomes and covariates and explored the distribution of each variable. We looked at the range of age values and calculated the mean before creating a categorical age variable as previously described.

We compared sociodemographic characteristics of cell phone owners and non-owners using chi-square tests. We then examined these bivariate associations using a logistic regression model to estimate the unadjusted odds of cell phone ownership by age, education, wealth and location.

We then conducted a multivariable logistic regression to estimate the adjusted odds of cell phone ownership by sociodemographic characteristics. We tested for multi-collinearity using the Variance Inflation Factor (VIF) command in STATA. [98]

$\text{Logit}(Y_i) = \beta_0 + \beta_n X_{ni}$

$Y_i$ = (outcome): cell phone ownership for subject i
$\beta_0$ = intercept
$\beta_n$ = log-odds associated with n-th variable X
$X_{ni}$= vector of sociodemographic characteristics for subject i

After examining the sociodemographic characteristics of female cell phone owners, we compared prevalence of modern contraceptive use between cell phone owners and non-owners using a chi-square test to assess the effect of sample selection by cell phone ownership status on modern contraceptive use estimates.

Finally, we assessed method mix, which is the percent distribution of modern contraceptive users by method, in a time period. [99] We compared the distribution of modern method mix among cell phone owners and non-owners, using a chi-square test. [99] Because of complex survey design, all aforementioned chi-square tests were adjusted with Rao and Scott's first order correction method.

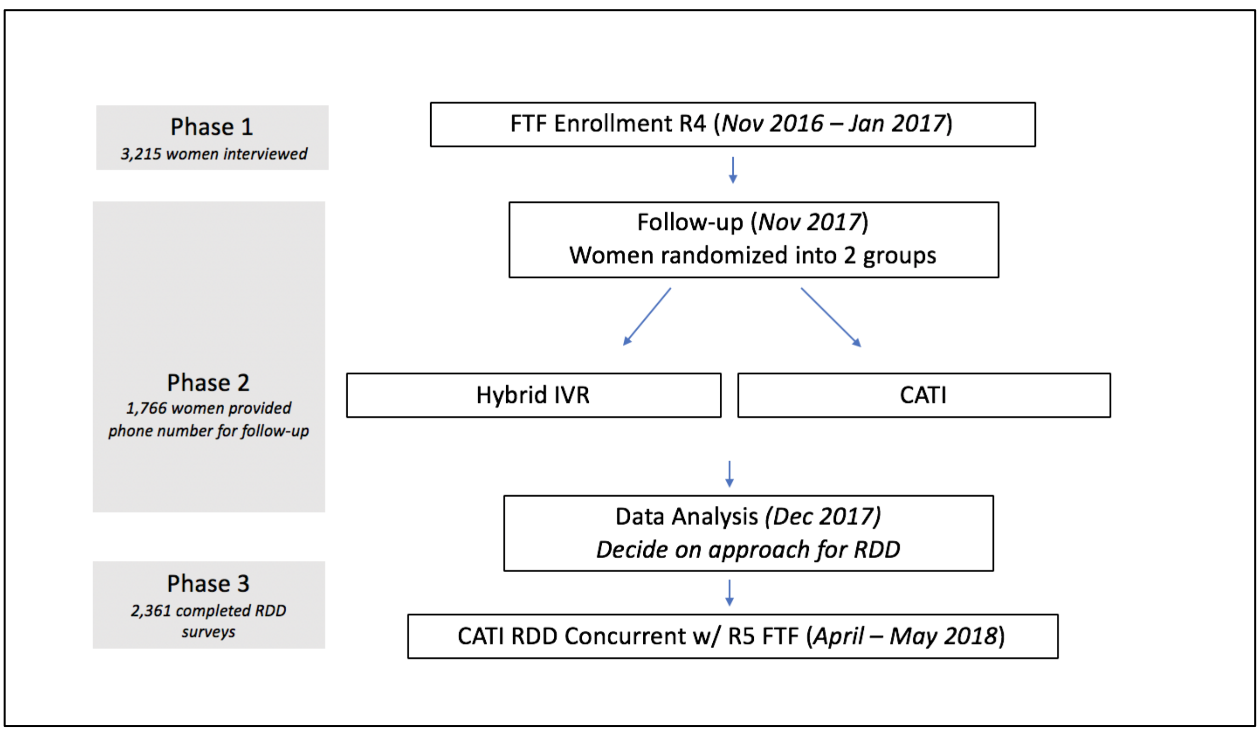
## 3.8 ***Aim 2***

Aim 2 examined the sociodemographic characteristics associated with consenting to and completing a follow-up phone survey, by survey mode (CATI and Hybrid IVR), among R4 female cell phone owners.

### 3.8.1. Study design

This analysis was based on a follow-up study of PMA2020 R4, randomizing female respondents who own a cell phone into two groups, one receiving a CATI phone survey, the other receiving a Hybrid IVR phone survey. Ethical approval was granted by the Institutional Review Board at the Johns Hopkins Bloomberg School of Public Health (IRB No. 00007961) and the by *Comité d'éthique pour la recherche en santé* (IRB No. 2017-8-0114).

*Figure 3-3. Study design for the follow-up survey*



### 3.8.2 Study Sample

The sample included 1,766 women interviewed in PMA2020 R4 who reported cell phone ownership (57% of initial sample) and consented to be followed-up (96.9%). The aforementioned 1,766 women were randomly assigned to two arms after stratification by survey language (five languages and one 'other' group): 882 were randomized to receive Hybrid IVR and 884 to receive CATI (as explained below).

### 3.8.3 Sample Size

We calculated our sample size by estimating the number of women we would need to complete the survey in order to detect a difference in education (defined in three categories; no education, primary and secondary education) between CATI and Hybrid IVR. Education was selected due to concerns over Hybrid IVR survey completion among a low literacy population.

Because our survey was restricted to a female population, who are more difficult to reach than men in SSA, [70, 100] we estimated a 30% loss to follow-up between the FTF and phone survey. We then assumed an 80% response rate in the CATI arm (n=494) and a 40% response rate in the Hybrid IVR arm (n=247) (Table 3-2). These assumptions were based on information from Viamo, indicating that 20 percent of people who provide a phone number complete an IVR follow-up survey. We expected a higher response rate due to the fact that the IVR survey opened with a human interviewer. A CATI study in Lebanon achieved an 80% response rate among follow-up respondents. [32] Finally, we hypothesized that women completing CATI and Hybrid IVR phone surveys would have a higher level of education than non-completers, but that the distortion would be greater in the Hybrid IVR arm. Starting with a baseline distribution of 21% women with a secondary education or higher among PMA2020 R4 cell phone owners, we assumed this proportion would rise to 25% in the CATI respondent sample and 35% in the hybrid IVR respondent sample).

With a power of 0.80, an alpha of 0.05, we ultimately estimated that we needed 231 women per group to detect an absolute difference of 10% in the proportion of respondents with secondary higher education between CATI respondents and Hybrid IVR respondents corresponding to p2=0.35. Thus we consider that the available sample size of phone owners in PMA2020 R4 is sufficient to conduct the study.

*Table 3-2. Follow-up sample size calculation*

| | Arm A – Hybrid IVR (n=882) | Arm B – CATI (n=884) |
|---|---|---|
| Number of women contacted for the phone follow-up survey (30% attrition) | 617 | 618 |
| Response rate | 40% | 80% |
| Number of women completed the phone follow-up survey | 247 | 494 |

### 3.8.4 Questionnaire

The CATI and Hybrid IVR surveys used the same wording in questions as the PMA2020 FTF questionnaire with only minor modifications. The phone survey (CATI and Hybrid IVR) included 17 questions: five introductory questions to identify the respondent, four demographic questions (age, area of residence, marital status and parity), five questions about the respondent's awareness of modern contraceptive methods (IUD, implant, condom, pills and injectables) and three questions on contraceptive use (current use, current method used or pregnancy intention).

Questions were identical prior to consent 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 drop. Following consent, women receiving the Hybrid IVR survey heard an explanation of what to expect during an IVR survey, and then were asked to press 1 on her keypad. If the respondent was unable to press 1, she was considered incapable of participating in the IVR survey, and the interview ended. If the woman pressed 1, the respondent 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 asked the respondent which region she lived in and the last IVR question asked the respondent to enter her age.

Of the 17 survey questions, one question was numeric (age), three were multiple choice, and 12 were binary with yes/no response option. The key question "What are you or your partner currently using to delay or avoid a pregnancy?" was field-coded by interviewers in CATI and a multiple choice question in the IVR survey. For the IVR survey, the school question was broken into two questions, first asking if the respondent had ever been to school, and if the respondent answered affirmatively, asking her to what level she had studied. Both the CATI and Hybrid IVR survey were conducted in five languages. The questionnaire was written in French in the software, and CATI interviewers translated in real-time, after having developed and refined consistent translations in the local language during the training.

### 3.8.5. Study procedures

*Software development*

Viamo developed and designed the CATI and Hybrid IVR software in close collaboration with PMA2020 researchers. 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 standard fixed telephone line. Each interview required two phone lines: one for the operator and one for the respondent, joined through the software. The software automatically dialed phone numbers and recorded the outcome of each call attempt. The questionnaire was displayed on the tablet screen. The data was stored on Viamo's cloud server and downloaded daily by PMA2020 staff.

*Pilot 1: Translation, Recording & Pre-Testing IVR Survey*

The automated nature of an IVR survey affords the respondent no opportunity to clarify a survey question. Therefore, quality translation is critical for the success of an IVR survey. The goal of the pilot was to identify any IVR questions that needed to be re-recorded due to poor translation or audio quality. We adhered to translation best practices as closely as possible, following the Translation, Review, Adjudication, Pretesting and Documentation (TRAPD) approach. [101] To translate the survey from French into four local languages: Dioula, Fulfulde, Gourmantchema and Moore, we convened a group of language experts: one ISSP staff, three PMA2020 interviewers, one translation specialist and the woman who would ultimately record the survey. The questions were subsequently recorded in five languages in a professional recording studio by women who worked as radio personalities (with the exception of the Gourmantchema-speaker).

Once the recordings were finalized and uploaded to Viamo's IVR software, we conducted one day of user testing with 3-4 participants per language. The testing identified questions that needed to be re-recorded due to poor translation or poor audio quality. Over half of the Gourmantchema questions had to be re-recorded, but the other languages had minimal edits. During the second round of recording, translators joined the recording session when possible to ensure quality of the translations.

*Pilot 2: RDD of 100 complete IVR Surveys*

In September 2017, we conducted an IVR random digit dial pilot survey on Viamo's platform, with the end goal of conducting 100 surveys. We used the information from this pilot to identify any questions leading to high hang-up or refusal and to estimate the length of the IVR survey. We also embedded an A/B test [102] to assess if two encouraging messages would increase response rates compared to no encouragement messages. The response rate was higher for those who did not receive the encouraging message, thus we decided not to include the encouraging message in the final version of the phone survey.

*Pilot 3: CATI Pre-Test*

With the goal of identifying any problematic questions, Kantar International, the call center implementing partner, used an internal list of valid phone numbers to call respondents and complete 10 CATI questionnaires the month before data collection. No major changes were made to the script after this pilot.

### 3.8.6 Data Collection

We trained 20 female interviewers and retained 15 for data collection. Interviewer training was four days, and interviewers piloted the survey for two days. We calculated the percent of female phone owners from PMA2020 R4 who spoke each of the five languages during the baseline FTF survey and hired interviewers according to this language distribution (Gourmanthcema – 1, Fulfulde – 1, Dioula – 3, Moore – 6, Exclusively French – 4).

Each interviewer was assigned approximately 120 respondents. Most interviewers were not married (87%). Most were currently enrolled or had finished university (73%) and were nulliparous (66%). Just under half (46%) had previous survey experience but none had conducted PMA2020 surveys in the past.

The call center, managed by JHU and ISSP researchers, in partnership with Kantar International, was in a centrally located conference center in Ouagadougou. Interviewers were split into three shifts, morning, mid-day and afternoon, and worked at the same time each day throughout data collection. Interviewers only administered one remote data collection mode (either CATI or Hybrid IVR) for the duration of the study.

Data collection took place November 5 – 17, 2017. All respondents were called up to six times at varying times of day and were called at least once during the weekend. If a respondent picked up and wanted to be called back within 15 minutes, the operator could accommodate the request, otherwise the respondent was called back the next day. A respondent was not called back if she refused to consent to the study or if someone answered the call that did not know the respondent. All respondents saw the same phone number appear on their cell phone, which was identifiable as a landline phone number. The call center phone number could not be called back. Women who completed the survey were sent the equivalent of $1 US dollar phone credit the day after completing the interview.

### 3.8.7 Measures

*Call Outcomes*

We used the American Association for Political Opinion Research (AAPOR) 9th edition disposition codes to classify phone survey participants into standardized categories. AAPOR's bi-annual document, *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, provides guidelines for classifying respondents, with the aim of standardizing key survey rates by using consistent language. [103] A final disposition code is the final call outcome for each woman. For example, if a woman is called three times, does not pick up the first call, picks up the second call but is unable to speak at that time during the second call, then finally refuses participation on the third call, the disposition code is refusal, per the outcome last (third) call. Disposition codes are useful for internal monitoring of phone survey procedures as well as for the comparison of levels and causes of phone sample selection across studies. [50] Finally, using AAPOR's standard definitions allows research to create transparent, replicable and comparable measures between surveys.

In this study, all participants (PMA2020 R4 cell phone owners) were eligible for the phone follow-up survey and were classified as respondents or non-respondents. Using the AAPOR call disposition classification, adapted for the specific purposes of our study comparing CATI and IVR response rates, we defined eight categories as shown in Table 3-3. The first six non-respondent categories included two non-contacts groups: calls that were never picked up after six attempts ("did not pick-up") and calls that were picked up but for which the eligible respondent was not found ("picked up but woman not found"). Next, we defined two refusal categories according to whether the woman refused participation before or at the time of consent. Finally, we identified two types of break-off. After consent, women randomized to the Hybrid IVR survey were asked to press 1 on their keypad before being transferred to the IVR survey. If the woman was unable to press 1 on her keypad, the survey ended and she was classified as an IVR break-off. All other women who had consented to participate but answered less than 50% of the 17 questions were classified as break-offs. Finally, survey respondents were divided into two categories, partial completers who answered between 50% and 80% of the questionnaire and completers who answered 80% of the questionnaire or more.

*Table 3-3. Individual Call Outcomes (Final Disposition Code) Definitions*

| **AAPOR Categorization** | **Final Disposition Code** | **Explanation** |
|---|---|---|
| **Non-Respondent** | | |
| NC (2.20) | Non-contact (did not pick up) | No phone calls were picked up over the 6 attempts |
| O (2.36) | Non-contact (someone picked up the phone call but interviewer never spoke with the woman) | A phone call was answered but the intended respondent was either unknown, or never available to speak to the interviewer |
| R (2.12) | Refusal pre-consent | The respondent refused to participate in the study before consent |
| R (2.111) | Refusal | The respondent refused to participate in the study at the time of consent |
| R (2.121) | IVR Break-off | The respondent consented to participate but was unable to push 1 on phone (IVR test) |
| R (2.12) | Break-off (consented but less than 50% completion) | The respondent consented to participate but answered less than 50% of the questions |
| **Respondent** | | |

| | | |
|---|---|---|
| P (1.2) | Partial completer (answered 50-80% of questions) | The respondent consented to participate and answered 50% to 80% of the questions |
| I (1.1) | Completer (answered more than 80% of questions) | The respondent consented to participate and answered more than 80% of the questions |

*+values based on final disposition code value in AAPOR 9th edition of Standard Definitions, page 75*

### 3.8.8 Variables of interest

Building on the AAPOR disposition codes, we defined two outcome measures to assess non-response distortion in our CATI and Hybrid IVR phone follow-up surveys: consenting (yes/no) and completing (yes/no). Consent was not a traditional survey outcome used by AAPOR, but was defined to reflect the fact that up to the consent question, Hybrid IVR and CATI followed an identical protocol (interviewer asking the same questions). A woman was considered as having consented if she said "yes" at the consent question. Non-consenters included the two non-contacts categories and the two refusal categories. The other dependent variable was survey completion, defined as answering 50% or more of the relevant survey questions corresponding to AAPOR partial completers and completers.

*Reproductive health indicators*

To evaluate the implications of phone survey non-completion on estimates of modern contraceptive use (non-response bias), we used the same indicator of modern contraceptive use in Aim 1. Women were categorized into those using any modern method vs. non-users of modern methods. Modern methods in Burkina Faso include pills, implants, injectables, intrauterine device, condoms, female and male sterilization, lactational amenorrhea method (LAM), emergency contraception, and standard days method. [92]

### 3.8.9 Independent variables

We considered the same sociodemographic factors as Aim 1, informed by the literature on cell phone survey response in LMIC [70, 72] and on modes of data collection in SSA. [62, 104-106] Sociodemographic information was collected during PMA2020 R4 and therefore available for all eligible women, regardless of whether or not the woman responded to the phone follow-up survey. The follow-up phone survey data was only used to assign final disposition codes and thus calculate survey outcome rates.

The key independent variables were age, which was categorized into seven groups based on descriptive analysis (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49), and residential area (urban vs. rural). Additional co-variates of interest included current union status (in union – i.e., currently married or living with a partner vs. not in union), highest school ever attended (none, primary, or secondary and higher), household wealth (lowest middle and highest tertiles), and having electricity (yes vs. no), and parity (ever given birth vs. never given birth). When survey completion was the outcome, we used four age groups (15-19, 20-29, 30-39, 40-49) since only a small number of women completed the Hybrid IVR survey.

### 3.8.10 Analytical plan

We first examined whether randomized groups (CATI and Hybrid IVR) were similar according to their sociodemographic characteristics at baseline (PMA2020 R4 survey).

We then used AAPOR guidelines to estimate the four call outcomes by mode of data collection: response, cooperation, refusal and contact rates (Table 3-4).

- AAPOR's ***response rate*** corresponds to the number of individuals who complete the phone interview, fully or partially, over all dialed phone numbers assuming all phone numbers are eligible for the study. AAPOR distinguishes two subcategories of response rates, based on whether respondents completed the 80% (Response rate 5) or 50% of relevant survey questions (Response rate 6).
- AAPOR's ***cooperation rate*** is similar to response rate, but excludes non-contacts from the denominator. Two cooperation rates are used in this study: Cooperation rate 1 only includes complete interviews in the numerator while Cooperation rate 2 also includes partial interviews.
- AAPOR's ***refusal rate*** corresponds to the number of individuals who refuse to be interviewed over all dialed phone numbers assuming they are all eligible.
- Finally, AAPOR's ***contact rate*** corresponds to the number of calls in which one member of the unit was reached (i.e. someone at the target phone number picked up the phone); among all eligible individuals.

*Table 3-4. AAPOR Survey Outcome Rates*

| Response Rates | Explanation |
|---|---|
| Response rate 5: $\frac{I}{I+P+R+NC+O}$ | All individuals who complete more than 80% of survey / All eligible individuals |
| Response rate 6: $\frac{I+P}{I+P+R+NC+O}$ | All individuals who complete more than 50% of survey / All eligible individuals |
| **Cooperation Rates** | |
| Cooperation rate 1 : $\frac{I}{I+P+R+O}$ | All individuals who complete more than 80% of survey / Eligible individuals who were ever contacted |
| Cooperation rate 2: $\frac{I+P}{I+P+R+O}$ | All individuals who complete more than 50% of survey / Eligible individuals who were ever contacted |
| **Refusal Rate** | |
| Refusal rate 3: $\frac{R}{I+P+R+NC+O}$ | All individuals who refused to complete the survey / All eligible individuals |
| **Contact Rate** | |
| Contact rate 3: $\frac{I+P+R+O}{I+P+R+NC+O}$ | All phone numbers that answered the call / All phone numbers |

Turning to our analysis of factors associated with phone survey consent and survey completion, we compared the distribution of the aforementioned sociodemographic characteristics between consenters and non-consenters and between completers and non-completers, using chi-square tests. Analysis were stratified by mode of data collection (CATI and Hybrid IVR).

We then conducted multivariable logistic regression models to identify the independent factors associated with survey consent. We performed one model for each survey mode (Hybrid IVR and CATI). Next, we used multivariable logistic regression models to identify the independent factors associated with survey completion, and stratified the analysis by survey mode (Hybrid IVR and CATI). We also directly compared Hybrid IVR and CATI completers to evaluate if sample distortion was different between the two modes, given the lower response rate in IVR. We did not include marital status nor parity in the multivariable models because these characteristics were not significantly related to consent or completion in the bivariate analyses. Electricity was also not included because of high correlation with wealth tertiles. We checked for co-linearity and found no variance inflation factor of more than four. [98] We conducted analyses in Stata version 15 (StataCorp 2017). The multivariable logistic regression is specified below.

$\text{Logit}(Y_i) = \beta_0 + \beta_n X_{ni}$

$\text{Logit}(Y_{ii}) = \beta_0 + \beta_n X_{ni}$

$Y_i$ = (outcome): ***consenting*** for subject i (per mode)
$Y_{ii}$ = (outcome): ***completing*** for subject i (per mode)
$\beta_0$ = intercept
$\beta_n$ = log-odds associated with each co-variate
$X_{ni}$= vector of sociodemographic characteristics for subject i

Finally, we compared the prevalence of modern contraceptive use between survey completers and non-completers using chi-square tests, for both modes of data collection, Hybrid IVR and CATI.

## *3.9. Aim 3*

Aim 3 compared the prevalence of modern contraceptive use among women of reproductive age in Burkina Faso between two samples obtained from different survey methodologies: a FTF national probability survey (PMA2020 Round 5) and an RDD cell phone survey administered using CATI.

### 3.9.1 Study design

PMA2020 R5 was a cross-sectional, national probability survey conducted in Burkina Faso at the end of 2017. The RDD CATI phone survey was conducted between April and May 2018. We chose to use CATI rather than Hybrid IVR based on results of Aim 2. This decision was reflected in the amended protocol submitted to the JHU Institutional Review Board. The ISSP team submitted a new protocol to their ethical committee for the RDD survey (IRB No. 2018-3-036).

### 3.9.2 Study sample

#### 3.9.2.1 PMA2020 R5 sample

Using the same procedures as PMA2020 R4 described in Section 3.6, PMA2020 R5 used a two stage probability sampling design to include 2,811 households (98.5% response rate) and 3,659 females aged 15 to 49 years from the selected households (97.8% response rate). [81]

#### 3.9.2.2 RDD Study sample

In most LMIC, a sampling frame of phone numbers is difficult to obtain, [48] and the same is true for Burkina Faso. Therefore, we chose to conduct the phone survey using RDD sampling. To create the phone numbers to randomly call, Viamo, an international mobile technology survey company, [107] used a list of the 25 existing prefixes provided from the three mobile network operators in Burkina Faso. Viamo randomly generated the remaining six digits to create a list of phone numbers. In addition, we implemented quotas to improve RDD sample representation by age and area of residence. [48] A quota is a limited quantity of a certain outcome, in our case, the number of completed interviews, by age (15-19 years old or 20-49 years old) and by geographic location (urban or rural). We established the quota groups by comparing the follow-up respondents and non-respondents from Aim 2 to see which groups were less likely to respond to the phone interview. We found that rural, young and uneducated women were least likely to respond. Our quota groups only included residence and age for ease of implementation. We then used our target sample data, PMA2020 R4 female respondents, to calculate the percent of women in each of the four groups (see Table 3-5). We proportionally selected our target sample size from the sample distribution of PMA2020 for each quota group. Once we reached the target number of completed interviews in a group, the questionnaire was programmed by the cellphone survey administration firm, Viamo, to automatically end the interview after thanking the respondent.

*Table 3-5. The number of completed surveys by quota group*

| | Rural | Urban |
|---|---|---|
| **15-19** | *388*<br>*(16.4%)* | *129*<br>*(5.5%)* |
| **20-49** | *1417*<br>*(60%)* | *427*<br>*(18.1%)* |
| **Total** | | **2,361** |

### 3.9.3 Sample size

This aim assessed the equivalence of modern contraceptive use estimates between an RDD survey and a concurrent FTF survey in Burkina Faso. We considered the FTF and RDD estimate to be equivalent if the outcome of interest (contraceptive prevalence rate) for both approaches were within a four percent margin of error. The FTF R5 (PMA2020) sample was projected to include 3,500 women; because of the complex survey design of PMA2020, we expected the effective sample size would be approximately 2,334 women, with an expected design-effect (deff) of 1.5. Our power calculation suggested that we needed a minimum 1,574 effective sample size in the RDD survey to detect equivalence in two estimated rates (with FTF modern contraceptive use expected to be 23%) with a margin-of-error of four percent between the FTF and RDD estimates. Post-stratification weighting increases deff, thus we targeted (1,574 X 1.5 (deff)) 2,361 complete interviewers for the RDD method. [108]

RDD has low response rates in SSA, mainly due to the lack of a sampling frame of existent numbers. Considering that a minimal number of phone numbers dialed would be valid, and that reaching women for cell phone surveys is more difficult than reaching men, we estimated that we needed to contact 38,578 phone numbers to attain the goal sample size of 2,361 women. A summary of the sample size calculation is presented in Table 3-6.

*Table 3-6. Estimated number of CATI RDD attempted calls in order to achieve sample size*

| | **Calls attempted** | **Calls connected** | **1st question answered (language)** | **Meets eligibility (female & 15 - 49)** | **Completed surveys** |
|---|---|---|---|---|---|
| Overall percent (denominator = calls attempted) | | 20% | 17% | 7% | 6% |
| % of calls from previous step | | 20% | 85% | 40% | 90% |
| (N) | 38,578 | 7,716 | 6,558 | 2,623 | 2,361 |

### 3.9.4 Questionnaire

The RDD questionnaire contained 17-19 questions (depending on skip patterns) to specifically address the research question related to comparison of modern contraceptive use estimates with an FTF survey. Most of the questions had been used in the CATI phone survey described in Aim 2. Four questions helped establish the eligibility of the respondent, followed by five to six demographic questions, five questions about contraceptive awareness, and three or four questions on contraceptive use. The RDD questions were identical to the FTF questions with a few adaptations to accommodate administration over the phone. The RDD survey was available in French and six local languages.

### 3.9.5. Study procedures

*Pilot #1: Questionnaire Pilot*

There were four objectives for the first pilot, which took place over two days in February 2018: 1) count the number of phone calls needed to compete 20 surveys when applying quotas 2) confirm the appropriateness of the opening of the survey, 3) ensure survey questions were well-written and that the answer options were exhaustive, and 4) document the duration of the survey. Two interviewers from the follow-up study, supervised by the ISSP project manager, called 723 phone numbers to complete 20 interviews; approximately 36 calls per completed interview. The majority (58%) of dialed phone numbers were invalid. Among the 42% of phone calls that were picked up, 51% were picked up by men, 23% were refused, 9% were ineligible (age or quota filled), 7% did not speak the same language as the interviewer and 10% of calls lead to a completed interview. The interviewers and supervisors identified several questions that needed to be re-worded. The second question, which asks the respondent which gender he or she is, was immediately identified as problematic, precipitating re-wording for the survey opening. The residence question (urban or rural) was simplified. The regional location question was changed because too few women could name their region, so instead we asked which chiefdom a woman lived nearest to. Finally, we broke the education question into two parts, first asking “Have you ever been to school?” then asking “What is the highest level of education you have attained?”

*Pilot 2: Software Pilot*

The second pilot took place in March 2017, with the overall objective of testing the software, by completing 20 interviews. The pilot objectives were to 1) identify problems or challenges with the software, and 2) ensure that the questionnaire was properly programmed. The pilot revealed technical problems with the software, which were subsequently addressed.

### 3.9.6 Data Collection

PMA2020 Baltimore and ISSP staff trained five supervisors for three days. Supervisors had at least a Bachelor degree and all had previous survey research experience. We trained 25 interviewers, of whom approximately eight had worked previously for PMA during FTF surveys, and 11 had worked during the follow-up phase of this study (Aim 2) but never for a PMA FTF survey. Interviewer training lasted four days and we retained 20 interviewers for data collection. All interviewers spoke French and at least one local language except for one interviewer who spoke exclusively French. Most interviewers spoke Moore (nine), and four interviewers spoke both Moore and Dioula, and three interviewers spoke only Dioula. We also had one interviewer that spoke Fulfulde and one that spoke Gourmantchema.

Data collection was managed by ISSP and took place from April 13 to May 17, 2018 in a call center located in an NGO building in Ouagadougou. Interviewers came in two shifts, the first from approximately 12 – 4 pm, the second group from 4 – 8 pm.

Because a significant proportion of phone numbers (58%) were identified as invalid during the pilot in February, we sent an IVR "validation/pre-notification" phone call before a phone number was released to an interviewer, with the goal of eliminating invalid numbers from interviewer's call lists. An invalid number was defined as a number that did not have an International

Telecommunications Union hang-up cause that indicated either the call was picked-up or the call was successfully placed. The pre-notification calls went out one to seven days in advance of an interviewer calling that number. The pre-notification call was recorded in Moore and said "Thank you for responding to our call. We will call you this week for a study. Please, pick up the phone when we call. Have a good day!"

For the first seven days of data collection, we called phone numbers that were identified with any one of the three call statuses during the validation calls: 1) No Answer, 2) Normal Clearing or 3) Normal Unspecified (Table 3-7). [109] However, after the first seven days of data collection, due to an insufficient number of completed interviews per day, we narrowed the definition of a valid phone number and no longer included phone numbers marked as 'No Answer' during the validation call. The narrower definition meant that interviewers were now only calling phone numbers that were answered during the validation calls.

*Table 3-7. Hang-up causes, defined by the International Telecommunications Union (ITU)*

| **Hang-up cause** | **ITU Definition[++]** | **Explanation** |
|---|---|---|
| No answer | The called party has been alerted but does not respond with a connect indication within a prescribed period of time. | No one picked up the call but the call went through |
| Normal clearing | This cause indicates that the call is being cleared because one of the users involved in the call has requested that the call be cleared. Under normal situations, the source of this cause is not the network. | Someone picked up the phone call |

| | | |
|---|---|---|
| Normal unspecified | This cause is used to report a normal event only when no other cause in the normal class applies. | No errors were detected, but it cannot be confirmed if the call was picked up. We only included phone numbers that had a call length of 13-14 seconds.+ |

+ The length of the prenotification message was 14 seconds. Thus only including calls of that length was an indirect approach to confirm the call was answered.
++ [109]

Once a phone number had been validated, it was added to a list that was then assigned to an interviewer. Lists were given to interviewers weekly and the majority of phone numbers were called once. The exception was when an interviewer exhausted her list of phone numbers and did not yet have a new list of numbers so she would call phone numbers back that did not pick up the first call she placed.

Once a female respondent picked up, she could either complete the survey at that time or be called back up to six times. The respondent could not specify specific times to be called back, except if she wanted to be called back within 30 minutes. If a respondent answered that did not speak the same language as the interviewer, she would be called back the same day by an interviewer that spoke the respondent's language. If a respondent explicitly refused the study, she was not called back. Men were not allowed to pass the phone to a female in their household. Respondents were unable to call back the phone number that called them and all respondents saw the same phone number on their cell phone, which was identifiable as a landline phone number. Women that completed the survey were sent the equivalent of $1 US dollar phone credit the day after completing the interview.

### 3.9.7 Measures

#### 3.9.7.1 Call outcomes

We used the 9$^{th}$ edition AAPOR final disposition codes to classify call outcomes. [103] The invalid phone numbers identified during the screening process were not assigned a disposition code. The 13 disposition codes divided into four groups (Not Eligible, Unknown Eligibility - non-interview; Eligible - non-interview; Interview) are presented in Table 3.8.

Non-eligible respondents were categorized into four codes. Respondents were ineligible if they were male, or were >49 years and <15 years or did not speak one of the seven survey languages. The fourth group consisted of women who spoke one of the survey languages and were between the ages of 15-49 but were ineligible due to quota restrictions.

Unknown eligibility was captured in four disposition codes and consisted of respondents whose eligibility was not known. Calls that were never picked up were classified as “No answer” and calls answered by a voice mailbox were classified as “Telephone answering device”. Respondents who answered but for whom age, gender or area of residence was not known were classified as “No screener completed”. Finally, respondents who spoke one of the seven survey languages but who did not speak the same language as the interviewer and were not reached during subsequent attempts were classified as “Other (language not matched with interviewer)”.

The next group “eligible, not interviewed”, was divided into three codes and consisted of women ages 15-49 who spoke one of the seven survey languages and were not excluded due to quota restrictions. The first was refusal before consent, after eligibility was established. The second was refusal at consent and the final was “break-off”, corresponding to a consenting respondent who completed less than 50% of questions.

The final group included women who completed the interview, classified as a partial interview when 50-80% of questions were answered and a complete interview when 80% of questions or more were answered.

*Table 3-8. Call Disposition codes for RDD Survey*

| AAPOR Code | Title | Definition |
|---|---|---|
| | **Not Eligible** | |
| (4.71) | Gender (not female) | Male |
| (4.72) | Age | Female and age <15 or >49 years |
| (4.73) | Language | Female and none of the 7 languages available in survey |
| (4.8) | Quota Filled | Respondent was female and age-eligible but due to quota restrictions was not interviewed |
| | **Unknown Eligibility, non-interview** | |
| UH (3.13) | No Answer | Phone call not picked-up |
| UH (3.14) | Telephone answering device | Phone call went to voice mail |
| UH (3.21) | No screener completed – talked with respondent but hung-up or refused | Respondent picked- up the call but interviewer was unable to confirm eligibility |
| UO (3.90) | Other (Language not matched with interviewer ) | Respondent spoke one of seven survey languages but the interviewer did not speak the same language |
| | **Eligible, non-interview** | |
| R (2.111) | Refusal pre-consent but confirmed female and 15-49 | Eligible respondent refused to participate before consent |
| R (2.11) | Refusal at consent | The respondent refused the study during consent |
| R (2.1) | Break-off (consented but less than 50% of relevant questions answered) | The respondent consented but answered less than 50% of the questions |
| | **Interview** | |
| P (1.2) | Partial (50-80% of relevant questions answered) | The respondent consented and answered between 50-80% of the questions |
| I (1.1) | Complete (more than 80% of relevant questions answered) | The respondent consented and answered more than 80% of the survey questions |

#### 3.9.7.2 Variable of interest

The outcome of interest was a binary measure of modern contraceptive use, based on two questions, that were asked identically in the two surveys. The first question asked whether the respondent or her partner was currently using a form of contraception ("Are you or your partner currently doing something or using any method to delay or avoid getting pregnant?"). If the respondent answered positively, she was asked to specify the type of method used ("Which method or methods are you using?"). If the respondent identified a modern method (as specified below), she was classified as a user of modern contraception.

Traditionally, measures of modern contraceptive use include all modern contraceptives available in a country. In Burkina Faso, the PMA20200 FTF survey asked respondents about 12 modern contraceptive methods: male and female sterilization, implant, IUD, injectables, pill, emergency contraception, male condom, female condom, diaphragm, foam/jelly and LAM. However, the RDD survey only collected data about five methods: implants, injectables, pills, condoms and IUDs, which covers the 98.8% of modern contraceptive methods reported during R5. Thus we limited the definition of modern contraceptive use to these five methods for both the FTF and RDD surveys in this study. [81] Long-acting reversible contraceptive methods include implants and IUDs, methods that provide protection years at a time but can be reversed. [93] Finally, we constructed a five-category indicator of method mix, distinguishing the following contraceptive users: implant, IUD, injectables, pills, and condoms.

#### 3.9.7.3 Independent variables

Although we had a wealth of data about the PMA2020 R5 respondents, the RDD survey was shorter thus limiting the use of FTF data for co-variates. We selected our models after examining the existing literature about modern contraceptive use and the literature about phone ownership and/or phone surveys. To assess the difference in contraceptive use between women participating in the R5 PMA2020 survey and those participating in the RDD survey, we pooled both surveys and defined mode of data collection (FTF or RDD-CATI) as our key independent variable.

Covariates included women's sociodemographic characteristics, such as age (in 5-year age intervals), current union status (in union, i.e., currently married or living with a partner vs. not in union), residential area (urban vs. rural), highest school ever attended (none, primary, or secondary and higher) and language of survey (Moore, French, Dioula, Fulfulde or Gourmantchema). Sexual and reproductive health variables included parity (ever vs. never). We included province and number of children in descriptive analyses but did not include them in our regression model.

### 3.9.8 Analytical Plan

*Call Outcomes*

After assigning a disposition code to each woman called in the RDD study, we calculated the four key AAPOR survey outcome rates. The four key rates are response, cooperation, refusal and contact (Table 3-9).

Unlike the Aim 2 follow-up survey, the eligibility of all phone numbers was not known in the RDD survey. To improve the specificity of survey outcome measures, AAPOR recommends calculating rates that exclude an estimated number of unknown eligibility phone numbers from the denominator. We used the pilot data to estimate the percent of unknown eligibility calls that would be eligible. Based on pilot data collected in February 2018, we estimated that 20% of calls with unknown eligibility would include an eligible woman. Applying this correction, we defined Response rates 3 & 4 (same numerator as Response rates 1 &2 but denominator excluded 80% of unknown eligibility calls) and Contact rate 2 (same numerator as Contact rate 1 but denominator excluded 80% of unknown eligibility calls).

*Missing Data*

The RDD data had item non-response due to internet outages at the call center. The internet was intermittently cut due to electricity brown-outs, and responses during electricity cuts were not recorded (missing completely at random assumed). We used the hot deck method [110] to impute missing values for three variables: age (43 missing values, 1.8%), residence (10 missing values, 0.4%) and education (10 missing values, 0.4%).

*Weighting RDD sample*

To address RDD sample distortion, we created RDD post-stratification weights, based on three sociodemographic characteristics of women of reproductive age in Burkina Faso (area of residence, age and education). These factors were chosen based on phone ownership characteristics identified in Aim 1. The PMA2020 R5 sample served as the reference population in the absence of census data in Burkina Faso (last available census was conducted in 2006). We calculated the ratio of RDD respondents to R5 in urban/rural groups and seven age groups (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49). We then calculated the ratio of RDD respondents to R5 respondents in three education groups: no education, primary education and secondary education. Post-stratification weights were computed by multiplying each education ratio group with each of the residence/age groups. There was no design weight for the cellphone sampling because all RDD calls were random.

*Table 3-9. AAPOR Survey Outcome Rates*

| Response Rates | |
|---|---|
| Response rate 1: $\frac{I}{I+P+R+NC+O+UH+UO}$ | *Minimum response rate.*<br>All individuals who complete more than 80% of survey /<br>All attempted calls |
| Response rate 2: $\frac{I+P}{I+P+R+NC+O+UH+UO}$ | All individuals who complete more than 50% of survey /<br>All attempted calls |
| Response rate 3: $\frac{I}{I+P+R+NC+O+0.2*(UH+UO)}$ | All individuals who complete more than 80% of survey /<br>All attempted calls minus 80% of calls with unknown eligibility |
| Response rate 4: $\frac{I+P}{I+P+R+NC+O+0.2*(UH+UO)}$ | All individuals who complete more than 50% of survey /<br>All attempted calls minus 80% of calls with unknown eligibility |
| **Cooperation Rates** | |
| Cooperation rate 1 : $\frac{I}{I+P+R+O}$ | All individuals who complete more than 80% of survey /<br>Eligible individuals who were ever contacted |
| Cooperation rate 2: $\frac{I+P}{I+P+R+O}$ | All individuals who complete more than 50% of survey /<br>Eligible individuals who were ever contacted |
| **Refusal Rate** | |
| Refusal rate 3: $\frac{R}{I+P+R+NC+O}$ | All individuals who refused to complete the survey /<br>All attempted calls |
| **Contact Rate** | |
| Contact rate 1: $\frac{I+P+R+O}{I+P+R+NC+O+UH+UO}$ | All phone numbers that answered the call /<br>All phone numbers |
| Contact rate 2: $\frac{I+P+R+O}{I+P+R+NC+O+0.2*(UH+UO)}$ | All phone numbers that answered the call /<br>All attempted calls minus 80% of calls with unknown eligibility |

*Descriptive Analyses*

To examine distributions of the aforementioned independent variables, all of which are categorical except age, we conducted univariate analysis, looking at patterns of response in the FTF sample and in the RDD sample. In the FTF sample, we examined distributions among all women (n=3,659), who represent the target population and among a selected sample of cell phone owners (n=2,027), who represent the sample frame of the RDD survey. The R5 data were adjusted for sampling weights, which address disproportionate two-stage cluster sampling and non-response rates. [91] Turning to the RDD data, we included women who were partial (50-80% of questions answered) or complete respondents (more than 80% of questions answered) (N=2,379) in the RDD sample and examined the distribution of women's sociodemographic characteristics using unweighted and weighted RDD data. Also, we compared the reporting of modern contraceptive use between four groups: FTF – full sample; FTF – phone owners; RDD unweighted; RDD weighted. We also constructed a five-category indicator of method mix, distinguishing the following contraceptive users: implant, IUD, injectables, pills, and condoms.

*Equivalence Test*

To examine the equivalence of modern contraceptive use prevalence estimates in the weighted RDD and FTF samples, we set the equivalence margin $\delta$ to +/- 4%. The null hypothesis assumed a difference of more than 4% between the two estimates of modern contraceptive use. Rejection of the null hypothesis would lead to the conclusion that weighted RDD and FTF estimates of modern contraceptive use were equivalent within a margin of 4%. We report the 90% confidence interval for the difference in point estimates, which simulates performing two one-sided tests. We also report a p-value from an adjusted Wald test. [111]

*Multivariable logistic regression*

For the logistic regressions, we combined the RDD and FTF data to compare modern contraceptive use by survey mode, adjusted for confounding covariates. We first assessed bi-variate relationships between each co-variate and modern contraceptive use. We then conducted multivariable logistic regression to assess the adjusted odds of modern contraceptive use by survey mode, adjusting for covariates. We also compared the RDD and FTF phone owner sample using multivariable logistic regression. We checked for co-linearity and found no variance inflation factor of more than two. [98] Analysis was performed using weighted RDD and FTF data. We conducted analyses in Stata version 15 (StataCorp 2017) and determined statistical significance using an alpha of 0.05.

# 4. Differences in sociodemographic characteristics of cell phone owners and non-owners and the implications of cell phone ownership on estimates of modern contraceptive use among a national sample of women in Burkina Faso

## *4.1 Abstract*

**Introduction:** In 2017, there were 176 million unique mobile subscribers in the 15 countries that comprise the Economic Community of West African States (ECOWAS). This translates to a 47% cell phone penetration rate, up from 28% at the beginning of the decade. [112] This dramatic increase in cell phone ownership presents the opportunity to collect public health data by calling respondents on their cell phone. However, because cell phone ownership varies by socioeconomic, demographic and geographical variables, there is a risk of frame error: selecting a sample population unreflective of the target population, leading to biased estimates. To assess potential frame error and its implications for modern contraceptive use estimates in Burkina Faso, we described the sociodemographic characteristics associated with cell phone ownership among women in Burkina Faso and evaluated how modern contraceptive use rates varied by cell phone ownership.

**Study Design:** We analyzed data from a 2016 Performance Monitoring and Accountability 2020 (PMA2020) cross-sectional survey, comprising a nationally probability sample of 3,215 women aged 15 to 49 in Burkina Faso.

**Results:** Less than half of women (47%) reported cell phone ownership. The odds of reporting cell phone ownership among women with secondary education or higher was four times the odds

of women with no formal education (OR = 4.3, 95% confidence interval 2.9 – 6.1). Overall, 22% of women reported current modern contraceptive use. Women who owned a cell phone were more likely to report modern contraceptive use than those who did not (29% versus 16%, p-value <0.001). Method mix was substantially more diverse among those who owned cell phones, compared to their counterparts.

**Conclusions:** This study examined the risk of frame error by identifying women that can and cannot be reached via cell phone in Burkina Faso based on cellphone ownership status. Cell phone owners reported higher modern contraceptive use, which suggests that cell phone surveys are likely to overestimate national contraceptive prevalence rates – a key indicator for tracking family planning programs.

## 4.2 Introduction

Cell phone surveys are a nascent but growing field in Sub-Saharan Africa (SSA). Over the past eight years, researchers have become increasingly interested in capitalizing on the exponential growth of cell phone ownership in SSA to collect data. [12, 46] Urbanization, expanded cell phone network coverage and the low cost of purchasing a cell phone have all contributed to increased cell phone ownership throughout the continent, but cell phone ownership is biased towards men, the educated and urban dwellers. [35] In 2017 there were 176 million unique mobile subscribers in the 15 member states that comprise the Economic Community of West African States (ECOWAS). The mobile penetration rate, which is the percent of unique users within a specific population, [38] is 47% in the ECOWAS, up from 28% at the beginning of the decade. [112] From 2018 to 2025, 72 million new mobile subscribers are expected in ECOWAS countries, driven by the youth population purchasing cell phones for the first time. [112]

Phone surveys have an established history in high income countries, and in particular the United States and Europe. Good data quality, low cost and rapid data collection made landline random digit dial (RDD) telephone surveys the preferred mode of data collection in the United States from the 1980s until the early 2000s. [74] However, the increase in both cell phone ownership and internet connection combined with increasing distrust of phone surveys have caused landline RDD response rates to plummet to single digits in the US. [74, 75]

Cell phone surveys have a shorter history in low and middle income countries (LMIC). Ten computer assisted telephone interview (CATI) studies were identified in a 2017 systematic review of population-based cell phone surveys in LMIC. [46] Response rates varied from 99% in Mali [113] and 98% in Tanzania, [114] to 35% in Brazil. [9]

An area of concern among survey researchers is whether to use a cell phone or landline frame, or to create a dual frame, consisting of both landline and cell phone numbers. [74] In survey research, the frame is the list of units from which to select the study sample. A sample frame should include every unit of the target population. [115] Units can be people, businesses or households. [116] Each unit in the frame should have a nonzero probability of being selected into the sample. [116] Having a complete, correct frame is important for multiple reasons. In probability surveys, the frame will affect sampling error and weighting adjustments. [74] Although demographers often use complex sampling methods (i.e., not simple random sampling) that preclude proportional representation of a population or equal representation of all units, they can produce representative estimates by adjusting their sample to the underlying frame.

If there is a difference between the target population and the frame population, the result is frame error. Frame error is a type of non-sampling error and arises in the process of "constructing, maintaining and using the sample frame(s) for selecting the survey sample." [10] There are three types of frame error: population specification error, selection error, and coverage error. [117] The first, population specification error, occurs when the sphere of the population is incorrectly defined. For example, the border of a state is incorrectly defined, either including units that should not be in the frame, or omitting units that should be included in the frame. The second type of frame error is selection error, which occurs when sample selection procedures are incomplete (under-selecting) or falsely identifying (over-selecting) those in the sample. The third type of error, the frame error this analysis is concerned with, is coverage error, which occurs when the source sample does not accurately reflect the population. [115, 117] Coverage error can be caused by over-coverage, which occurs when participants in the source sample are ineligible but included in the frame, or when participants are included more than once in the frame. Coverage error can also be caused by under-coverage, which occurs when members of the target population are not included in the frame. Under-coverage is considered a more serious problem than over-coverage primarily because under-coverage is more difficult to identify than over-coverage. For example, over-coverage can occur if individuals own several cell phones. This multiplicity error can be corrected by asking respondents about the number of cell phone numbers he or she has. [49] On the other hand, identifying missed units is costly and difficult and often requires external sources. [116]

In the context of cell phone surveys in SSA, under-coverage is a concern because only 50% of the population owns a cell phone. [54] In other words, if an individual in the target population does not own a cell phone and the survey exclusively samples via phones, individuals without cell phones are erroneously excluded from the sample frame (under-coverage), which results in frame error. [115] Under-coverage error will lead to bias (systematic error) in outcome estimates if the cause of coverage error is associated with the outcome of interest. Both descriptive and analytic statistics are affected by frame bias. [49]

When considering the use of cell phone surveys to track sexual and reproductive health indicators, and in particular family planning metrics among women in SSA, it is important to identify the factors associated with cell ownership among the target population, as these factors may also be related to contraceptive use. While a number of population studies track household cell phone ownership, such as the Demographic and Health Surveys (DHS) [118] or Afrobarometer surveys, [119] few examine individual cell phone ownership among women in SSA. The scant literature that does exist about individual cell phone ownership reveals a substantial difference by gender: cell phone ownership data from Niger reported a 45 percent point gender gap. [96] Two studies from Kenya found that cell phone ownership is not uniform across the population. [120, 121] Poor rural women were the least likely to own cell phones, and in rural areas, phone sharing practices were more common than in urban areas. [120] In Rwanda, cell phone ownership was highest among males, the educated and the urban population. [122] Thirty percent of the population in the Rwanda study was estimated to share their cell phones, with women being more likely than men to share a phone, but unlike the pattern in Kenya, sharing was uniform across geographic areas.

While examination of the sociodemographic characteristics associated with female cell phone ownership is rare, the study of cell phone ownership implications for reproductive health estimates is even less common. One study in Nigeria assessed the relationship between maternal health service utilization, including contraceptive use and cell phone ownership, but included only women who had children under two years of age in five states in the country. The results showed that women who did not have phone access had half the odds of reporting modern contraceptive use compared to women that had phone access. [123]

According to a nationally representative survey, Performance Monitoring and Accountability 2020 (PMA2020), an estimated 47% of women in Burkina Faso owned a cell phone at the end of 2016. [124] How factors associated with cell phone ownership map with factors associated with modern contraceptive use, including age, education, household wealth and area of residence [125] has yet to be investigated. The objective of this study is to identify differences in the sociodemographic and reproductive health characteristics of women who own and who do not own a cell phone in Burkina Faso. We evaluate the implications of cell phone ownership on estimates of modern contraceptive use by comparing the prevalence of modern contraceptive use among cell phone and non-cell phone owners. The study results will help to understand the risk of frame error in a cell phone survey in Burkina Faso and its implication on modern contraceptive prevalence rates estimates.

## *4.3 Methods*

### ***Data source & Study population***

We analyzed data from the 4th round (R4) PMA2020 survey in Burkina Faso. PMA2020 is a nationally representative survey, primarily designed to track key family planning indicators under the Family Planning 2020 initiative. [8] Since 2014, PMA2020 has conducted five rounds of data collection in Burkina Faso. Ethical approval for human subject research was obtained from the Johns Hopkins Bloomberg School of Public Health's Institutional Review Board as well as the *Comité d'éthique pour la recherche en sante* in Ouagadougou, Burkina Faso. Datasets are publicly available. [89, 126]

Each PMA2020 survey round follows the same cross-sectional design in the form of a two-stage stratified cluster design. This design starts with a selection of geographic clusters based on probability proportional to size in each of the urban and rural strata, and is followed by a random selection of households within each sample cluster. Detailed sampling methods and procedures are in the methods section of this dissertation.

The PMA2020 surveys use a network of female resident interviewers to collect data through Face to Face (FTF) interviews with the head of the household and with female household members (*de jure)* and women who slept in the selected households the night before data collection (*de facto).* The household survey lists household members, durable assets and livestock, and documents house structure, water, sanitation, and hygiene conditions. [89] Interviewers record responses directly on cell phones and upload the data on a protected cloud server when cellular network is available. The female survey records sociodemographic characteristics and reproductive health measures, including current, recent and ever use of contraception, and current or previous method used.

The fourth round of the survey in Burkina Faso was conducted between November 2016 and January 2017. A total of 2,751 households and 3,215 eligible women aged 15 to 49 years were interviewed. The response rates were 97.9% for the household interview and 95.4% for the female interview.

***Measures***

The key dependent variable in this analysis was cell phone ownership. During the female survey, interviewers asked women “How many phone numbers do you have?” Women were considered cell phone owners if they reported having one or more phone numbers.

We chose independent variables based on the literature that shows education and wealth are consistently correlated with cell phone ownership [94, 95] and urban-rural ownership differences are profound among women in LMIC. [96] All considered, independent variables included women’s sociodemographic characteristics such as current union status (in union – i.e., currently married or living with a partner vs. not in union), residential area (urban vs. rural), highest school ever attended (none, primary, or secondary and higher), household wealth (asset score quintiles: lowest quintile vs. all other quintiles; highest quintile vs. all other quintiles), and having electricity (yes vs. no). Age was categorized into seven groups (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49) when comparing phone owners and non-owners in descriptive and regression analysis. Household wealth in PMA2020 surveys are based on a summary measure of household assets, based on Principal Component Analysis (PCA) method, which is also used in the DHS surveys. [97] Although electricity was used to calculate the household wealth index in the DHS, in this analysis electricity was also a separate co-variate because cell phone use requires access to electricity.

When assessing for frame bias, the dependent variable of interest was use of modern contraceptive method(s). Modern contraceptive methods, as defined by the World Health Organization, include pills, implants, injectables, intrauterine devices, condoms, female and male sterilization, lactational amenorrhea method (LAM), emergency contraception, and standard days method. [92] Using PMA2020 questions on current use of contraception, including type of method used, women were categorized into two groups: users of modern method and non-users of modern methods. Long-acting reversible contraceptive methods include implants and IUDs, methods that provide protection years at a time but can be reversed. [93] We also constructed a five-category indicator of method mix, distinguishing the following contraceptive users: implant, IUD, injectables, pills, and condoms.

### *4.4 Analysis*

We examined differences in sociodemographic composition of cell phones owners and non-cell phone owners among all women in the sample using chi-square tests. We then estimated the odds of cell phone ownership by age, education, wealth and area of residence using bivariate followed by multivariable logistic regression models. Parity and marital status variables were not significantly associated with phone ownership in bi-variate analyses and were not included in multivariable regression models.

After establishing the characteristics associated with cell phone ownership, we explored frame bias by comparing the prevalence of modern contraceptive use among cell phone owners and non-owners. In addition, we compared the distribution of modern method mix among cell phone owners and non-owners, using a chi-square test. [99] All analyses were adjusted for sampling weights and complex survey design, which account for two-stage cluster sampling and non-response rates. [91]

## *4.5 Results*

Study sample characteristics are presented in Table 4-1. The 3,215 women interviewed in PMA2020 R4 survey were on average 28.6 years old, 70% were married, and 79% had children. Three-quarters of the sample lived in rural areas and only a third of women had ever attended school.

Less than half (47%) of respondents reported cell phone ownership. Among cell phone owners, 83% reported only one phone number, 14% had two numbers and the remaining 3% had three or more phone numbers. The mean cell phone numbers among cell phone owners was 1.2.

Comparison of sociodemographic characteristics of cell phone owners and non-owners show no differences in age composition, marital status or parity (Table 4-2). Conversely cell phone owners and non-owners differed according to education, area of residence and language used to complete the survey. Specifically, 28% of cell phone owners had secondary education versus 11% of non-owners. A greater percentage of cell phone owners resided in urban areas compared to non-owners (37% versus 13%). More cell phone owners completed the survey in French (18%) compared to non-owners (3%), but fewer cell phone owners completed the survey in an 'other' language (10%) compared to non-owners (25%).

Both household characteristics were significantly different between cell phone owners and non-owners (p-values <0.05). More cell phone owners were in the highest wealth quintile (35%

compared to 10% among non-owners). Half of cell phone owners had electricity, whereas only 30% of non-owners had electricity.

The positive association between cell phone ownership and education remained significant in the multivariable analysis (Table 4-3). Compared to women with no education, who represent 65% of the sample, the odds of cell phone ownership were twice as high among women with primary education (Adjusted Odds Ratio (aOR): 1.9; 95% Confidence Interval (CI): 1.4 – 2.6), and over four times higher for women with secondary education (aOR: 4.3; 95% CI: 2.9 – 6.1), after adjusting for other sociodemographic characteristics. All age groups were more likely to report cell phone ownership compared to adolescents aged 15-19. The association of wealth and cell phone ownership was attenuated in the multivariable logistic regression compared to the bivariate results, but women in the highest wealth quintile still had almost twice the odds of owing a phone compared to women in the four lower wealth quintiles (aOR 1.8, 95% CI: 1.2 – 2.6).

We now describe the differences in  modern contraceptive use rates between cell phone owners and non-owners. Overall, 22% of women reported current use of a modern method in PMA2020 R4 (Table 4-4). We found a 13 percentage point difference between the prevalence of modern contraceptive use by cell phone ownership: 29% among cell phone owners and 16% among non-owners, p-value <0.001). This suggests the potential risk of over-estimating national modern contraceptive use prevalence in Burkina Faso using cell phone surveys due to substantial frame bias.

Not only was the percent of women using a modern method of contraception different among cell phone owners and non-owners, but the type of method used also differed (p-value <0.001). Overall, the order of most to least frequently used methods was the same among cell phone owners and non-owners, but the modern method mix was more diversified among cell phone owners compared to non-owners. Two methods – implants and injectables – accounted for 88% of the method mix among non-cell phone owners, while they accounted for 70% of modern method mix among cell phone owners (Figure 4-1).

## 4.6 Discussion

Phone surveys are an effective approach to collecting data when enough units (households or people) in the population have a phone, or when the difference between phone owners and non-owners is not related to the study outcome. [116] This analysis shows cell phone ownership among women of reproductive age in Burkina Faso varies substantially, especially by area of residence, education, and wealth, resulting in frame error due to under-coverage of groups who may not be accessed in cell phone surveys. This under-coverage is likely to bias national estimates of modern contraceptive use as the prevalence of modern contraceptive use is 13% point higher among cell phone owners than among non-owners.

Cell phone frame error stemming from greater phone ownership among the urban, educated and wealthiest has been reported in previous surveys [35, 123] and are confirmed in our study among a representative population of women of reproductive age in Burkina. We further contribute to the literature by showing that cell phone ownership is also associated with modern contraceptive use, suggesting substantial frame bias in the production of national family planning estimates if we were to track these indicators using phone surveys.

Not only is modern contraceptive use different by cell phone ownership, the distribution of methods among modern contraceptive users varies by cell phone ownership. Therefore, not only would our point estimate of modern contraceptive use be biased, the methods used by cell phone owners would not reflect target population use. Moreover, cell phone owners were less likely to use a long-acting contraceptive method (IUD or implant) compared to non-cell phone owners, so method switching may be more common among cell phone owners.

While rapid expansion of cell phone coverage may close the sociodemographic gap between owners and non-owners and reduce frame error, other strategies could be considered to address under-coverage frame bias until phone coverage is more common. [115] The most feasible option is to use a linking procedure, also known as the half-open interval. [117] The approach is implemented by establishing a rule that links persons missing from the frame (non-cell phone owners) to a person who is included in the frame (cell phone owners), with equal inclusion probability to their linked counterpart. The linking technique could be possible by asking a selected cell phone owner to pass the phone to a female household member who does not own a phone. However, the half-open interval method has traditionally been used for FTF surveys, in a US setting, and was deemed only useful if there was severe under-coverage. [127] Thus the effectiveness of this approach is unknown for phone surveys in LMIC. Another option would be to complement the phone frame with another frame that includes non-phone owners. However, there would be potential for frame overlap. [115] Using multiple frames, such as including cell phones and landlines has been a popular approach in Western countries where landline phones preceded cell phones. [128] However, in Burkina Faso, much like the rest of SSA, less than five percent of homes have a landline. [129] The sole option for a second frame in SSA would be a frame comprised of people available for a FTF survey. A FTF survey frame mitigates the main advantage of cell phone surveys, which is reduced cost. Furthermore, the population hardest to reach with a cell phone frame, rural uneducated women, are among the most challenging and expensive to reach with a FTF survey.

To our knowledge, this is the first study to employ nationally representative data to identify the groups of women who may be excluded from cell phone surveys in West Africa and to assess the implications of cell phone ownership on reproductive health estimates.

Limitations of this study include measurement error regarding women's cell phone ownership. Women may have reported a family member's phone number rather than a personal phone number, which could cause overestimation of female phone ownership. This potential mis-report has implications for reaching a woman. Her husband may be a gatekeeper to the phone, or if she shares the phone, her access may be limited. More importantly, cell phone ownership does not necessarily mean a woman is easily reached via cell phone. She may have a cell phone number but does not have cell phone network except on market days, or insufficient electricity to charge the cell phone regularly. Also, with the rapid increase in phone ownership, our findings may have limited relevance in the near future in Burkina Faso.

## 4.7 Conclusion

This analysis identified women that can and cannot be reached via cell phone in Burkina Faso and quantifies the potential frame bias when estimating modern contraceptive use from a cell phone survey. While these results suggest caution in using cell phone surveys for tracking family planning metrics due to substantial frame bias, we suggest monitoring the rise in female cell phone ownership in population based surveys to determine when frame bias is no longer a significant issue. In the meantime, we also recommend that our results be used to define quota strategies to improve on cell phone sample representation.

## 4.8 Tables for Chapter 4

*Table 4-1. Characteristic of study sample*

| | Total N = 3,215 |
|---|---|
| | N (%) |
| **Age** | |
| Mean | 28.6 years |
| 15-19 | 711 (22.1%) |
| 20-24 | 556 (17.3%) |
| 25-29 | 556 (17.3%) |
| 30-34 | 460 (14.3%) |
| 35-39 | 405 (12.6%) |
| 40-44 | 296 (9.2%) |
| 45-49 | 228 (7.1%) |
| | |
| **Residential area** | |
| Rural | 2,431 (75.6%) |
| Urban | 784 (24.4%) |
| | |
| **Marital status** | |
| Currently not in union | 981 (30.5%) |
| Currently in union | 2,234 (69.5%) |
| | |
| **Highest school attended** | |
| No education | 2,074 (64.5%) |
| Primary | 518 (16.1%) |
| Secondary or higher | 624 (19.4%) |
| | |
| **Parity** | |
| Never given birth | 682 (21.2%) |
| Ever given birth | 2,533 (78.8%) |
| | |
| **Language** | |
| Dioula | 469 (14.6%) |
| French | 334 (10.4%) |
| Fulfulde | 196 (6.1%) |
| Gourmantchema | 312 (9.7%) |
| Moore | 1,328 (41.3%) |
| Other | 575 (17.9%) |
| | |

| | |
|---|---|
| **HH Wealth (quintile)** | |
| Lowest | 691 (21.5%) |
| Lower | 614 (19.1%) |
| Middle | 666 (20.7%) |
| Higher | 559 (17.4%) |
| Highest | 685 (21.3%) |
| | |
| **HH Electricity** | |
| No | 1,286 (40.0%) |
| Yes | 1,929 (60.0%) |

**Note:** *% estimates are adjusted for sampling weight*

*Table 4-2. Characteristic of study sample by cellphone ownership*

| | Cell phone ownership (N = 3,215) | | p-value |
|---|---|---|---|
| | Not an Owner | Owner | |
| | N (%)* | N (%) * | |
| Total | 1,707 (53.1%) | 1,508 (46.9%) | |
| **Age** | | | |
| Mean | 28.2 years | 28.8 years | |
| 15-19 | 428 (25.1%) | 284 (18.8%) | 0.0059 |
| 20-24 | 273 (16.0%) | 284 (18.8%) | |
| 25-29 | 279 (15.8%) | 284(18.8%) | |
| 30-34 | 230 (13.5%) | 230 (15.2%) | |
| 35-39 | 215 (12.6%) | 190 (12.6%) | |
| 40-44 | 159 (9.3%) | 138 (9.1%) | |
| 45-49 | 131 (7.7%) | 98 (6.5%) | |
| | | | |
| **Residential area** | | | |
| Rural | 1,487 (87.1%) | 944 (62.6%) | <0.001 |
| Urban | 229 (12.9%) | 564 (37.4%) | |
| | | | |
| **Marital status** | | | |
| Currently not in union | 485 (28.4%) | 496 (32.9%) | 0.1679 |
| Currently in union | 1,222 (71.6%) | 1,012 (67.1%) | |
| | | | |
| **Highest school attended** | | | |
| No education | 1,280 (75.0%) | 807 (53.5%) | <0.001 |
| Primary | 241 (14.1%) | 276 (18.3%) | |
| Secondary or higher | 186 (10.9%) | 427 (28.3%) | |
| | | | |
| **Parity** | | | |
| Never given birth | 335 (19.6%) | 345 (22.9%) | 0.1647 |
| Ever given birth | 1,372 (80.4%) | 1,163 (77.1%) | |
| | | | |
| **Language** | | | |
| Dioula | 251 (14.7%) | 219 (14.5%) | <0.001 |
| French | 58 (3.4%) | 270 (17.9%) | |
| Fulfulde | 169 (9.9%) | 32 (2.1%) | |
| Gourmantchema | 241 (14.1%) | 75 (5.0%) | |
| Moore | 551 (32.3%) | 766 (50.8%) | |
| Other | 439 (25.7%) | 146 (9.7%) | |
| | | | |

| | | | |
|---|---|---|---|
| **HH Wealth (quintile)** | | | |
| Lowest | 483 (28.3%) | 210 (13.9%) | <0.001 |
| Lower | 387 (22.7%) | 228 (15.1%) | |
| Middle | 399 (23.4%) | 265 (17.6%) | |
| Higher | 273 (16.0%) | 285 (18.9%) | |
| Highest | 166 (9.7%) | 520 (34.5%) | |
| | | | |
| **HH Electricity** | | | |
| No | 1,183 (69.3%) | 742 (49.2%) | <0.001 |
| Yes | 524 (30.7%) | 766 (50.8%) | |

***Note:** *% estimates are adjusted for sampling weight*

** P-value for Rao and Scott's chi-square test for differential distribution of cell phone ownership by background characteristics*

*Table 4-3. Odds of cell phone ownership by background characteristics: bivariate and multivariable logistic regression analyses*

| | **Bivariate** | **Multivariable** |
|---|---|---|
| | Crude OR (95% CI) | Adjusted OR (95% CI) |
| **Age group** | | |
| 15-19 *(reference)* | | |
| 20-24 | **3.2 (2.2 – 4.5)** | **2.4 (1.8 – 3.4)** |
| 25-29 | **3.6 (2.5 – 5.1)** | **2.9 (2.1 – 4.0)** |
| 30-34 | **3.9 (2.5 – 6.0)** | **3.0 (2.2 – 4.2)** |
| 35-39 | **3.2 (2.2 – 4.8)** | **2.8 (1.8 – 4.3)** |
| 40-44 | **2.7 (1.7 – 4.3)** | **3.0 (2.0 – 4.4)** |
| 45-49 | 1.2 (0.7 – 2.0) | **2.7 (1.9 – 4.1)** |
| | | |
| **Residential area** | | |
| Rural *(reference)* | | |
| Urban | **4.5 (3.0 – 6.7)** | **1.9 (1.3– 2.9)** |
| | | |
| **Marital Status** | | |
| Currently not in union *(reference)* | | |
| Currently married | 0.76 (0.6 – 1.0) | |
| | | |
| **Highest school attended** | | |
| No education *(reference)* | | |
| Primary | **2.0 (1.5 – 2.7)** | **1.9 (1.4 – 2.6)** |
| Secondary or more | **4.3 (3.0 – 6.4)** | **4.3 (2.9 – 6.1)** |
| | | |
| **Parity** | | |
| Never given birth *(reference)* | | |
| Ever given birth | 1.0 (0.97 – 1.0) | |
| | | |
| **Household wealth quintile** | | |
| Lowest *(reference other 4 quintiles)* | **0.6 (0.4 – 0.8)** | **0.6 (0.4 – 0.9)** |
| Highest *(reference other 4 quintiles)* | **4.5 (3.2 – 6.4)** | **1.8 (1.2 – 2.6)** |
| | | |
| **Electricity** | | |
| No household electricity *(reference)* | | |
| Household electricity | **2.5 (1.7 – 3.7)** | 1.3 (0.8 – 1.9) |

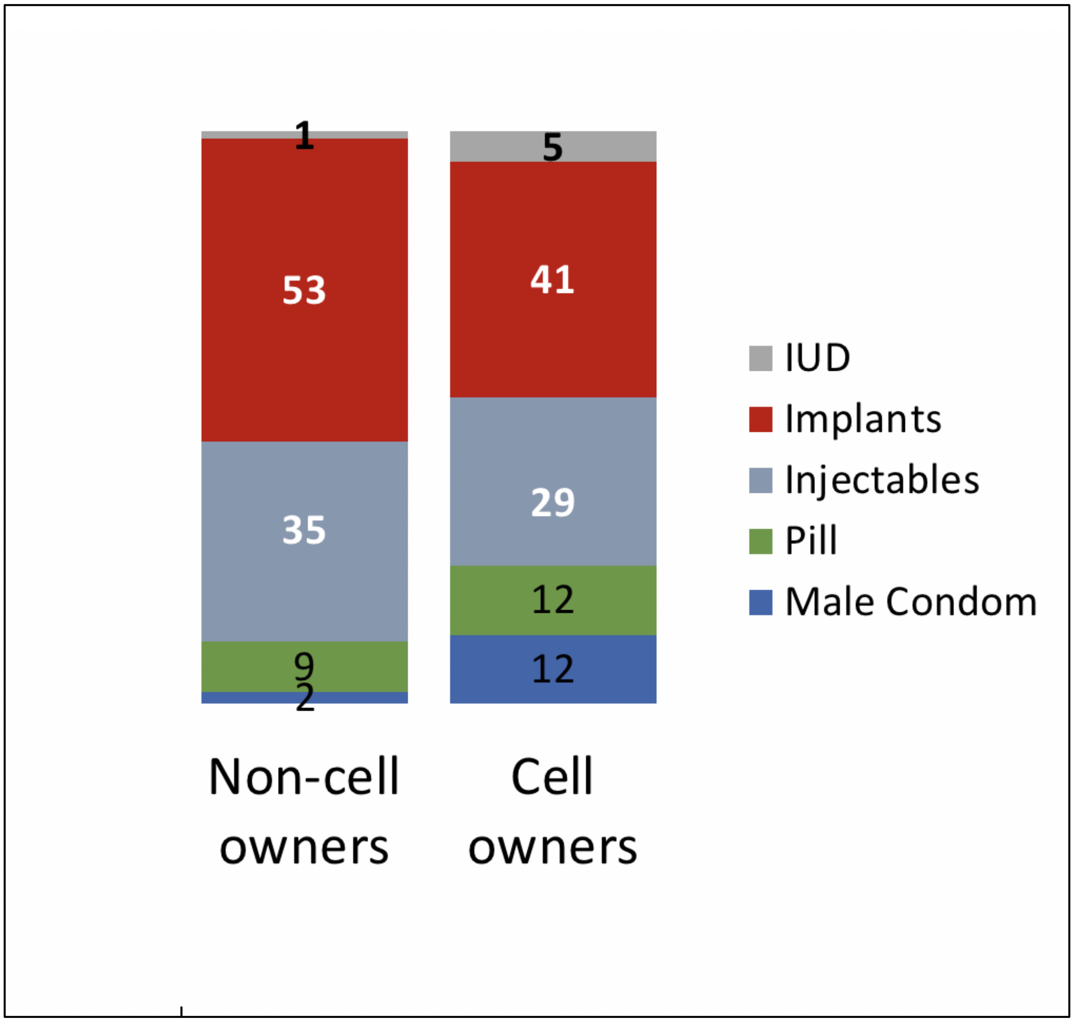
**Note:** *Analyses adjusted for sampling weight. Only statistically significant variables in bi-variate analysis were included in multivariable analysis.*

*Table 4-4. Prevalence of modern contraceptive use by female cell phone ownership*

| | All women | Cell phone owners | Cell phone non owners | p-value |
|---|---|---|---|---|
| All | 21.9 (19.3 – 24.7) | 29.4 (26.0 – 32.9) | 15.9 (12.9 – 19.3) | <0.001 |

**Note:** *% estimates are adjusted for sampling weight; 95% CI is adjusted for higher intraclass correlation due to clustering of subjects in enumeration areas (ICC >0/deff>1).*

*Figure 4-1. Distribution of method mix among women currently using modern contraception by cell phone ownership status*




*The most effective method currently used, if multiple methods were reported
**Note:** *% estimates are adjusted for sampling weight*

# 5. Sociodemographic characteristics related to cell phone survey completion among female phone owners in Burkina Faso and the implications of non-response on contraceptive use estimate

## *5.1 Abstract*

**Introduction:** The exponential growth of cell phone ownership in Low and Middle Income Countries (LMIC) has generated significant interest in using cell phones for conducting population surveys remotely in these geographies, using either Computer Assisted Telephone Interviews (CATI) or interactive voice response (IVR) methods. However, the use of cell phones for data collection raises concerns of survey quality and the representativeness of the sample. The objectives of this study are to compare the sociodemographic characteristics of cell phone owners between women who consent to participate and those who do not, and between women who complete a phone survey to those who don't. We also assess the implications of cell phone survey completion on modern contraceptive use estimates in Burkina Faso. We assess these questions using two phone survey modes: CATI and Hybrid IVR. Hybrid IVR is an IVR survey that opens with a human operator who acquaints and consents the participant before the participants begins the IVR survey.

**Study Design:** The study draws on a nationally representative sample of women 15-49 years old who own a cell phone in Burkina Faso, identified in Round 4 of the Burkina Faso Performance Monitoring and Accountability 2020 (PMA2020) 2016 Survey. Female cell phone owners who consented to participate in a follow-up phone survey were randomly allocated to participate in a CATI or a Hybrid IVR follow-up phone survey, 11 months following PMA2020 baseline interviews. We defined phone survey completion as responding to 50% of questions or more in the phone follow-up survey and investigated sociodemographic characteristic associated with cell phone completion using multivariable logistic regression models, stratifying analysis by survey mode (CATI and Hybrid IVR). We compared the proportion of modern contraceptive users among phone completers and non-completers using chi-square tests.

**Results:** Among the 1,766 women called for the follow-up survey, 20% completed the Hybrid IVR survey and 50% completed the CATI survey. Women in urban communities were more likely to complete the CATI survey compared to their rural counterparts (Odds Ratio (OR) 1.7, 95% Confidence Interval (CI): 1.1 – 2.5) as were women with a secondary education or more (OR 1.4, 95% CI: 1.1 – 2.0). Sample distortion of Hybrid IVR completers was even more pronounced, resulting in higher representation of women with secondary education in the Hybrid IVR completion sample than in the CATI completion sample (OR: 1.7, 95% confidence interval: 1.1 – 2.6). There was no difference in modern contraceptive use rates between female cell phone owners who completed the phone survey and those who did not, regardless of survey mode (CATI or Hybrid IVR).

**Conclusions***:* In Burkina Faso, the response rate was higher for the CATI survey, compared to Hybrid IVR. Although both modes introduced non-response errors, the reporting of contraceptive use was similar between the survey completers and non-completers suggesting no non-response

bias. We conclude CATI is the preferred method of conducting cell phone surveys in low literacy countries such as Burkina Faso.

## 5.2 Introduction

The projection that over the next 75 years the majority of francophone West African countries will quadruple in population has engendered interest in increasing contraceptive use. [2] With almost 30 percent of married women of reproductive age in francophone West Africa not using contraception but reporting the desire to delay or avoid a pregnancy, there is a need for contraceptives. [24] To track progress towards increasing contraceptive use, there is a need for rapid, cost-effective data collection approaches to supplement the five-year periodicity of the Demographic and Health Survey (DHS), which has been the main source of reproductive health data in low- and middle-income countries (LMIC) for the past 30 years. Since 2013, Performance Monitoring and Accountability 2020 (PMA2020) surveys have supplemented the DHS by providing timely, frequent data for decision makers by conducting face-to-face (FTF) surveys every six months to one year in eleven priority countries. [89] Although PMA2020 has been a successful compliment to the DHS, PMA2020 only operates in four francophone West African countries: Burkina Faso, Democratic Republic of Congo, Ivory Coast and Niger.

The exponential growth of cell phone ownership and claims of low cost data collection have precipitated interest in cell phone surveys in SSA. [35] In Burkina Faso, the site for this study, the 2016 PMA2020 survey found that 86% of households owned a cell phone and 47% of women reported personal cell phone ownership. [124] Although less than half of women owned a phone, a woman may have access to a phone through a household member. [96]

A transition from FTF to cell phone data collection raises issues of survey quality and comparability of indicators. In order to operationalize survey quality, survey researchers devised the concept of Total Survey Error (TSE). [10, 30] Survey mode impacts all five components of TSE: frame, non-response, specification, measurement, and data processing errors.

Whereas the previous chapter addressed frame error, this chapter examines non-response error. Non-response error occurs when people who are sampled but not interviewed differ in a non-negligible way from those who are interviewed. [10] In telephone surveys, non-response generally stems from three causes: 1) failure to contact sampled respondents; 2) refusal to participate; 3) ability or language constraints. [130]

Traditionally, response rates have been used to gauge non-response error and survey quality, but a high non-response rate does not necessarily translate to biased estimates if the propensity to respond is not correlated with the outcome of interest. [50, 51] Non-response bias occurs when people who are sampled but who don't complete the survey systematically differ from those who successfully complete regarding the outcome of interest. [10, 51] Non-response bias is concerning for both point and variance estimations. [131]

Non-response error and bias in cell phone surveys have been minimally explored in SSA due to the recency of the approach in the region, but there is a well-established body of literature on phone survey research in high income countries. [56, 57, 73, 132-134] The seminal studies in the

US exploring phone non-response took place in the 1980s and 1990s, using administrative data to construct the profile of non-respondents. [51] Characteristics related to landline survey response in the US included education, age, minority status and civic engagement. [73] Fewer studies have examined non-response error associated with cell phone surveys in the US, as noted by both the 2010 [73] and 2017 [74] American Association for Public Opinion Research reports on cell phone surveys. [128] One study of non-response to a US cell phone survey found that people who reported high civic engagement and volunteerism were more likely to participate in the survey, introducing bias to estimates of behaviors, such as church attendance, contact with elected officials and attending campaign events. [75]

Non-response to phone surveys, including cell phones has also been extensively studied in Europe. A recent health survey of 15,635 people using a mobile and landline frame in France had much higher response rates than in the US: the response rate was 47.3% for landlines and 37.7% for cell phones. Regardless of mode, the sample was representative of the general population, with the exception of education: those who participated in the phone survey were more educated than the general population. [76] In Europe, compared to those with only a landline, the mobile phone-only population is younger, better educated, earns higher income and has more advanced technology competencies. [77]

There is a dearth of rigorously conducted research on cell phone surveys in LMIC and among the research that does exist, the representativeness of a sample is rarely considered. Rather, studies have mainly addressed feasibility and measurement error. [12] Studies that attempt to profile non-respondents often assess sample distortion by comparing phone survey respondents to a reference population based on census or DHS data or comparing early and late responders. [30, 71, 72] Two recent random digit dial (RDD) studies in SSA, one in Ghana using Interactive Voice Response (IVR) and one in Ivory Coast using computer assisted telephone interview (CATI), both found that completers of cell phone surveys were more likely to be young, urban, educated and male as compared to DHS population distributions. [71, 72] Although comparing the phone respondent to the DHS sample can provide some insight on non-responders, these studies cannot distinguish frame (cell phone ownership) from non-response error because the frame is unknown.

Three CATI studies, in Tanzania, [135] Peru and Honduras, [32] were designed to specifically assess non-response to cell phone surveys. Similar to the study presented in this chapter, these studies enrolled respondents during a FTF survey who were then followed-up via cell phone. The study in Dar es Salaam, Tanzania was a 33-round panel phone survey. Respondents were asked about a range of issues including health, education, food security and governance. The target population was people living in urban households in Tanzania. In each of the 550 selected households, one respondent was randomly selected for the study and 418 participants who had a phone were enrolled. Wealth was a significant predictor of participation, when comparing survey participants (people who completed at least one survey over the 33 rounds) and non-participants (people who did not answer any calls or did not have a cell phone) among all 550 participants. Area of residence (peri-urban versus urban) and a proxy for wealth, having a premium cell phone provider, were significant predictors of participation when restricting analysis to cell phone owners (n=418), suggesting these factors are related to non-response. [135] Another study about living conditions and poverty, conducted by the World Bank in

Honduras and Peru, enrolled a national random sample of households with baseline FTF surveys and gave phones to households that had none. Attrition during follow-up calls was highest among poorer, less educated, older and rural participants. In both countries women were slightly more likely than men to respond to the first wave of follow-up phone calls. [32]

All considered, issues of non-response are highly relevant when conducting phone surveys, but over the past four decades survey methodologists have amassed strong evidence, including a meta-analysis, that there is no consistent relationship between non-response rates and non-response bias in high income countries. [51, 136] Over the last 15 years in SSA, studies of non-reponse bias has primarily focused on HIV related behaviors. There is evidence of higher refusal of HIV testing when the respondent has been recently tested, but the impact was negligible (less than one percentage point) on HIV prevalence estimates in most countries studied. [137, 138] Cell phone non-response bias on other reproductive health indicators such as modern contraceptive use has yet to be tested among female populations in SSA.

To fill this gap in the literature, we conducted a study in Burkina Faso, a country with ambitious goals for increasing contraceptive use among women of reproductive age. The 2017- 2020 Burkina Faso National Family Planning Acceleration Plan set the national target to 32% contraceptive prevalence by 2020. [82] Recent trends indicate accelerated progress from 15% to 26% modern contraceptive prevalence rate among women in union between 2010 and 2018, but sustaining this two percent annual increase will be challenging. [1] Since 2014, PMA2020 has contributed towards tracking these annual estimates and evaluating the impact of specific FP programs at the national level. For example, PMA2020 data was used to identify a doubling of subcutaneous injectables over a six month period, prompting scale-up of injectable availability. [139] PMA2020 has completed five rounds of FTF data collection in Burkina Faso and is exploring new approaches to collect quality data at lower cost – particularly cell phones surveys. [12, 30, 72]

The paucity of remote data collection studies in SSA, and in particular studies using a known representative sampling frame to assess non-response bias is a notable knowledge gap. The objective of this research is to identify sociodemographic characteristics related to phone survey completion among a representative sample of female phone owners in Burkina Faso and to evaluate the implications of non-response on contraceptive use estimates at the national level in Burkina Faso. We examine these questions for two types of phone survey mode: CATI and Hybrid IVR.

## *5.3 Methods*

The study utilized a nationally representative sample of women of reproductive age who owned a cell phone, identified in Burkina Faso PMA2020 Round 4 (R4), who were randomized to receive a phone follow-up survey, either using CATI or Hybrid IVR modes of data collection. CATI involves a live interviewer who administers the full survey. Hybrid IVR is an IVR survey that starts with a human operator. The human operator confirms eligibility of the respondent, carries out the consent process and explains how an IVR survey works before transferring the respondent to the automated IVR survey. Hybrid IVR was preferred over IVR because of the low literacy of women in Burkina Faso (30% literacy). [87] It was assumed that having a live interviewer open the IVR survey would increase response rates and improve data quality. Phone follow-up occurred 11 months after the PMA2020 R4 survey. The phone survey was a shortened version of PMA2020's standard FTF questionnaire. The study received IRB approval from the Johns Hopkins Bloomberg School of Public Health and from the ethical committee in Burkina Faso, *Comité d'éthique pour la recherche en santé*.

***Study population***

Performance Monitoring and Accountability 2020 (PMA2020) R4 survey in Burkina Faso used a two-stage stratified cluster design, starting with a selection of 83 geographic enumeration areas using probability proportional to size sampling, stratified by rural/urban geographies, followed by a random selection of 35 of the approximately 200 households within each enumeration area. Within each household, all eligible women aged 15 to 49 years are invited to participate. A total of 3,215 female respondents were interviewed during PMA2020 R4, with a 95.4% response rate. Among women in the study sample, 1,839 owned a cell phone and 1,766 (96%) consented to be contacted again for the current study (Figure 5-1). The aforementioned 1,766 women were randomly assigned to two arms after stratification by survey language (five languages and one 'other' group): 882 were randomized to receive Hybrid IVR and 884 to receive CATI (as explained below).

***Study procedures***

As previously indicated, women were asked to respond to the survey with one of two modes: CATI or Hybrid IVR. Fifteen trained interviewers conducted the phone follow-up survey, operating from a call center in Ouagadougou seven days a week between the hours of 8 am and 8 pm, in

November 2017. Interviewers called all eligible women (consenting female cell phone owners from PMA2020 R4). Although participants had consented to be re-contacted during PMA2020 R4, all participants were re-consented orally during the phone interview. Women who didn't answer were attempted six times before being classified as non-contacts.

***Measures***

<u>Final Disposition Codes for each Eligible Woman</u>

Using the American Association for Public Opinion Research (AAPOR) call disposition classifications and considerations specific to our study, we defined eight categories (Table 5-1). all participants were eligible and were classified as respondents or non-respondents.

Table 5-1. *Definition of final disposition codes*

| AAPOR Categorization | Final Disposition Code | Explanation |
|---|---|---|
| **Non-Respondent** | | |
| NC (2.20) | Non-contact (did not pick up) | No phone calls were picked up over the 6 attempts |
| O (2.36) | Non-contact (someone picked up the phone call but interviewer never spoke with the woman) | A phone call was answered but the intended respondent was either unknown, or never available to speak to the interviewer |
| R (2.12) | Refusal pre-consent | The respondent refused to participate in the study before consent |
| R (2.111) | Refusal | The respondent refused to participate in the study at the time of consent |
| R (2.121) | IVR Break-off | The respondent consented to participate but was unable to push 1 on phone (IVR test) |
| R (2.12) | Break-off (consented but less than 50% completion) | The respondent consented to participate but answered less than 50% of the questions |
| **Respondent** | | |
| P (1.2) | Partial completer (answered 50-80% of questions) | The respondent consented to participate and answered 50% to 80% of the questions |
| I (1.1) | Completer (answered more than 80% of questions) | The respondent consented to participate and answered more than 80% of the questions |

***Dependent variables***

For the analysis assessing the sociodemographic factors related to phone survey response, we defined two outcome measures. Our first indicator "consenting to the phone survey", is not a traditional survey measure used by AAPOR, but has significance in this study because phone survey participants heard the same questions, regardless of mode (CATI or IVR), until after the consent questions. All non-contacts and refusals (before or at the time of consent) were classified as non-consenters while women who consented to participate were all classified as consenters. Consenters included women who completed little, some or all the questions (break-offs, partial completers and full completers). Our second measure, "completing phone survey," was defined as answering 50% or more of the relevant survey questions - partial and complete respondents.

To evaluate the implications of non-completion on estimates of contraceptive use (non-response bias), we defined modern contraceptive use as a dichotomous variable (yes vs. no). Women were categorized into those using any modern method vs. non-users of modern methods. Modern methods in Burkina Faso, as defined by the World Health Organization, include pills, implants, injectables, intrauterine device, condoms, female and male sterilization, lactational amenorrhea method (LAM), emergency contraception, and standard days method. [92]

**Independent Variables**

Our sociodemographic independent variables were chosen by reviewing the literature from LMIC about responsiveness to cell phone surveys [70, 72] and literature about modes of data collection in SSA. [62, 104-106] Sociodemographic information was collected during PMA2020 R4 and therefore available for all eligible women, regardless of whether or not the woman responded to the phone follow-up survey. The follow-up data was only used to assign final disposition codes and thus calculate survey outcome rates. Sociodemographic variables were age, which was categorized into seven groups based on descriptive analysis (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49), and residential area (urban vs. rural). Additional co-variates included current union status (in union – i.e., currently married or living with a partner vs. not in union), highest school ever attended (none, primary, or secondary and higher), household wealth (lowest, middle and highest tertiles), having electricity (yes vs. no), and parity (ever given birth vs. never given birth). When survey completion was the outcome, we used four age groups (15-19, 20-29, 30-39, 40-49) since a small number of women completed the Hybrid IVR survey.

### *5.4 Analysis*

To examine whether randomized groups (women assigned to CATI phone survey and women assigned to Hybrid IVR survey) were similar according to their background characteristics at baseline (PMA2020 R4 survey), we explored the distribution of selected demographic and socioeconomic variables by mode of data collection arm.

We used the final disposition codes to calculate follow-up phone survey outcome rates. AAPOR offers four essential outcome rates (see Table 5-2).

*Table 5-2.* AAPOR Survey Outcome Rates

| Response Rates | Explanation |
|---|---|
| Response rate 5: $\frac{I}{I+P+R+NC+O}$ | All individuals who complete more than 80% of survey / All eligible individuals |
| Response rate 6: $\frac{I+P}{I+P+R+NC+O}$ | All individuals who complete more than 50% of survey / All eligible individuals |
| **Cooperation Rates** | |
| Cooperation rate 1 : $\frac{I}{I+P+R+O}$ | All individuals who complete more than 80% of survey / Eligible individuals who were ever contacted |
| Cooperation rate 2: $\frac{I+P}{I+P+R+O}$ | All individuals who complete more than 50% of survey / Eligible individuals who were ever contacted |
| **Refusal Rate** | |
| Refusal rate 3: $\frac{R}{I+P+R+NC+O}$ | All individuals who refused to complete the survey / All eligible individuals |
| **Contact Rate** | |
| Contact rate 3: $\frac{I+P+R+O}{I+P+R+NC+O}$ | All phone numbers that answered the call / All phone numbers |

- AAPOR's ***response rate*** corresponds to the number of individuals who complete the phone interviews, fully or partially, over all dialed phone numbers. AAPOR distinguishes two subcategories of response rates, based on whether respondents completed the 80% (Response rate 5) or 50% of relevant survey questions (Response rate 6).
- AAPOR's ***cooperation rate*** is similar to response rate, but excludes non-contacts from the denominator. Two cooperation rates are used in this study: Cooperation rate 1 only includes complete interviews in the numerator, and Cooperation rate 2 also includes partial interviews in the numerator.

- AAPOR's ***refusal rate*** corresponds to the number of individuals who refuse to be interviewed, among all eligible individuals.
- Finally, AAPOR's ***contact rate*** corresponds to the number of calls in which one member of the unit was reached (i.e. someone at the target phone number picked up the phone); among all eligible individuals.

Turning to our analysis of factors associated with survey consent and survey completion, we compared the distribution of the aforementioned sociodemographic characteristics between consenters and non-consenters and between completers and non-completers, using chi-square tests. Analysis were stratified by mode of data collection (CATI and Hybrid IVR).

We then conducted multivariable logistic regression models to identify the independent factors associated with survey consent. We performed one model for each survey mode (Hybrid IVR and CATI). Next, we used multivariable logistic regression models to identify the independent factors associated with survey completion, and stratified the analysis by survey mode (Hybrid IVR and CATI). We also directly compared Hybrid IVR and CATI completers to evaluate if sample distortion was different between the two modes, given the lower response rate for Hybrid IVR.

The five multivariable logistic regressions generated odds ratios and 95% confidence intervals quantifying the relationship between consent or completion and women's sociodemographic characteristics. We did not include marital status nor parity in the multivariable models because these characteristics were not significantly related to consent or completion in the bivariate analyses. Electricity was also not included because of high correlation with wealth tertiles. We checked for collinearity among other covariates and found no variance inflation factor (VIF) of more than four. [98]

Finally, we compared the reporting of modern contraceptive use between survey completers and non-completers using chi-square tests, for both Hybrid IVR and CATI.

All analyses were conducted in Stata version 15 (StataCorp 2017). Survey weights were not used for analyses.

## *5.5 Results*

### *Characteristics of the sample*

The average age of the 1,766 eligible women (female cell phone owners who consented to be followed-up) enrolled in this study was 28.5 years (Table 5-3). While 42% of women had never attended school, 20% had attended primary school and 38% had secondary or higher education. 65% were married and 72% were parous. Two-thirds of women lived in rural areas and 62% had electricity in the household. Finally, 67% were in the highest wealth tertile, with 18% in the middle and 15% in the lowest tertile. After randomization, women in the Hybrid IVR and CATI groups were similar across all the sociodemographic characteristics examined in this study (Table 5-3).

Phone Survey Outcomes

***Call Disposition Codes***

The percent of eligible women who did not answer any of the six calls was the same among the two groups (26%) but among those in the Hybrid IVR arm, more women were unavailable or not found even when the phone call was answered, compared to the CATI arm (20% Hybrid IVR vs. 15% CATI, p-value = 0.005) (Table 5-4).

Overall, 832 (47%) women consented to participate in the phone interview when called: 43% of Hybrid IVR women and 53% of women in the CATI arm (p-value <0.001). Break-off, defined as consenting but answering less than 50% of the questions, was substantially higher for the Hybrid IVR arm (24%) than for the CATI arm (2%). Altogether, among all women randomized to the Hybrid IVR arm, 20% completed the survey (2% partial interviews, 18% complete interviews), while 51% of those randomized to the CATI arm completed the survey (1% partial interviews, 50% complete interviews).

***Survey Rates***

The four essential AAPOR survey outcome rates (response, cooperation, refusal and contact) based on disposition codes are presented in Table 5-5. The response rate, a measure of the number of partial or completed interviews over all eligible respondents, was twice as high in the CATI arm compared to the Hybrid IVR arm. Specifically, the CATI response rate was 50.3% compared to 19.8% for Hybrid IVR (response rate #6). The cooperation rate, defined as completed interviews over contacted respondents, for CATI (85%) dwarfs the cooperation rate for Hybrid IVR (36%) (cooperation rate 2). The refusal rate for Hybrid IVR (12%) was almost double the refusal rate of CATI (7%), although the difference was not statistically significant (p-value = 0.226). Finally, the contact rate, the proportion of all cases in which a responsible member of the unit was reached by the survey, was similar among the two modes (55% for Hybrid IVR and 59% for CATI).

Characteristics associated with phone survey consent and completion

Overall, 42.5% of women consented to the Hybrid IVR survey. This proportion was lower among women 15-19 years (27%) and higher among the oldest age group (55%) (Table 5-6). The percentage of consenters was higher among urban women compared to rural women (46% versus 35% p-value = 0.002), and likewise higher among wealthier women (45% in the highest wealth tertile and 44% in the middle wealth tertile versus 29% in the lowest wealth tertile, p-value = 0.001). A lower percentage of women who spoke an 'other' language or Fulfulde when completing PMA2020 R4 FTF survey consented (22% and 17%) compared to over 40% of women who spoke French, More or Dioua (p-value = 0.008).

Among all 882 women randomized to Hybrid IVR, only 19.8% completed the survey. Survey completion was not dependent on age (p-value = 0.479) but was lower among rural women compared to women living in urban areas (9% versus 25%, p-value = 0.002). Likewise, the proportion of women completing the Hybrid IVR survey was lower among women with no education (12% versus 29% for those with secondary education or higher, p-value <0.001). Twice as many women who had household electricity completed the survey compared to women with no electricity (24% versus 13%, p-value <0.001). Finally, fewer women who spoke Gourmantchema, Fulfulde or an 'other' language completed the survey, compared to French speakers or More speakers (p-value = 0.005).

Many of the patterns of consent and completion described for the Hybrid IVR arm were also found in the CATI arm, although overall consent and completion were higher in the CATI arm (Table 5-7). Few rural women consented to the CATI survey compared to women living in urban areas (38% versus 58%, p-value <0.001). Only a third of teenagers consented to the CATI survey, but at least 45% of all other age groups consented (p-value <0.001). CATI consent ranged from 36% to 57% according to wealth tertile (p-value <0.001) and also varied by household electricity status (55% compared to 44% without electricity, p-value = 0.003). A third of women who spoke Fulfulde consented to the CATI survey and 30% of women who spoke Gourmantchema, compared to over 50% of women who spoke Dioula, French and Moore (p-value =0.002).

Whereas a considerable number of Hybrid IVR consenters did not complete the IVR survey, this was not the case for the CATI arm, where only 24 women who consented to CATI did not complete the CATI survey. Therefore, the proportion of CATI completers was very similar to the proportion of CATI consenters among all women and among subgroups of women.

**Results of multivariable analysis comparing phone survey consenters and non-consenters**

Results from the multivariable logistic regression were similar to bivariate results. The odds of consenting to the Hybrid IVR survey were higher among women 20-49 compared to women 15-

19 (Table 5-8). The odds of consent were lower among women in the lowest wealth tertile compared to the middle and highest wealth groups (Odds Ratio (OR): 0.5, 95% Confidence Interval (CI): 0.3 – 0.8).

Similar to the Hybrid IVR results, women in the CATI arm were more likely to consent if they were 20 years or older. Urban women had higher odds of consenting, relative to rural women (OR 1.6, 95% CI: 1.1 – 2.4), and women with secondary education or higher had higher odds of consenting to the CATI, compared to women with no education (OR 1.6, 95% CI: 1.1 – 2.3). Wealth was no longer a predictor of CATI consent, after adjusting for other factors.

**Results of Multivariable analysis comparing phone survey completers and non- completers**

Results from multivariable logistic regression models comparing women who completed and who did not complete Hybrid IVR and CATI are presented in Table 5-9. Women who were 30 years and older were more likely to complete the Hybrid IVR survey than women 15-19, as well as women with secondary education, who had 2.5 times the odds of completion relative to women with no education (95% CI: 1.6 – 3.9). Urban women had twice the odds of completing compared to rural women (95% CI: 1.1 – 3.6). The odds of CATI completion were also elevated for women who were 20 years and older compared to younger teenagers. Women in urban communities were more likely to complete the CATI survey compared to their rural counterparts (OR 1.7, 95% CI: 1.1 – 2.5) as were women with a secondary education or more who had 40% higher odds of completing the survey than women with no education (OR 1.4, 95% CI: 1.1 – 2.0).

***Direct comparison of Hybrid IVR and CATI completers***

In multivariable logistic regression comparing Hybrid IVR completers with CATI completers, the only significant difference was education: Hybrid IVR women completers were more likely to have a secondary education or higher than CATI completers OR=1.7 (95% CI: 1.1 – 2.6) (Table 5-10).

**Implications of phone survey Completion on modern contraceptive use estimates**

Finally, we considered modern contraceptive use among completers and non-completers by data collection mode. The difference in modern contraceptive use among women who completed the Hybrid IVR survey (36.1%, 95% CI: 28.9% – 43.7%) was not statistically significantly different from the percent of women reporting modern contraceptive use who did not complete the Hybrid IVR survey (34.5%, 95% CI: 31.0% – 38.2%, p-value = 0.708) (Table 5-11). In line with the Hybrid IVR results, the difference in modern contraceptive use was not statistically significantly different between CATI completers (28.3%, 95% CI: 24.1% – 32.8%) and non-completers (29.5%, 95% CI: 25.2% – 34.0%, p-value = 0.692) (Table 5-12).

## *5.6 Discussion*

This chapter offers three main findings. First, we found CATI response and cooperation rates were more than double the Hybrid IVR rates due to high break-off post-consent among women assigned to the Hybrid IVR arm. Second, the low contact rates resulted in sample distortion for both modes. Third, non-response did not affect estimates of modern contraceptive based on R4

PMA2020 data, suggesting neither mode introduced non-response bias for estimates of modern contraceptive use among female cell phone owners.

Although evidence of non-response bias on estimates of modern contraceptive use was not established for either mode, we retained CATI as a better approach for phone surveys in settings similar to Burkina Faso. The two main reasons are higher break-off of the Hybrid IVR arm, and more profound distortion of Hybrid IVR completers than CATI completers. The sample distortion is expected: Hybrid IVR requires participants to answer the questionnaire without assistance and therefore has a higher cognitive burden than CATI, making it more difficult for the uneducated or for people who do not speak a majority language. [67] In this study, the response rate was among women randomized to the Hybrid IVR survey was only 12.3% among women with no education compared to 28.5% among women with secondary or higher level of education. In CATI survey, the response rates were high in all education groups.

This study illustrates that CATI phone follow-up surveys among women are feasible but suffer from noteworthy non-response in addition to frame error as described in the previous chapter. Failure to contact the sampled participants was the main cause of non-response, with 46% of Hybrid IVR and 41% of CATI participants classified as non-contacts. Non-contact is the main cause of non-response in developed countries as well. [51] Refusal to participate was minimal among both arms, with most of the refusals happening before consent. Finally, we saw that Hybrid IVR increased non-response due to lack of ability to respond with a phone keypad; with 3.5% of contacted participants willing to participate but unable to navigate the phone keypad to begin the survey. Recent RDD studies in the US lost 9% [66] and 7% [67] of the sample when transferring respondents from CATI to IVR. Among Hybrid IVR participants who were able to press 1 (and be transferred to IVR), there was substantial break-off, with 20% answering less than 50% of the survey questions. These cognitive constraints were not relevant to CATI but language constraints were apparent in both modes, with consented women who did not speak one of the five main languages in the R4 survey being less likely to complete the study: only 18% of women who consented to Hybrid IVR and spoke an 'other' language completed the survey, whereas 85% of women who consented and spoke an 'other' language completed the CATI survey.

No published studies to date that we know of have directly compared two remote cell phone survey interview modes in SSA. A study that compared CATI and IVR surveys among an RDD sample of respondents in Tanzania was completed in 2017, but the results have not yet been published. [34] The two modes used in our research, IVR and CATI, are also rarely compared in high income countries. [63] The studies that do exist used RDD or list-assisted RDD and sought to examine differences in reports of sensitive behaviors such as substance abuse, [63-65] youth risk behaviors, [66] and child maltreatment. [67] Not only is the direct comparison of two remote modes rare, it is uncommon for IVR studies in SSA to have a known sampling frame. Having a known sampling frame allows to anticipate loss to follow-up in phone follow-up surveys and provides the characteristics of non-respondents, allowing us to quantify the difference between respondents and non-respondents.

Studies using CATI for remote data collection in Lebanon, [41] Honduras and Peru [32] have examined response rates and related sample distortion. These studies enrolled men and women who completed a baseline FTF survey and were then followed-up using CATI. The study in Lebanon had an 82% response rate but did not compare respondent and non-respondent profiles. [41] The profile of completers in our study had a similar pattern as completers of phone surveys in Peru and Honduras, with the wealthy, educated and urban also more likely to complete phone follow-up surveys. [32] The response rate at first contact was 33% in Peru and 59% in Honduras. These surveys conducted as part of a World Bank study, however, found that young participants were more likely to complete the survey whereas our results found an opposite association. [32]

The study has notable strengths. Our sampling frame was a population-based FTF sample that allowed us to assess the characteristics related to consent and completion regardless of the woman's participation in the phone follow-up survey. The FTF survey provided identical measures for phone survey respondents and non-respondents, whereas many studies evaluating non-response error rely on surveys that do not have directly comparable indicators for non-responders. These alternate data sources, such as administrative data, census or DHS, can describe sociodemographic sample distortion but are unable to assess the relationship between non-response and behavioral measures of interest (such as contraceptive use in our survey). Knowledge of the characteristics of phone non-response is also valuable for survey design, in order to oversample women who are less likely to respond and adjust remote data collection estimates through weighting. Beyond the use of baseline data to characterize non-respondents, the randomization of respondents to CATI or Hybrid IVR is a strength when comparing two data collection modes, allowing a more robust comparison of non-response by phone survey mode.

Our four survey outcomes rates must be interpreted with caution. The follow-up design of the survey meant women had already participated in PMA2020 survey and consented to follow-up yielding potentially higher response rates than an RDD survey. Phone follow-up surveys conducted after a FTF survey in SSA have only investigated response rates among populations including both male and female respondents but have generally found higher response rates than in our CATI survey: 69% in a four- wave panel survey in South Sudan, [140] 98% across 14 waves of data collection among farmers in Tanzania [114], 99% response rate over six waves in Mali [113], and 75% response rate in urban Tanzania. [46, 135] Our CATI response rate was likely much lower due to the 11 month span between enrollment and follow-up, and because women are harder to reach than men via cell phones. [96] Research from LMIC show that rapid follow-up (defined as less than a month since first contact) after enrolling in a FTF survey is key to reduce non-contact, the main cause of non-response in our survey. [32, 141] Conversely, the response rate for the Hybrid IVR survey was higher than other IVR surveys in SSA, most likely due to the selection of cell phone owners in our survey and to the human introduction. The only response rates available for comparison are from surveys that use RDD sampling, not follow-up surveys. Surveys in Mozambique and Zimbabwe had 9% and 8% response rates [46, 70], respectively, and a more recent RDD in Ghana had a 21% response rate. [72] Another study design limitation was the difference in supervision quality by mode as reflected in the differences in refusal and contact rate between CATI and Hybrid IVR arms. The CATI interviewers had a supervisor that was better able to coach them to both find and persuade respondents to participate compared to the Hybrid IVR supervisor. Such differences highlight the importance of study implementation as a key factor contributing to data quality.

The generalizability of the study is limited by geography and time. Burkina Faso has some unique traits related to cell phone data collection including high language fractionalization, making remote data collection more difficult because interviewers are not able to speak all languages of respondents. Low female literacy reduces cell phone survey options to interviewer based interactions. However, these traits are mirrored in many West African countries. Finally, as cell phone ownership is rapidly increasing throughout SSA, our findings may have limited relevance in the near future in Burkina Faso and West Africa more broadly, as the profile of phone owners is likely to change over time.

## *5.7 Conclusion*

We identified the characteristics related to CATI and Hybrid IVR consent and completion and concluded that among PMA2020's target population, women of reproductive age, CATI results in a more representative sample. We did not find evidence of non-response bias for our outcome of interest, modern contraceptive use, for either mode of data collection. This study is one of the first to analyze phone survey non-response error and to compare CATI and Hybrid IVR in SSA. This study will inform remote data collection efforts in West Africa, particularly among women.

## 5.8 Tables for Chapter 5

*Figure 5-1. Study enrollment flow chart*

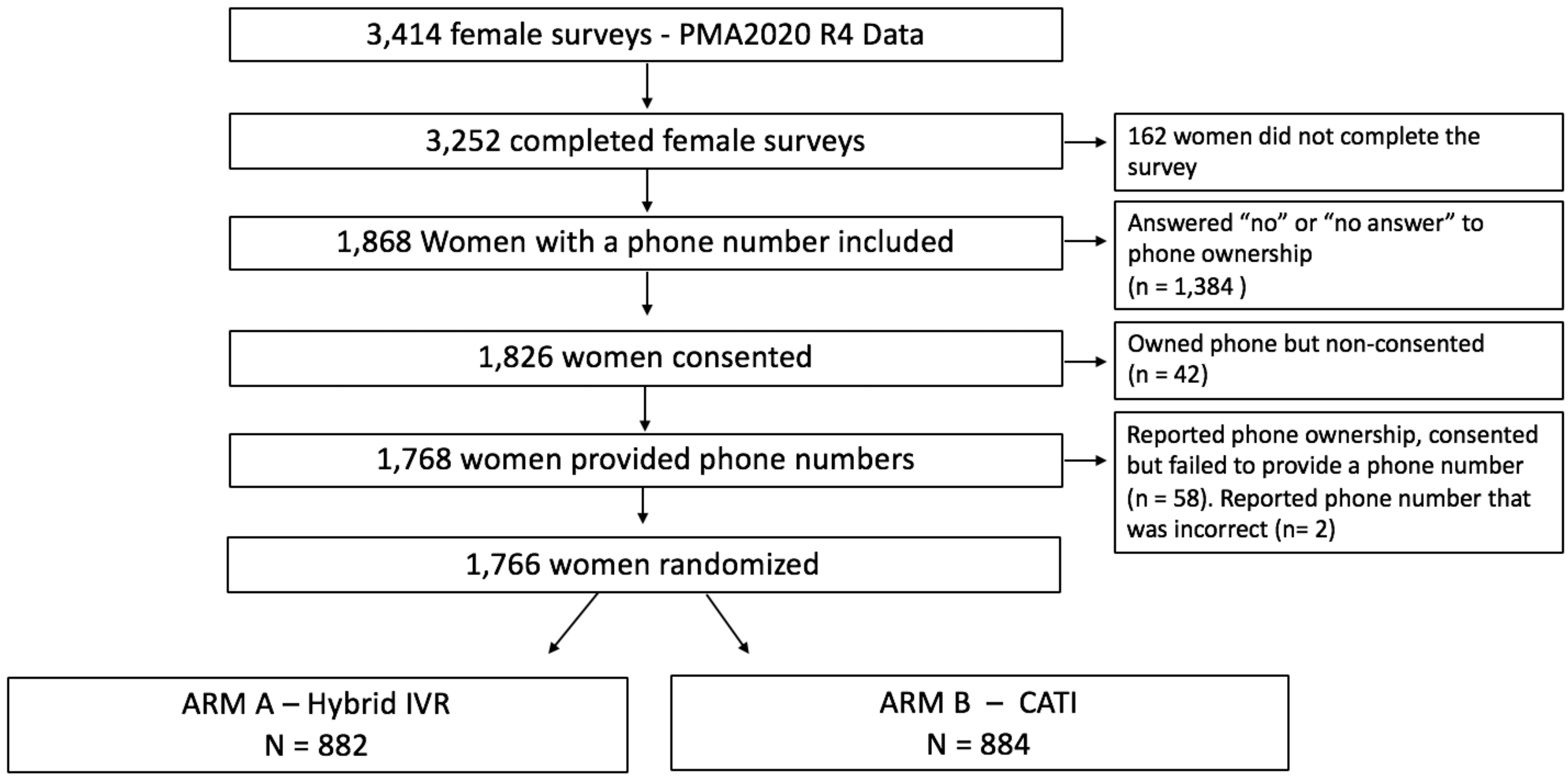


*Table 5-3. Sample characteristics of female cell phone owners in Burkina Faso overall and by study arm*

| | % of total study population | % of Hybrid IVR arm | % of CATI only arm |
|---|---|---|---|
| Total | n=1766 | n=882 | n=884 |
| **Age** | | | |
| Mean | 28.5 | 28.6 | 28.4 |
| 15-19 | 19.1 | 18.5 | 19.8 |
| 20-24 | 19.8 | 19.8 | 19.8 |
| 25-29 | 19.4 | 19.9 | 18.8 |
| 30-34 | 15.0 | 14.3 | 15.7 |
| 35-39 | 11.7 | 11.6 | 11.9 |
| 40-44 | 9.0 | 9.3 | 8.6 |
| 45-49 | 6.0 | 6.6 | 5.4 |
| | | | |
| **Urban/rural** | | | |
| Urban | 33.4 | 33.3 | 33.4 |
| Rural | 66.6 | 66.7 | 66.6 |
| | | | |
| **Marital status** | | | |
| Currently not in union | 34.7 | 35.4 | 34.1 |
| Currently in union | 65.3 | 64.4 | 65.9 |
| | | | |
| **Highest school attended** | | | |
| Never | 42.0 | 42.0 | 41.9 |
| Primary | 19.9 | 20.5 | 19.2 |
| Secondary or higher | 38.1 | 37.4 | 38.9 |
| | | | |
| **HH Wealth (tertile)** | | | |
| Lowest | 15.3 | 15.5 | 15.1 |
| Middle | 18.1 | 17.6 | 18.7 |
| High | 66.6 | 66.9 | 66.3 |
| | | | |
| **Parity** | | | |
| Yes | 71.9 | 72.5 | 71.3 |
| No | 28.1 | 27.5 | 28.7 |
| | | | |
| **HH Electricity** | | | |
| Yes | 62.0 | 62.1 | 62.0 |
| No | 38.0 | 37.9 | 38.0 |

*Table 5-4. Final Disposition Code by study arm (CATI versus Hybrid IVR) among female cell phone owners in Burkina Faso*

| | Hybrid IVR | | CATI | |
|---|---|---|---|---|
| | **N** | **%** | **N** | **%** |
| **Non-Respondents** | | | | |
| Non-contact (didn't pick up) | 225 | 25.5 | 228 | 25.8 |
| Non-contact (someone picked up the phone call but interviewer never spoke with the woman) | 173 | 19.6 | 134 | 15.2 |
| Refusal pre-consent | 90 | 10.2 | 54 | 6.1 |
| Refusal | 14 | 1.6 | 6 | 0.7 |
| Break-off IVR | 31 | 3.5 | N/A | N/A |
| Break-off (consented but less than 50% of relevant questions answered) | 174 | 19.7 | 17 | 1.9 |
| **Respondents** | | | | |
| Partial (50-80% of relevant questions answered) | 18 | 2.0 | 5 | 0.6 |
| Complete (more than 80% of relevant questions answered) | 157 | 17.8 | 440 | 49.8 |
| **Total** | 882 | | 884 | |

*Table 5-5. Follow-up phone call outcome rates by study arm (CATI versus Hybrid IVR) among female cell phone owners in Burkina Faso*

| Outcome Rates | Hybrid IVR | CATI |
|---|---|---|
| **Response Rates** | | |
| Response Rate 5 | 17.7% | 49.7% |
| Response Rate 6 | 19.8% | 50.3% |
| **Cooperation Rates** | | |
| Cooperation Rate 1 | 32.2% | 84.0% |
| Cooperation Rate 2 | 35.9% | 85.2% |
| **Refusal Rate** | | |
| Refusal Rate 3 | 11.9% | 6.8% |
| **Contact Rate** | | |
| Contact Rate 3 | 54.8% | 59.1% |

*Table 5-6. Percent consenting to or completing the Hybrid IVR survey by women's sociodemographic characteristics among female cell phone owners randomized to the Hybrid IVR Survey arm*

| | Hybrid IVR | | | | | |
|---|---|---|---|---|---|---|
| | **% who did not consent** | **% who consented** | p-value | **% did not complete** | **% who completed** | p-value |
| Total | 57% (N= 507) | 43% (N = 375) | | 80% (N= 707) | (20%) N= 175 | |
| **Age** | | | <0.001 | | | 0.479 |
| 15-19 | 73.0 | 27.0 | | 84.1 | 15.9 | |
| 20-24 | 60.6 | 39.4 | | 78.3 | 21.7 | |
| 25-29 | 52.8 | 47.2 | | 82.4 | 17.6 | |
| 30-34 | 52.4 | 47.6 | | 74.6 | 25.4 | |
| 35-39 | 52.9 | 47.1 | | 82.4 | 17.6 | |
| 40-44 | 52.4 | 47.6 | | 79.3 | 20.7 | |
| 45-49 | 44.8 | 55.2 | | 77.6 | 22.4 | |
| | | | | | | |
| **Urban/rural** | | | 0.002 | | | 0.002 |
| Rural | 64.6 | 35.4 | | 90.8 | 9.2 | |
| Urban | 53.9 | 46.1 | | 74.8 | 25.2 | |
| | | | | | | |
| **Marital status** | | | 0.052 | | | 0.282 |
| Currently not in union | 61.9 | 38.1 | | 78.2 | 21.8 | |
| Currently in union | 55.1 | 44.9 | | 81.2 | 18.8 | |
| | | | | | | |
| **Highest school attended** | | | 0.149 | | | <0.001 |
| Never | 60.9 | 39.1 | | 87.6 | 12.3 | |
| Primary | 52.5 | 47.5 | | 80.7 | 19.3 | |
| Secondary or higher | 56.4 | 43.6 | | 71.5 | 28.5 | |
| | | | | | | |

| **HH Wealth (tertile)** | | | <0.001 | | | <0.001 |
|---|---|---|---|---|---|---|
| Lowest | 71.5 | 28.5 | | 91.2 | 8.8 | |
| Middle | 56.1 | 43.9 | | 89.0 | 11.0 | |
| Highest | 54.6 | 45.4 | | 75.3 | 24.7 | |
| | | | | | | |
| **Parity** | | | 0.011 | | | 0.502 |
| No | 64.2 | 35.8 | | 78.7 | 21.3 | |
| Yes | 54.8 | 45.2 | | 80.7 | 19.3 | |
| | | | | | | |
| **HH Electricity** | | | 0.122 | | | <0.001 |
| No | 60.8 | 39.2 | | 87.4 | 12.6 | |
| Yes | 55.5 | 44.5 | | 75.7 | 24.3 | |
| | | | | | | |
| **Language** | | | 0.008 | | | 0.005 |
| Dioula | 59.3 | 40.7 | | 82.7 | 17.3 | |
| French | 57.7 | 42.3 | | 75.0 | 25.0 | |
| Fulfulde | 83.3 | 16.7 | | 91.7 | 8.3 | |
| Gourmantchema | 51.5 | 48.5 | | 90.9 | 9.1 | |
| Moore | 53.6 | 46.4 | | 79.1 | 20.9 | |
| Other | 78.4 | 21.6 | | 96.1 | 3.9 | |

*Table 5-7. Percent consenting to or completing the CATI survey by women's sociodemographic characteristics among female cell phone owners randomized to the CATI survey arm*

| | CATI | | | | | |
|---|---|---|---|---|---|---|
| | **% who did not consent** | **% who consented** | p-value | **% did not complete** | **% who completed** | p-value |
| Total | 48% (N = 425) | 52% (N = 459) | | 49% (N = 439) | 51% (N = 445) | |
| **Age** | | | <0.001 | | | <0.001 |
| 15-19 | 65.7 | 34.3 | | 66.3 | 33.7 | |
| 20-24 | 52.6 | 47.4 | | 54.9 | 45.1 | |
| 25-29 | 37.4 | 62.6 | | 39.8 | 60.2 | |
| 30-34 | 37.4 | 62.6 | | 39.6 | 60.4 | |
| 35-39 | 52.4 | 47.6 | | 53.3 | 46.7 | |
| 40-44 | 35.5 | 64.5 | | 36.8 | 63.2 | |
| 45-49 | 45.8 | 54.2 | | 45.8 | 54.2 | |
| | | | | | | |
| **Urban/rural** | | | <0.001 | | | <0.001 |
| Rural | 61.0 | 38.9 | | 62.0 | 38.0 | |
| Urban | 41.6 | 58.4 | | 43.5 | 56.5 | |
| | | | | | | |
| **Marital status** | | | 0.919 | | | 0.941 |
| Currently not in union | 47.8 | 52.2 | | 49.8 | 50.2 | |
| Currently in union | 48.2 | 51.8 | | 49.6 | 50.4 | |
| | | | | | | |
| **Highest school attended** | | | 0.50 | | | 0.84 |
| Never | 52.7 | 47.3 | | 54.1 | 45.9 | |
| Primary | 47.1 | 52.9 | | 47.1 | 52.9 | |
| Secondary or higher | 43.6 | 56.4 | | 46.2 | 53.8 | |
| | | | | | | |
| **HH Wealth (tertile)** | | | <0.001 | | | <0.001 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Lowest | 63.9 | 36.1 | | 64.7 | 35.3 | |
| Middle | 54.6 | 45.4 | | 55.8 | 44.2 | |
| Highest | 42.7 | 57.3 | | 44.5 | 55.5 | |
| | | | | | | |
| **Parity** | | | 0.119 | | | 0.127 |
| No | 52.1 | 47.9 | | 53.6 | 46.4 | |
| Yes | 46.4 | 53.6 | | 48.0 | 52.0 | |
| | | | | | | |
| **HH Electricity** | | | 0.003 | | | 0.005 |
| No | 54.5 | 45.5 | | 55.7 | 44.3 | |
| Yes | 44.2 | 55.8 | | 46.0 | 54.0 | |
| | | | | | | |
| **Language** | | | 0.002 | | | 0.001 |
| Dioula | 45.3 | 56.7 | | 45.3 | 54.7 | |
| French | 43.9 | 56.1 | | 46.8 | 53.2 | |
| Fulfulde | 63.6 | 36.4 | | 54.6 | 45.4 | |
| Gourmantchema | 60.6 | 30.4 | | 63.6 | 36.4 | |
| Moore | 56.7 | 53.3 | | 48.0 | 52.0 | |
| Other | 74.5 | 25.5 | | 78.3 | 21.6 | |

*Table 5-8. Odds of consenting to the phone follow-up survey by background characteristics: multivariable logistic regression analyses among the two study arms*

| | **Adjusted Odds Ratio Hybrid- IVR Consenters vs. Non-Consenters** | **Adjusted Odds Ratio CATI Consenters vs. Non-Consenters** |
|---|---|---|
| **Age group** | N= 882 | N= 884 |
| 15-19 (reference) | | |
| 20-24 | **1.7 (1.1 – 2.7)** | **1.8 (1.1 – 2.8)** |
| 25-29 | **2.7 (1.7 – 4.2)** | **3.7 (2.3 – 6.0)** |
| 30-35 | **2.9 (1.7 – 4.8)** | **3.9 (2.4 – 6.4)** |
| 35-39 | **2.8 (1.6 – 4.8)** | **2.1 (1.2 – 3.6)** |
| 40-44 | **3.0 (1.7 – 5.5)** | **4.7 (2.6 – 8.6)** |
| 45-49 | **4.0 (2.1 – 7.8)** | **3.3 (1.7 – 6.7)** |
| **Residential area** | | |
| Rural (reference) | | |
| Urban | 1.4 (0.9 – 2.0) | **1.6 (1.1 – 2.4)** |
| **Highest school attended** | | |
| No education (reference) | | |
| Primary | 1.4 (0.9 – 2.1) | 1.2 (0.8 – 1.8) |
| Secondary or more | 1.4 (0.9 – 2.0) | **1.6 (1.1 – 2.3)** |
| **HH Wealth (tertile)** | | |
| Lowest (reference all other groups) | **0.5 (0.3 – 0.8)** | 0.7 (0.4 – 1.1) |
| Highest (reference all other groups) | 0.8 (0.5 – 1.2) | 1.1 (0.7 – 1.7) |

*Table 5-9. Odds of completing the phone follow-up survey by background characteristics: multivariable logistic regression analyses among the two study arms*

| | **Adjusted Odds Ratio Hybrid- IVR Completers vs. Non-Completers** | **Adjusted Odds Ratio CATI Completers vs. Non-Completers** |
|---|---|---|
| **Age group** | N= 882 | N= 884 |
| 15-19 (*reference*) | | |
| 20-29 | 1.4 (0.8 – 2.3) | **2.3 (1.5 – 3.4)** |
| 30-39 | **2.1 (1.2 – 3.8)** | **2.7 (1.8 – 4.2)** |
| 40-49 | **2.1 (1.1 – 4.0)** | **3.8 (2.3 – 6.4)** |
| **Residential area** | | |
| Rural (*reference*) | | |
| Urban | **2.0 (1.1 – 3.6)** | **1.7 (1.1 – 2.5)** |
| **Highest school attended** | | |
| No education (*reference*) | | |
| Primary | 1.5 (0.9 – 2.4) | 1.3 (0.9 – 1.9) |
| Secondary or more | **2.5 (1.6 – 3.9)** | **1.4 (1.1 – 2.0)** |
| **HH Wealth (tertile)** | | |
| Lowest (reference all other groups) | 0.8 (0.4 – 1.8) | 0.7 (0.4 -1.2) |
| Highest (reference all other groups) | 1.3 (0.7 – 2.5) | 1.1 (0.7 – 1.7) |

*Table 5-10. Odds of completing the Hybrid IVR survey compared to completing the CATI survey by background characteristics: multivariable logistic regression analyses*

| | **Adjusted Odds Ratio Hybrid IVR vs. CATI among survey completers** *(Reference group CATI)* |
|---|---|
| **Age group** | |
| 15-19 (*reference*) | |
| 20-29 | 0.9 (0.5 – 1.5) |
| 30-39 | 0.9 (0.5 – 1.7) |
| 40-49 | 1.1 (0.5 – 2.1) |
| **Residential area** | |
| Rural (*reference*) | |
| Urban | 1.4 (0.8 – 2.4) |
| **Highest school attended** | |
| No education (*reference*) | |
| Primary | 1.3 (0.8 – 2.2) |
| Secondary or more | **1.7 (1.1 – 2.6)** |
| **HH Wealth (tertile)** | |
| Lowest (reference all other groups) | 1.1 (0.5 – 2.5) |
| Highest (reference all other groups) | 1.4 (0.8 - 2.7) |

*Table 5-11. Modern contraceptive use by Hybrid IVR completion among all women randomized to Hybrid IVR*

| **Percent of women using modern method (95% Confidence Internal)** | |
|---|---|
| All women randomized to Hybrid IVR | 34.8 (31.7 – 38.1) |
| Completed survey | |
| No | 34.5 (31.0 – 38.2)* |
| Yes | 36.1 (28.9 – 43.7) |

**P-value = 0.708*

*Table 5-12. Modern contraceptive use by CATI completion among all women randomized to CATI*

| **Percent of women using modern method (95% Confidence Internal)** | |
|---|---|
| All women randomized to CATI | 28.9 (25.9 – 32.0) |
| Completed survey | |
| No | 29.5 (25.2 – 34.0)* |
| Yes | 28.3 (24.1 – 32.8) |

**P-value = 0.692*

# 6. Comparability of modern contraceptive use estimates between a face-to-face survey and a cellphone survey among women of reproductive age in Burkina Faso

## *6.1 Abstract*

**Introduction:** The proliferation of cell phone ownership in Sub-Saharan Africa (SSA) presents the opportunity to collect public health indicators at a lower cost by transitioning from face-to face (FTF) to cell phone surveys. However, this substitution of survey modes raises questions of sample representativeness and data quality. This analysis assesses the equivalence of modern contraceptive prevalence estimates between a nationally representative FTF survey and a cell phone survey using random digit dialing (RDD) among women of reproductive age in Burkina Faso.

**Methods:** We analyzed data from two surveys conducted in Burkina Faso between December 2017 and May 2018. The FTF survey, a nationally representative sample of 3,556 women of reproductive age (15-49 years), was conducted by PMA2020 (Round 5). The RDD survey was a cross-sectional survey using computer-assisted telephone interviewing (CATI). We used quota restrictions so the RDD sample would mirror the underlying age and residence (urban/rural) distribution of the female population of Burkina Faso. Ultimately, 2,379 women completed the RDD survey. Our outcome measure was current use of modern contraception, limited to the five main methods used in Burkina Faso (IUD, implant, injectable, pill and condom). Our key independent variable of interest was survey mode (FTF versus RDD). We first evaluated differences in the distribution of sociodemographic characteristics between the two samples then generated RDD post-stratification weights to account for these differences. We then compared use of modern contraception by survey mode, using unweighted and weighted RDD estimates and tested for the equivalence of modern contraceptive use rates between the FTF sample and the RDD sample within a margin of +/-4%. Finally, we conducted a multivariable logistic regression to evaluate if the effect of survey mode on contraceptive use remained significant, after adjusting for women's sociodemographic factors.

**Results:** Compared to FTF respondents, women in the RDD sample were younger, a higher proportion had a secondary degree and spoke French. Altogether, 40% of RDD respondents reported modern contraceptive use (unweighted) versus 26% of FTF respondents. This difference, which remained unchanged after applying post-stratification weights to the RDD sample (39% weighted) surpassed the equivalence margin of 4%. The RDD sample also produced higher estimates of contraceptive use than the subsample of women who owned a phone in the FTF sample (32%). After adjusting for women's sociodemographic factors, the odds of contraceptive use were 1.9 times higher (95% CI: 1.6- 2.2) in the weighted RDD survey compared to the full FTF survey and 1.6 times higher (95% CI: 1.3 – 1.8) compared to FTF phone owners.

**Conclusions**: Modern contraceptive prevalence in Burkina Faso is over-estimated when using a cell phone RDD survey. Sample distortion associated with area of residence and age may affectively be addressed by applying quotas but doesn't prevent substantial bias in contraceptive use estimates. Further research should aim to explore causes of differential reporting of modern contraceptive use by survey modes.

## 6.2 Introduction

Reproductive health indicators are a cornerstone of public health policy and program planning in low- and middle-income countries (LMIC). The data inform government decisions in countries undergoing or beginning the demographic transition, as they forecast resources such as school, infrastructure and plan cities for young, growing populations. Demographers use these estimates to assess family planning trends, to compare progress within and between countries, and to generate population projections. [142, 143] Data from population-based surveys such as the Demographic and Health Surveys (DHS) and UNICEF's Multiple Indicator Cluster Survey (MICS) help monitor progress towards the Sustainable Development Goals 3 (Good Health & Wellbeing) and 5 (Gender Equity). [144]

Although the DHS and MICS have provided invaluable data for the past 30 years, there are two main challenges associated of face-to-face (FTF) surveys: cost and extensive field time to attain the needed population coverage. [37, 145] The challenges of FTF surveys in LMIC are contrasted with the three primary advantages of phone surveys identified in the 1990s that remain pertinent today: speed of data collection, cost efficiency, and ability to supervise interviewers throughout data collection. [133] Cost reduction is a major consideration in LMIC; one study conducted in Honduras showed a decrease from $40 USD per interview using FTF to $17 USD using a cellphone survey with interactive voice response (IVR). [32]

Cell phone ownership is growing exponentially in Sub-Saharan Africa (SSA), mostly driven by urbanization and low cost of purchasing a cell phone and airtime. [35, 146] There are currently over 690 million cell phone subscriptions in SSA and over 1 billion subscriptions are anticipated by 2021. [36] By 2021, mobile phone penetration, which is the percent of unique mobile phone users within a specific population, [38] is expected to be at least 50 percent in SSA. [39] In this context, demographers are considering cell phone surveys to track national health indicators [34, 48, 72] but the quality of phone survey estimates needs to be compared with FTF before substituting data collection modes. Five types of survey error dictate survey quality: frame and non-response, which relate to survey representation, and specification, measurement and processing errors, which relate to the quality of data collected. These errors are impacted by mode of data collection, which also determines sampling approaches. [54] Cell phone surveys are traditionally collected using one of three modes – computer assisted telephone interviews (CATI), IVR or short message service (SMS), but we direct our attention to the assessment of survey errors comparing CATI to FTF surveys, which is the focus of our study.

Few studies comparing cell phone surveys to FTF studies have been conducted in LMIC. [30] Those that have compared cell phone and FTF survey results have used a census or a FTF reference population-based survey such as DHS, rather than a concurrent study with the same questions as the cell phone survey. Two recent random digit dial (RDD) studies, an IVR survey in Ghana about general health [72] and a CATI in the Ivory Coast about HIV risk behaviors [71] have followed that approach. In Ghana, two-thirds of RDD respondents were male and more than half were 15-24 years of age whereas only 48% of the population is male and 30% ages 15-24 in the 2017 census. [72] In the Ivory Coast study, the composition of the RDD sample was different than the DHS sample distribution, with over-representation of urban individuals and males. Self-reported HIV prevalence was similar after stratification by sex, age group, level of education and place of residence. [71]

Assessment of sampling frame and non-response errors between CATI and FTF were explored in the two previous chapters of this dissertation. Frame error occurs when the sample population does not correctly reflect the target population, with certain units excluded or erroneously included in the sample frame. [117] Frame error is of serious concern in LMIC as a number of studies report different distributions of sample characteristics between phone owners and non-owners. [48, 147, 148] In Aim 1, we showed that cell phone owners and non-owners differed in many respects, including reproductive health behaviors such as contraceptive use.

Non-response error occurs when the population that responds to a survey is different from those who did not respond to the survey but were in the sample frame. [73, 136] The risk of non-response is higher for cell phone surveys than for FTF surveys because individuals are likely not to pick-up the call or if they do, are less likely to participate than if an interviewer comes to their house. This also leads to lower response rates in cell phone surveys than in FTF surveys. [12]

There has been concerted effort to reduce non-response in RDD phone surveys. For example, studies have explored increasing participation of females, by looking at the impact of incentive amount on response rate and identifying whether a female or male voice increases women's response rates to phone surveys. [34] A randomized study assessing the effect of incentives on phone survey participation conducted in four countries (Afghanistan, Ethiopia, Mozambique, Zimbabwe) found higher response rates in the group who were offered compensation but the impact of the incentive varied by country. [70] The same study found that the countries with higher mobile penetration (Ethiopia and Zimbabwe) produced samples that more closely resembled the actual population, and also produced more precise outcome estimates compared to the countries with lower mobile phone penetration. However, over-representation of individuals speaking the majority languages (Amharic and Shona) was still evident in RDD samples in Ethiopia and Zimbabwe.

The third aim of this dissertation examines whether an RDD cell phone survey produces equivalent estimates of modern contraceptive use as a reference FTF survey conducted in the general population after applying post-stratification weights to account for RDD sample distortion.

## *6.8 Methods*

We used two datasets for this analysis. The first dataset is the Burkina Faso PMA2020 Round 5 (R5) survey, designed to track key family planning indicators under the Family Planning 2020 initiative. [8] PMA2020 Round 5 implemented a two-stage stratified cluster design to select a national sample of women of reproductive age in Burkina Faso. This design starts with a probability proportional to size selection of 83 geographic clusters stratified by urban and rural areas followed by a random selection of about 35 households within each sample cluster. Detailed sampling methods and procedures are available in the methods section of this dissertation. The fifth round of the survey in Burkina Faso was conducted between November

2017- January 2018 and included a total of 2,811 households (98.5% response rate) and 3,659 females (97.8% response rate). [81]

Data collection was performed by a network of trained female interviewers who conducted FTF interviews with the head of selected households and with all eligible females 15-49 years from the selected households. Interviewers recorded responses directly on cell phones and uploaded the data when cellular network was available.

The second dataset was from the RDD CATI survey, which took place four months after PMA2020 Round 5 (R5), from April to May 2018. RDD was chosen because the mobile network operators in Burkina Faso do not share lists of valid phone numbers. CATI was chosen over Hybrid IVR based on Aim 2 results, indicating higher response rates and less sample distortion in the CATI sample than Hybrid IVR sample. Results from Aim 2 also informed the sampling strategy of the RDD cross-sectional survey, which included quotas by age and area of residence to improve representativeness and reduce design effects. [48] A quota is a limited quantity of a certain outcome, in our case, the number of completed interviews, by age (15-19 years old or 20-49 years old) and by geographic location (urban or rural). We established the quota groups by comparing the follow-up respondents and non-respondents from Aim 2 to see which groups were less likely to respond to the phone interview. We found that rural, young and uneducated women were least likely to respond. We included only residence and age in our quota groups for ease of implementation. We then used our target sample data, PMA2020 Round 4 female respondents, to calculate the percent of women in each of the four groups (see Table 6-1). We proportionally selected our target sample size from the sample distribution of PMA2020 for each quota group. Once we reached the target number of completed interviews in each group, the questionnaire was programmed by the cellphone survey administration firm, Viamo, an international mobile technology survey company, [107] to automatically end the interview after thanking the respondent.

*Table 6-1. The number of completed surveys needed by quota group*

| | **Rural** | **Urban** |
|---|---|---|
| **15-19** | 388 (16.4%) | 130 (5.5%) |
| **20-49** | 1417 (60%) | 428 (18.1%) |
| **Total** | | **2363** |

In the absence of active phone number lists, a significant proportion of phone numbers generated by RDD are invalid. The pilot conducted in February 2018 estimated that 58% of RDD generated numbers were not assigned to a subscriber. To address this issue, we sent out an IVR "validation/pre-notification" phone call to all generated RDD numbers in order to eliminate invalid phone numbers so the interviewers would not waste time calling non-existent phone numbers. We defined a phone number as valid if the outcome of the call, as recorded by the mobile phone company, had one of three following outcomes (1- No Answer, 2- Normal Clearing or 3- Normal Unspecified), as described in detail in the Methods Section of this dissertation.

## Questionnaires

The PMA2020 Burkina Faso female questionnaire has approximately 200 questions and usually takes less than 40 minutes to complete. [149] The questions are standard, based on the DHS. The household survey lists household members, counts assets and livestock, and documents house, water, sanitation, and hygiene conditions. The female survey collects sociodemographic information including cell phone ownership, and reproductive health measures, including current use of contraception as described in the next section on measurement. [89]

The RDD questionnaire was limited to 19 questions that allowed comparison of modern contraceptive use estimates with the FTF survey. Four questions helped establish the eligibility of the respondent, followed by 5-6 demographic questions, five questions were about contraceptive awareness, and 3-4 questions on contraceptive use. The RDD questions were identical to the FTF questions with a few adaptations to accommodate the questionnaire phone format. The RDD survey was available in French and six local languages.

### *Measures – defining final disposition codes for RDD survey*

We used the 9th edition American Association for Political Opinion Research (AAPOR) final disposition codes to classify call outcomes. [103] The invalid phone numbers identified during the screening process were not assigned a disposition code. The 13 disposition codes are divided into four groups (Not Eligible; Unknown Eligibility – non-interview; Eligible – non-interview; Interview) and are presented in Table 6-2.

Non-eligible respondents were categorized into four codes. Respondents were ineligible if they were male, or were >49 years and <15 years or did not speak one of the seven survey languages. The fourth group consisted of women who spoke one of the survey languages and were between the ages of 15-49 but were ineligible due to quota restrictions.

Unknown eligibility was captured in four disposition codes and consisted of respondents whose eligibility was not known. Calls that were never picked up were classified as "No answer" and calls answered by a voice mailbox were classified as "Telephone answering device". Respondents who answered but for whom age, gender or area of residence was not known were classified as "No screener completed". Finally, respondents who spoke one of the seven survey languages but who did not speak the same language as the interviewer and were not reached during subsequent attempts were classified as "Other (language not matched with interviewer)".

The next group "eligible, not interviewed", was divided into three codes and consisted of women ages 15-49 who spoke one of the seven survey languages and were not excluded due to quota restrictions. The first was refusal before consent, after eligibility was established. The second was refusal at consent and the final was "break-off", corresponding to a consenting respondent who completed less than 50% of questions.

The final group included women who completed the interview, classified as a partial interview when 50-80% of questions were answered and a complete interview when 80% of questions or more were answered.

*Table 6-2. Call Disposition codes for RDD Survey*

| AAPOR Code | Title | Definition |
|---|---|---|
| | **Not Eligible** | |
| (4.71) | Gender (not female) | Male |
| (4.72) | Age | Female and age <15 or >49 years |
| (4.73) | Language | Female and none of the 7 languages available in survey |
| (4.8) | Quota Filled | Respondent was female and age-eligible but due to quota restrictions was not interviewed |
| | **Unknown Eligibility, Non-Interview** | |
| UH (3.13) | No Answer | Phone call not picked-up |
| UH (3.14) | Telephone answering device | Phone call went to voice mail |
| UH (3.21) | No screener completed – talked with respondent but hung-up or refused | Respondent picked- up the call but interviewer was unable to confirm eligibility |
| UO (3.90) | Other (Language not matched with interviewer ) | Respondent spoke one of seven survey languages but the interviewer did not speak the same language |
| | **Eligible, Non-Interview** | |
| R (2.111) | Refusal pre-consent but confirmed female and 15-49 | Eligible respondent refused to participate before consent |
| R (2.11) | Refusal at consent | The respondent refused the study during consent |
| R (2.1) | Break-off (consented but less than 50% of relevant questions answered) | The respondent consented but answered less than 50% of the questions |
| | **Interview** | |
| P (1.2) | Partial (50-80% of relevant questions answered) | The respondent consented and answered between 50-80% of the questions |
| I (1.1 | Complete (more than 80% of relevant questions answered) | The respondent consented and answered more than 80% of the survey questions |

*Measures – Independent & Dependent Variables*

The outcome of interest was a binary measure of modern contraceptive use, based on two questions that were asked identically in the two surveys. The first question asked whether the respondent or her partner was currently using a form of contraception ("Are you or your partner currently doing something or using any method to delay or avoid getting pregnant?"). If the respondent responded positively, she was asked to specify the type of method used ("Which method or methods are you using?"). If the respondent identified a modern method (as specified below), she was classified as a user of modern contraception. Long-acting reversible contraceptive methods include implants and IUDs, methods that provide protection years at a time but can be reversed. [93]

Traditionally, measures of modern contraceptive use include all modern contraceptives available in a country. In Burkina Faso, the PMA20200 FTF survey asked respondents about 12 modern contraceptive methods: male and female sterilization, implant, IUD, injectables, pill, emergency contraception, male condom, female condom, diaphragm, foam/jelly and LAM. However, the RDD survey only collected data about five methods: implants, injectables, pills, condoms and IUDs, which covers 98.8% of modern contraceptive methods reported during R5. Thus we limited the definition of modern contraceptive use to these five methods for both the FTF and RDD surveys in this study. [81] In addition, we constructed a five-category indicator of method mix, including the following contraceptive users: implant, IUD, injectables, pills, and condoms.

We selected independent measures related to modern contraceptive use cited in the literature as well as factors related to phone ownership, that were collected in both surveys to conduct our analysis. [124, 125, 141] The independent variable of interest was mode of data collection. The FTF survey was the reference group, and RDD was the comparison group. Women's sociodemographic characteristics included age (15-19, 20-24, 25-29, 30-34, 35-39, 40-45, 45-49), current union status (in union – i.e., currently married or living with a partner vs. not in union), residential area (urban vs. rural), educational attainment (none, primary, or secondary and higher) and language of survey (Moore, French, Dioula, Fulfulde, Gourmantchema, Birifor or Bwamu). We also included parity (ever vs. never).

*Missing Data*

The RDD data had item non-response due to internet outages at the call center. The internet was sporadically cut due to electricity brown-outs, and as a result some answers were not recorded when interviews occurred during these electricity cuts (missing completely at random assumed). We used the hot deck method [110] to impute missing values for three variables: age (43 missing values, 1.8%), residence (10 missing values, 0.4%) and education (10 missing values, 0.4%).

## *6.3 Analysis*

*Call Outcomes*

Replicating the analysis conducted in Aim 2, we also created four key call outcome indicators – response rate, cooperation rate, refusal rate and contact rate – using the aforementioned call disposition codes aligned with AAPOR standards (Table 6-3).

- AAPOR's ***response rate*** corresponds to the number of individuals who complete the phone interview (complete or partial), over all attempted calls. AAPOR distinguishes two subcategories: response rate 1 includes individuals who complete 80% of the interview or more, while response rate 2 includes individuals who complete 50% of the interview or more.

- AAPOR's ***cooperation rate*** is similar to response rate, but includes only identified eligible individuals in the denominator. Cooperation rate 1 includes individuals who complete 80% of the interview or more, and Cooperation rate 2 includes individuals who complete 50% of the interview or more.
- AAPOR's ***refusal rate*** corresponds to the number of individuals who refuse to be interviewed, among all attempted calls.
- Finally, AAPOR's ***contact rate*** corresponds to the number of calls in which one member of the unit was reached (someone answered the call, regardless of that person's eligibility) among all attempted calls.

To improve on the specificity of these outcome measures, AAPOR also recommends calculating rates that exclude an estimated number of unknown eligibility phone numbers from the denominator for response and contact rates. Based on pilot data collected in February 2018, we estimated that 20% of calls with unknown eligibility would in fact include an eligible woman. Applying this correction, we defined Response rates 3 & 4 (same numerator as Response rates 1 & 2 but the denominator excluded 80% of attempted calls with unknown eligibility) and Contact rate 2 (same numerator as Contact rate 1 with the denominator excluding 80% of attempted calls with unknown eligibility).

*Table 6-3. Survey Outcome Rates*

| Response Rates | | Explanation[+] |
|---|---|---|
| Response rate 1: | $\frac{I}{I+P+R+NC+O+UH+UO}$ | *Minimum response rate.*<br>All individuals who complete more than 80% of survey /<br>All attempted calls |
| Response rate 2: | $\frac{I+P}{I+P+R+NC+O+UH+UO}$ | All individuals who complete more than 50% of survey /<br>All attempted calls |
| Response rate 3: | $\frac{I}{I+P+R+NC+O+0.2*(UH+UO)}$ | All individuals who complete more than 80% of survey /<br>All attempted calls minus 80% of calls with unknown eligibility |
| Response rate 4: | $\frac{I+P}{I+P+R+NC+O+0.2*(UH+UO)}$ | All individuals who complete more than 50% of survey /<br>All attempted calls minus 80% of calls with unknown eligibility |
| **Cooperation Rates** | | |
| Cooperation rate 1 : | $\frac{I}{I+P+R+O}$ | All individuals who complete more than 80% of survey /<br>Eligible individuals who were ever contacted |
| Cooperation rate 2: | $\frac{I+P}{I+P+R+O}$ | All individuals who complete more than 50% of survey /<br>Eligible individuals who were ever contacted |
| **Refusal Rate** | | |
| Refusal rate 3: | $\frac{R}{I+P+R+NC+O}$ | All individuals who refused to complete the survey /<br>All attempted calls |
| **Contact Rate** | | |
| Contact rate 1: | $\frac{I+P+R+O}{I+P+R+NC+O+UH+UO}$ | All phone numbers that answered the call /<br>All phone numbers |
| Contact rate 2: | $\frac{I+P+R+O}{I+P+R+NC+O+0.2*(UH+UO)}$ | All phone numbers that answered the call /<br>All attempted calls minus 80% of calls with unknown eligibility |

*Weighting RDD sample*

To address RDD sample distortion, we created RDD post-stratification weights, based on three sociodemographic characteristics of women of reproductive age in Burkina Faso (area of residence, age and education). These factors were chosen based on Aim 1 & 2 results within the constraints of the limited sociodemographic information collected in the RDD survey. The PMA2020 R5 sample served as the reference population in the absence of census data in Burkina Faso (last census was conducted in 2006). We calculated the ratio of RDD respondents to R5 in urban/rural groups and seven age groups (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49). We then calculated the ratio of RDD respondents to R5 respondents in three education groups: no education, primary education and secondary education. Post-stratification weights were computed by multiplying each education ratio group with each of the residence/age groups. There was no design weight for the cellphone sampling because all RDD calls were randomly generated.

*Descriptive Analyses*

To examine distributions of the aforementioned independent variables, all of which are categorical except age, we conducted univariate analysis, looking at patterns of response in the FTF and RDD samples. In the FTF sample, we examined distributions among all women (n=3,659), who represent the target population and among a selected sample of cell phone owners (n=2,027), who represent the sample frame of the RDD survey. The R5 data were adjusted for sampling weights, which address disproportionate two-stage cluster sampling and non-response rates. [91] Turning to the RDD data, we included women who were partial (50-80% of questions answered) or complete respondents (more than 80% of questions answered) (N=2,379) in the RDD sample and examined the distribution of their sociodemographic characteristics using unweighted and weighted RDD data. Also, we assessed modern contraceptive use in the following groups: FTF – full sample; FTF – phone owners; RDD (weighted and unweighted). We also constructed a five-category indicator of method mix, distinguishing the following contraceptive users: implant, IUD, injectables, pills, and condoms.

*Equivalence Test*

To examine the equivalence of modern contraceptive use prevalence estimates in the weighted RDD and FTF samples, we set the equivalence margin $\delta$ to +/- 4%. The null hypothesis assumed a difference of more than 4% between the two estimates of modern contraceptive use. Rejection of the null hypothesis would lead to the conclusion that weighted RDD and FTF estimates of modern contraceptive use were equivalent within a margin of 4%. We report the 90% confidence interval for the difference in point estimates, which simulates performing two one-sided tests. We also report a p-value from an adjusted Wald test. [111]

*Multivariable logistic regression*

We further compared modern contraceptive use by mode of data collection by conducting multivariable logistic regression adjusting for additional sociodemographic factors. We first assessed bi-variate relationships between each co-variate and modern contraceptive use among the pooled FTF-RDD data. We then conducted multivariable logistic regression to assess the odds of modern contraceptive use by survey mode, adjusting for covariates. We also compared the

RDD and FTF phone owner sample using multivariable logistic regression. We checked for co-linearity and found no variance inflation factor more than two. [98] Analysis was performed using weighted RDD and FTF data. We conducted analyses in Stata version 15 (StataCorp 2017) and determined statistical significance using an alpha of 0.05.

## *6.4 Results*

Call Outcomes

The cell phone survey was implemented by Viamo, which screened 202,295 unique phone numbers, of which 21% were deemed valid (Table 6-4). The 42,726 valid phone numbers were called by interviewers over the course of about a month and constituted our sample size for all survey response outcome analyses.

Overall, 45% of the 42,726 calls were categorized as ineligible, mostly due to a man answering the call (36%), and marginally to age (1%) or not speaking one of the seven survey languages (0.8%). Approximately 6% of calls were non-eligible because of quota restrictions (Table 6-5).

Another 49% of the 42,726 calls fell in the "unknown eligibility" category, mostly due to call no answer (43%), and marginally due to incomplete screening (5%), language mismatch with interviewer (i.e. respondent spoke one of seven languages but the interviewer did not speak the same language as the respondent and the respondent did not answer any subsequent calls) (1%) or answering machine (0.9%). Overall, less than 1% of eligible women did not complete the survey: total refusals before consent consisted of less than one percent of calls (0.09%). There were even fewer refusals during consent (0.01%) and only 0.07% of respondents were break-offs (consented but answered less than 50% of relevant questions).

Altogether, 2,379 women completed 50% or more of the survey questions and comprised our RDD study population for contraceptive prevalence analysis (n=54 were partial completers).

***Survey Rates***

The minimum response rate (Response rate 1), which includes only interviews where more than 80% of questions were answered and includes all attempted calls in the denominator was 9.9% (Table 6-6). This percentage rose to 10.1% when including partial survey completion (50-80%). When excluding 80% of unknown eligibility calls from the denominator, response rates increased substantially, to 68% when considering complete interviews (Response rate 3) and 70% when also counting partial interviews (Response rate 4).

The cooperation rate, which includes only eligible calls in the denominator, was 94.7% when counting complete interviews (Cooperation rate 1), and 97.1% when also counting partial interviews (Cooperation rate 2). The refusal rate was 3.1%.

<u>Characteristics of FTF & RDD samples</u>

***FTF Sample***

Among the 3,659 women who completed PMA2020 R5 FTF survey, 78% were rural, 72% were married and 75% had given birth (Table 6-7). The average age of the respondents was 28.6 years.

Close to two-thirds of women (63%) had never been to school; thus only one in five had a secondary education or higher. Moore was the most common language used for interviews (44%) while 10% of women completed the survey in French and 20% completed the interview in another language.

***FTF Sample – Phone owners only***

46% of women (N=1,671) in the FTF survey indicated that they owned a phone. Phone owners in the FTF survey were on average 28.9 years. The percent of women married (69%) and who had ever given birth (74%) were similar to the full FTF sample. A slightly greater percent of female phone owners lived in an urban area (35%) than the full FTF sample (20%). Two-thirds of FTF phone owners (65%) lived in rural settings and 30% had secondary education or higher. More than half of phone owners completed the survey in Moore (53%) and 19% completed the survey in French.

***RDD Sample***

The two characteristics used to create the quota groups shaped the unweighted RDD sample (N=2,379) as expected: the average age was 27.5 years old and 75% lived in rural areas. Three-quarters were married (74%), and had given birth at least once (76%). Half of women have never attended school and 32% had secondary education or higher. Two-thirds of women spoke Moore (65%) and a quarter spoke French (26%).

Compared to the target population of FTF respondents, women in the RDD sample were younger, more likely to have a secondary degree and more likely to respond in French. Characteristics of women in the RDD sample more closely reflected the FTF phone owner sample than the full FTF sample. The distribution of education and language was similar between the RDD sample the FTF phone owner sample, although women speaking French were over-represented in the RDD sample (26%) compared to the FTF phone owner sample (19%). A quarter of the RDD sample lived in the Center province (where Ouagadougou, the capital city is located) which was only 5% more than the FTF phone owner sample (19.5%) but 2.5 times the percent of respondents in the FTF survey.

After weighting the RDD sample to better reflect the attributes of the target sample of women of reproductive age in Burkina Faso (PMA2020 R5), the weighted RDD age distribution shifted to increased representation of women 40 years and older and increased the representation of women with lower education (63% never attending school). Weighting of the sample also increased the proportion of women who had ever given birth (80%). Distribution of language changed minimally from the unweighted to weighted sample. The percent of women who completed the survey in Moore increased five percent to 70% and deceased by 6% for French to 19% of the sample. There were no major changes in the distribution of women by province.

Contraceptive use by survey mode

***Contraceptive use among FTF, phone owner and RDD weighted samples***

A quarter of women (26% (95% CI: 22.7 – 29.6)) in the FTF survey reported contraceptive use (Table 6-8). This proportion increased to 31.7% (95% CI: 30.0% – 35.6%) among FTF phone owners. The proportion of women who reported contraceptive use was 40.2% (95% CI: 38.2% – 42.2%) in the unweighted RDD sample and remained mostly unchanged when applying RDD post-stratification weights (38.7% (95% CI: 36.7% – 40.8%).

***Method Mix among FTF, FTF Cell Phone Owners & RDD Sample***

Altogether 46% of contraceptive users in the FTF sample used a short-acting method (injectables, pills, or condoms) and 54% used long acting reversible contraception (LARC) (Figure 6-1). When restricting to contraceptive users who owned a cell phone in PMA2020 R5, the method mix shifted to 55% using a short-acting method and 45% using a LARC. Among RDD contraceptive users, 59% reported using short-acting methods and 41% used LARCs.

***Equivalence Test***

The difference of 12.7% in the prevalence of modern contraceptive use between weighted RDD and FTF full sample was greater than the 4% equivalence margin (90% CI of difference: 0.15 – 0.10). This led us to reject the null hypothesis (p-value <0.001) that the samples are similar and conclude that the weighted RDD modern contraceptive use prevalence was not equivalent to the FTF survey estimate.

***Multivariable Logistic Regression – odds of reporting modern contraceptive use by survey mode***

After adjusting for confounding covariates, results from the multivariable logistic regression model indicated that women in the RDD sample had almost twice the odds of reporting modern contraceptive use compared to women in the FTF sample (Odds Ratio (OR): 1.9, 95% Confidence Interval (95% CI): 1.6 – 2.1) (Table 6-9). Significant differences also remained when excluding non-cell phone owners from the FTF sample: women in the RDD sample had 60% higher odds (OR: 1.6 95% CI: 1.3 – 1.8) of reporting modern contraceptive use compared to FTF cell phone owners.

## *6.6 Discussion*

This study assessed the feasibility of conducting RDD phone surveys for monitoring family planning metrics in Burkina Faso and the associated bias in estimating national modern contraceptive prevalence. We found that six percent of valid phone numbers resulted in a complete interview. The RDD sample resulted in 14% point over-estimation of modern contraceptive use, which remained substantial and significant, after post-stratification adjustments of the RDD sample and further adjustment for confounding (OR: 1.9, 95% CI: 1.6 – 2.2).

As expected from previous studies on phone ownership [35, 96] and from our previous work comparing CATI respondents to non-respondents as seen in Aim 2, the RDD sample was distorted in favor of more educated women. In addition, the RDD sample over-represented women speaking Moore and living in Ouagadougou. Conversely, distortions by age and residential area were not apparent in our RDD study, due to the application of quotas, which excluded only six percent of eligible respondents.

The application of post-stratification weights to address educational, age and urban/rural distortions only reduced the difference by 1.5% point in comparison of contraceptive prevalence rates between RDD and FTF (from a 14.2% point difference using unweighted RDD estimates to 12.7% point difference using weighted RDD estimates). Furthermore, the differences between the FTF phone sample and weighted RDD sample's contraceptive prevalence rate estimates remained large (7% point difference), suggesting that variables other than three post-stratification weight variables caused modern contraceptive rate estimate distortion. A trio of articles from Brazil, all using data from VIGITEL, an annual and continuous phone survey in the 26 state capitals monitoring a host of non-communicable diseases, compared RDD samples with concurrent FTF samples. [9, 33, 69] Two of the articles used post-stratification weights and reported the weights reduced the difference between phone survey estimates and FTF survey estimates. [9, 33]

To better understand the reasons for the RDD over-estimation of modern contraceptive use in our study, we assessed the impact of frame bias by comparing RDD estimates to FTF phone owner estimates. The difference between the RDD unweighted sample and the FTF subsample of phone owners was significantly reduced (8.5% point difference) compared to the 14.2% difference observed with the full FTF sample suggesting frame bias may contribute for more than half of the over-estimation of modern contraceptive prevalence in the RDD sample. These results are in line with our Aim 1 results, showing greater use of contraception among cell phone owners compared to non-owners. They are also consistent with the findings of one of the aforementioned Brazilian studies showing little difference between RDD estimates and FTF estimates among phone owners. [33]

Nonetheless, the difference in estimates of contraceptive use between FTF phone owners and RDD respondents (unweighted) suggests additional effects of non-response or measurement bias. These two components may be interpreted as differences in population composition (non-response) versus differences in survey response by mode of data collection (measurement error). Based on Aim 2 results, non-response among cell phone owners led to sociodemographic distortion of the follow-up CATI survey sample, compared to the target population of FTF phone owners, but had no effect on contraceptive use estimates. The RDD sample showed greater representation of women with secondary education or higher, women under 34 and French- and Moore-speaking women compared to the FTF sample. However, after adjusting for these factors in the multivariable analysis, the odds of contraceptive use remained significantly higher in the unweighted RDD sample compared to the FTF phone owner sample (OR: 1.6 95% CI: 1.3 – 1.8).

Without a feasible non-response error explanation of over-report of modern contraceptive use among RDD respondents, we looked to measurement error. As the difference in RDD and FTF estimates varied by area of residence, we examined potential misclassification of area of residence due to the application of the quota design in the RDD survey. Specifically, we hypothesized that urban women, who were easier to identify in RDD screening than rural women, were intentionally mis-classified as rural by interviewers once the urban quota groups were filled, in order to speed up data collection. However, comparison of modern contraceptive use before and after the urban quota groups closed showed no increase in modern contraceptive use prevalence after the urban quota groups closed (data not shown). Results of the follow-up phone survey examined in Aim 2 however are worth noting. Among women who were classified as urban during the FTF survey, 16% changed to rural classification in the follow-up phone survey. Change in classification may explain some of the over-report of modern contraceptive use among rural women. Measurement error could also be caused by social desirability bias. The few studies addressing social desirability and data collection mode impact in LMIC have reported inconsistent results. [101] A study conducted in the Dominican Republic documented over-reporting of modern contraceptive use by rural women when responding to urban interviewers, in an effort to appear more modern. While social norms about contraception are likely different in Burkina Faso than in the Dominican Republic, social desirability may partly contribute to the over-reporting of modern methods, particularly since the RDD interviewers often disclosed they lived in Ouagadougou, the capital, and were highly educated, which can be detected through speech by the respondents. [150] Results from the follow-up phone survey examined in Aim 2 may support this hypothesis as reporting of modern contraceptive use in the follow-up phone survey was higher than the initial estimate based on the FTF survey collected 11 months earlier among the same women (45%, versus 35%, data not shown).

This study has a number of strengths. It is among the first in SSA to compare health estimates from concurrent surveys using different modes of data collection and as such, provides a firsthand investigation of opportunities and challenges of using phone surveys in a context of rapid demographic change. Use of similar questionnaires limited measurement error while the almost concurrent timing of the surveys also improved comparability of survey estimates. The sample sizes were large enough to allow equivalence testing of modern contraceptive use prevalence with a relatively low margin of equivalence of 4%.

The study also has a number of limitations, including the small number of demographic variables available in the RDD sample, which limited our capacity to systematically investigate differences between the RDD sample and the FTF target population. As a result, post-stratification weights were limited to a few demographics, leaving out potential unobserved differences that could better explain the difference between RDD and FTF modern contraceptive use estimates. Although we used multiple weights to correct for the difference, the modern contraceptive use estimate from the RDD survey remained higher than the estimate from the FTF survey. The study could not explain all the factors that may cause the differences in estimates between two surveys. Another limitation was not to have allowed men to pass the phone to a female respondent as 84% of households own a phone but fewer females own their own phone. Although the decision to not pass the phone was made to reduce the complexity of weighting (i.e. avoiding weighting the sample for women in a household that were not sampled), the volume of calls we had to place due to men answering the majority of calls (36.4% of calls picked up by men) made the project very challenging to implement and may have impacted data quality.

## 6.7 *Conclusion*

An RDD survey in Burkina Faso did not yield an estimate of modern contraceptive use that was equivalent to FTF reference estimate, even after applying post-stratification weights. Over-estimation of modern contraceptive use in our RDD survey originated both from a truncated sample frame, excluding non-cell phone owners, but also from non-response and measurement error, which need further examination as cell phone ownership expands in the SSA region.

## 6.8 Tables for Chapter 6

*Table 6-4. Total phone numbers called for IVR pre-notification and number validation*

| | | N | % |
|---|---|---|---|
| | In-valid phone number | 157,569 | 78.7 |
| | Valid phone number | 42,726 | 21.3 |
| | **Total** | 200,295 | |

*Table 6-5. Final disposition codes for RDD survey, among valid phone numbers*

| | | **N** *(42,726)* | **%** |
|---|---|---|---|
| | **Not Eligible** | | *(44.9%)* |
| | 4.7 No eligible respondent | | |
| | 4.71 Gender (male) | 15,570 | 36.4 |
| | 4.72 Age (<15 or >49) | 479 | 1.1 |
| | 4.73 Language (not one of the 7 languages available in survey) | 326 | 0.8 |
| | 4.8 Quota Filled | 2,812 | 6.6 |
| | **Unknown Eligibility, Non-Interview** | | *(49.3%)* |
| UH | 3.13 No Answer | 18,182 | 42.5 |
| UH | 3.14 Telephone answering device | 370 | 0.9 |
| UH | 3.21 No screener completed – talked with respondent but hung-up or refused | 1,984 | 4.6 |
| UO | 3.90 Other (Language not matched with interviewer ) | 549 | 1.3 |
| | **Eligible, Non-Interview** | | *(0.17%)* |
| R | 2.111 Refusal pre-consent | 37 | 0.09 |
| R | 2.11 Refusal at consent | 6 | 0.01 |
| R | 2.10 Break-off (consented but less than 50% of relevant questions answered) | 32 | 0.07 |
| | **Interview** | | *(5.6%)* |
| P | 1.2 Partial (50-80% of relevant questions answered) | 54 | 0.1 |
| I | 1.1 Complete (more than 80% of relevant questions answered) | 2,325 | 5.4 |

*Table 6-6. Call Outcome Rates for RDD survey based on Final Disposition Distributions*

| Response Rates | Explanation | Result |
|---|---|---|
| Response rate 1: $\frac{I}{I+P+R+NC+O+UH+UO}$ | *Minimum response rate.*<br>All individuals who complete more than 80% of survey /<br>All eligible individuals | 9.9% |
| Response rate 2: $\frac{I+P}{I+P+R+NC+O+UH+UO}$ | All individuals who complete more than 50% of survey /<br>All eligible individuals | 10.1% |
| Response rate 3: $\frac{I}{I+P+R+NC+O+e(UH+UO)}$ | *Estimates what proportion of cases of unknown eligibility are actually eligible.*<br>All individuals who complete more than 80% of survey /<br>All eligible individuals | 68.0% |
| Response rate 4: $\frac{I+P}{I+P+R+NC+O+e(UH+UO)}$ | Estimates what proportion of cases of unknown eligibility are actually eligible.<br>All individuals who complete more than 50% of survey /<br>All eligible individuals | 70.4% |
| **Cooperation Rates** | | |
| Cooperation rate 1 : $\frac{I}{I+P+R+O}$ | All individuals who complete more than 80% of survey /<br>Eligible individuals who were ever contacted | 94.7% |
| Cooperation rate 2: $\frac{I+P}{I+P+R+O}$ | All individuals who complete more than 50% of survey /<br>Eligible individuals who were ever contacted | 97.1% |
| **Refusal Rate** | | |
| Refusal rate 3: $\frac{R}{I+P+R+NC+O}$ | All individuals who refused to complete the survey /<br>All eligible individuals | 3.1% |
| **Contact Rate** | | |
| Contact rate 1: $\frac{I+P+R+O}{I+P+R+NC+O+UH+UO}$ | All phone numbers that answered the call /<br>All phone numbers | 10.4% |
| Contact rate 2: : $\frac{I+P+R+O}{I+P+R+NC+O+e(UH+UO)}$ | Same as Contact rate #1 but only estimated eligible cases are included in the denominator as undetermined cases. | 74.9% |

*Table 6-7. Characteristics of women by survey mode and cell phone ownership*

| | FTF All Respondents | | FTF Phone Owners | | RDD Unweighted | | RDD Weighted* | |
|---|---|---|---|---|---|---|---|---|
| Variable | N = 3,659 | % | N= 1,671 | % | N = 2,379 | % | N= 2,379 | % |
| **Age** | | | | | | | | |
| Mean (standard error) | 28.6 (0.24) | | 28.9 (0.26) | | 27.5 (0.18) | | 28.7 (0.21) | |
| 15-19 | 820 | 22.4 | 316 | 18.9 | 519 | 22.0 | 502 | 21.3 |
| 20-24 | 622 | 17.0 | 311 | 18.6 | 476 | 20.2 | 389 | 16.5 |
| 25-29 | 622 | 17.0 | 316 | 18.9 | 419 | 17.8 | 387 | 16.4 |
| 30-34 | 527 | 14.4 | 246 | 14.7 | 375 | 15.9 | 347 | 14.7 |
| 35-39 | 428 | 11.7 | 204 | 12.2 | 229 | 9.7 | 288 | 12.2 |
| 40-44 | 388 | 10.6 | 172 | 10.7 | 203 | 8.6 | 269 | 11.4 |
| 45-49 | 252 | 6.9 | 100 | 6 | 138 | 5.9 | 177 | 7.5 |
| | | | | | | | | |
| **Urban/rural** | | | | | | | | |
| Rural | 2,869 | 78.4 | 1079 | 64.6 | 1,776 | 75.3 | 1,769 | 75.0 |
| Urban | 790 | 21.6 | 592 | 35.4 | 583 | 24.7 | 590 | 25.0 |
| | | | | | | | | |
| **Marital status** | | | | | | | | |
| Currently not in union | 1,025 | 28.0 | 521 | 31.5 | 601 | 25.6 | 514 | 21.8 |
| Currently in union | 2,634 | 72.0 | 1145 | 68.5 | 1,748 | 74.4 | 1,845 | 78.2 |
| | | | | | | | | |
| **Highest school attended** | | | | | | | | |
| Never | 2,334 | 63.8 | 869 | 52 | 1,210 | 51.3 | 1,484 | 62.9 |
| Primary | 593 | 16.2 | 296 | 17.7 | 387 | 16.4 | 392 | 16.6 |
| Secondary or higher | 732 | 20.0 | 508 | 30.3 | 762 | 32.3 | 484 | 20.5 |
| | | | | | | | | |
| **Parity** | | | | | | | | |
| Avg # of kids among parous women | 3.0 | | 2.7 | | 3.2 | | 3.6 | |
| No | 918 | 25.1 | 434 | 26 | 573 | 24.4 | 488 | 20.7 |
| Yes | 2,741 | 74.9 | 1237 | 74 | 1,774 | 75.6 | 1,871 | 79.3 |
| | | | | | | | | |
| **Language** | | | | | | | | |
| Dioula | 388 | 10.2 | 373 | 11.1 | 169 | 7.2 | 193 | 8.2 |
| French | 1,603 | 10.6 | 388 | 18.5 | 600 | 25.6 | 455 | 19.3 |
| Fulfulde | 377 | 4.9 | 179 | 0.67 | 27 | 1.15 | 31 | 1.3 |
| Gourmantchema | 179 | 10.3 | 377 | 6.2 | 25 | 1.1 | 28 | 1.2 |

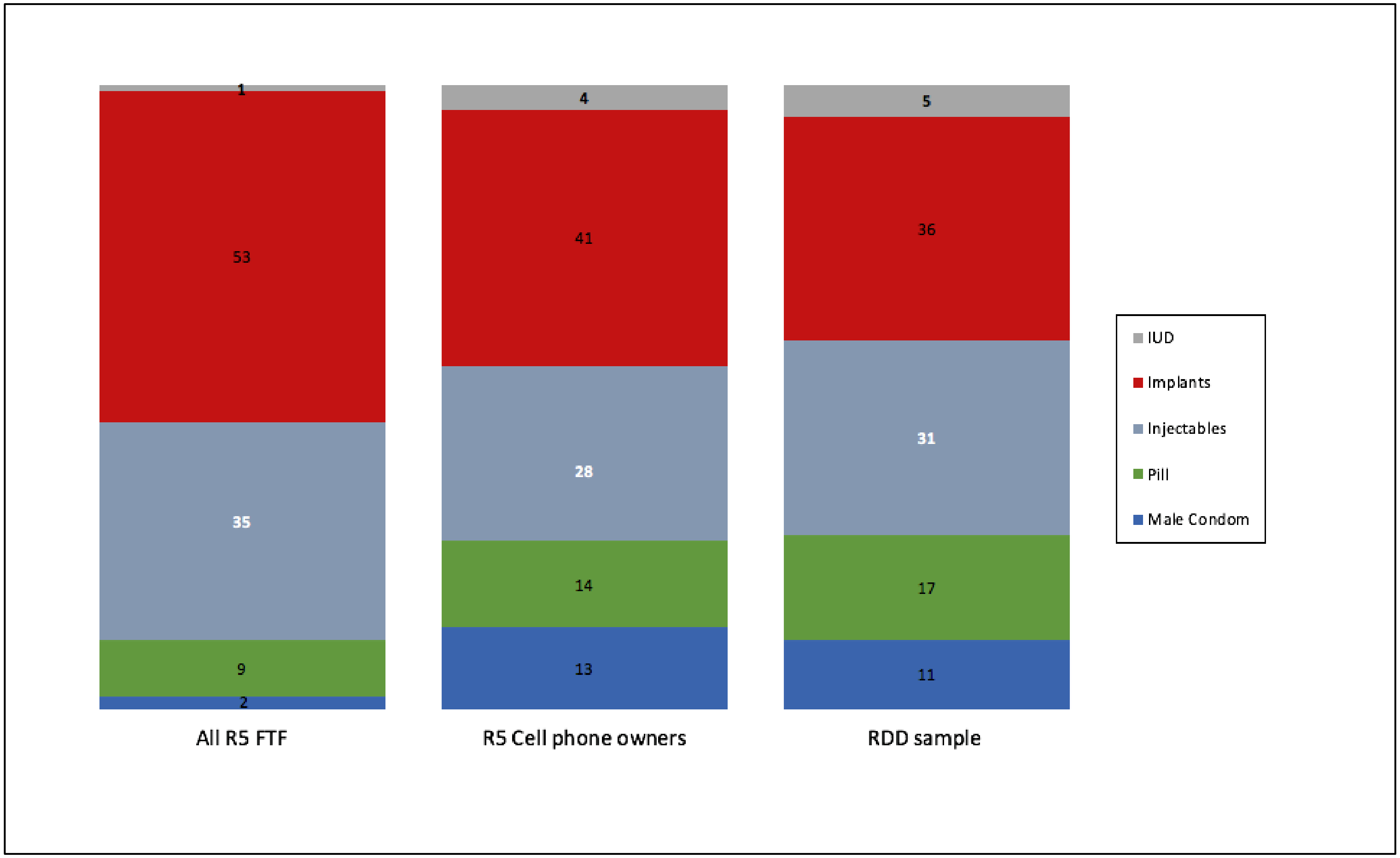
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Moore | 373 | 43.8 | 1602 | 53.3 | 1,521 | 64.9 | 1,651 | 70.0 |
| Other | 743 | 20.3 | 169 | 10.1 | - | - | - | - |
| | | | | | | | | |
| **Province** | | | | | | | | |
| Boucle du Mouhoun | 410 | 11.2 | 130 | 7.8 | 134 | 5.8 | 134 | 5.7 |
| Cascades | 168 | 4.6 | 42 | 2.5 | 73 | 3.2 | 73 | 3.1 |
| Centre | 395 | 10.8 | 326 | 19.5 | 574 | 25.0 | 578 | 24.5 |
| Centre-Est | 267 | 7.3 | 125 | 7.5 | 206 | 9.0 | 217 | 9.2 |
| Centre-Nord | 351 | 9.6 | 155 | 9.3 | 240 | 10.4 | 255 | 10.8 |
| Centre-Ouest | 417 | 11.4 | 207 | 12.4 | 170 | 7.4 | 172 | 7.3 |
| Centre- Sud | 99 | 2.7 | 38 | 2.3 | 99 | 4.3 | 104 | 4.4 |
| Est | 424 | 11.6 | 137 | 8.2 | 131 | 5.7 | 130 | 5.5 |
| Hauts-Bassins | 333 | 9.1 | 204 | 12.2 | 206 | 9.0 | 210 | 8.9 |
| Nord | 304 | 8.3 | 170 | 10.2 | 177 | 7.7 | 189 | 8.0 |
| Plateau-Central | 124 | 3.4 | 75 | 4.5 | 158 | 6.9 | 170 | 7.2 |
| Sahel | 238 | 6.5 | 30 | 1.8 | 78 | 3.4 | 78 | 3.3 |
| Sud-Ouest | 1 | 0.04 | 30 | 1.8 | 33 | 2.31 | 50 | 2.1 |

*Note: FTF analyses are weighted for survey design-weight.*
**post-stratification weights included age, residence and level of education*

*Table 6-8. Prevalence of Modern Contraceptive use by survey mode*

| | **FTF** *Full sample* | **FTF** *Phone owners* | **RDD** *Unweighted* | **RDD** *Weighted (age, residence, education)* |
|---|---|---|---|---|
| Prevalence of Modern contraceptive use | 26.0 (22.7 – 29.6) | 31.7 (30.0 – 35.6) | 40.23 (38.2 – 42.2) | 38.7 (36.7- 40.8) |

*Figure 6-1. Modern method mix among current users among full FTF sample, FTF cell phone owners, and RDD respondents (%)*




*The most effective method currently used, if multiple methods were reported

**Note:** *% estimates are adjusted for sampling weight*

*Table 6-9. Odds of reporting modern contraceptive use by women's characteristics and survey mode, using full FTF sample and FTF phone owner sample*

| | **RDD vs. Full FTF sample** | | **RDD vs. FTF phone owner sample** |
|---|---|---|---|
| | **Unadjusted OR (95% CI)** | **Adjusted OR (95% CI)** | **Adjusted OR (95% CI)** |
| **Mode** | | | |
| FTF *(reference)* | | | |
| RDD | **1.8 (1.6 – 2.0 )** | **1.9 (1.6 – 2.2)** | **1.6 (1.3 – 1.8)** |
| **Age group** | | | |
| 15-19 *(reference)* | | | |
| 20-24 | **2.4 (2.0 – 3.0)** | **1.4 (1.1 – 1.8)** | **1.4 (1.1 – 1.8)** |
| 25-29 | **3.0 (2.5 – 3.8)** | **1.6 (1.2 – 2.1)** | **1.6 (1.3 – 2.1)** |
| 30-35 | **3.4 (2.7 – 4.2)** | **1.7 (1.3 – 2.3)** | **1.6 (1.2 – 2.1)** |
| 35-39 | **3.3 (2.6 – 4.2)** | **1.6 (1.2 – 2.2)** | **1.7 (1.2 – 2.3)** |
| 40-44 | **2.2 (1.7 – 2.8)** | 1.1 (0.8 – 1.5) | 1.1 (0.8 – 1.6) |
| 45-49 | 1.0 (0.7 – 1.4) | 0.5 (0.4 – 0.7) | 0.5 (0.4 – 0.8) |
| **Residential area** | | | |
| Rural *(reference)* | | | |
| Urban | **1.3 (1.1 – 1.4)** | 1.0 (0.9 – 1.2) | 0.9 (0.8 – 1.2) |
| **Highest school attended** | | | |
| No education *(reference)* | | | |
| Primary | **1.5 (1.3 – 1.8)** | **1.7 (1.5 – 2.1)** | **1.7 (1.4 – 2.1)** |
| Secondary or more | **1.3 (1.1 – 1.5)** | **1.9 (1.5 – 2.3)** | **1.6 (1.3 – 2.0)** |
| **Survey Language** | | | |
| French *(reference)* | | | |
| Moore | **0.6 (0.6 -0.7)** | **0.6 (0.5 – 0.8)** | **0.6 (0.5 – 0.8)** |
| Gourma | **0.7 (0.5 – 0.9)** | 0.9 (0.7 – 1.3) | 1.0 (0.7 – 1.6) |
| Fulfulde | **0.2 (0.1 – 0.3)** | **0.3 (0.2 – 0.5)** | 0.5 (0.3 – 1.3) |
| Dioula | 0.8 (0.6 – 1.0) | 0.8 (0.7 – 1.1) | 0.9 (0.7 – 1.3) |
| Other | **0.6 (0.5 – 0.8)** | 0.9 (0.7 – 1.3) | 0.9 (0.6 – 1.5) |
| **Ever Birth** | | | |
| Never given birth *(reference)* | | | |
| Ever given birth | **3.7 (3.2 - 4.4)** | **4.8 (3.8 – 6.0)** | **3.7 (3.0 – 4.6)** |
| **Ever married** | | | |
| Not married (reference) | | | |
| Married | **1.7 (1.5 – 2.0)** | 0.9 (0.7 – 1.0) | **0.8 (0.6 – 0.9)** |

# 7. Conclusions

## *7.1 Summary of results*

This dissertation used a sequential design to evaluate opportunities and challenges of transitioning from traditional FTF probability sample surveys to phone surveys to monitor family planning indicators in a country experiencing rapid change in reproductive health indicators. Each study aim built on the information ascertained in the previous aim, informing the contribution of different sources of survey errors on estimates of modern contraceptive use. Specifically, the first two aims examined frame error and non-response error and their implications for estimates of contraceptive use based on a representative population of women who owned a phone, while the third aim evaluated phone versus FTF contraceptive estimates, using different sampling strategies (RDD for the phone survey and cluster probability sampling for FTF survey).

In addressing Aim 1, we identified major sociodemographic differences between female cell phone owners and non-owners among a representative sample of women of reproductive age in Burkina Faso, showing the demographic make-up of cell phone owners was skewed towards urban and educated women. The sample distortion of cell phone owners led to significant over-estimation of modern contraceptive use. As cell phone ownership increases in this region, we anticipate phone sample representation will improve thus reducing bias, [48] such as experienced in developed country settings in the 1990s. [73]

The second aim followed-up female phone owners identified in Aim 1 to estimate levels and correlates of non-response and identify the phone data collection mode (CATI or Hybrid IVR) that had less sample distortion. CATI respondents were more representative of the target FTF population of cell phone owners than Hybrid IVR respondents, but CATI respondents and the target FTF population were still noticeably different. Interestingly, non-response to both CATI and Hybrid IVR had little impact on estimates of contraceptive use. Given the lower response rate and greater distortion associated with Hybrid IVR, we concluded that CATI was a better choice for cell phone surveys in contexts where literacy rates are low (30%), as in Burkina Faso.

Finally, in Aim 3, we found that the RDD estimate of modern contraceptive use was significantly higher (40%) than the FTF reference estimate (26%) and was not reduced by applying post-stratification weights (39%) to account for age, education and area of residence distortion. These results differ from recent cell phone studies conducted in other LMIC where post-stratification weighting was generally successful in aligning phone survey estimates with reference FTF survey estimates. [9, 33, 71, 72, 151] These studies included men who are more likely to own phones than women, perhaps resulting in a phone owner population that is more representative of the underlying population. [35, 96] Over-estimation of modern contraceptive estimates in the RDD sample is likely related to a combination of frame, non-response and measurement error. More research is needed to tease out these effects in SSA contexts. [74]

Taken together, our results show that phone surveys among women introduces significant bias of modern contraceptive use. The distortion is primarily due to low cell phone diffusion among the female population in Burkina Faso (frame error), which should raise caution in replacing FTF probability surveys with phone surveys to monitor FP metrics. We encourage tracking cell phone ownership in FTF probability surveys to detect when frame bias is reduced as cell phone ownership reaches a larger population of women of reproductive age.

### *7.2 Strengths and limitations*

This multi-phased study provided a holistic perspective on the feasibility of conducting cell phone surveys in a low-resource setting where cell phone ownership is not yet universal. The design of the study, based on a representative sample of women of reproductive age is a strength of the study. Previous research compared cell phone respondents to a reference population from DHS surveys or census data, making it difficult to pinpoint the error source. [70-72] However, our study design allowed dissecting the cause of error, from frame error, to non-response error and measurement error, which most previous research did not. Specifically, the use of baseline PMA2020 data provided identical measures for cell phone owners and non-owners, and for phone survey respondents and non-respondents. This information was not only used to illustrate sample distortion at different stages of sample selection (cell phone ownership and non-response), but also to evaluate the implications of sample distortion on estimates of the outcome of interest.

Another strength is the direct comparison of Hybrid IVR and CATI. Known published studies comparing two or more remote modes of data collection were conducted in South America [32] (Peru and Honduras) and India. [152] By comparing the profile of respondents and key survey outcomes by mode, we identified CATI as the most appropriate data collection mode in this low resource setting, where levels of literacy are relatively low.

The target population, women, is a strength. The lack of research on innovative strategies to collect rapid turn-around, low cost estimates on women's health and well-being, including family planning metrics, is a limitation for tracking progress towards achieving the Sustainable Development Goals, including access to universal sexual health services and women and girl's empowerment. [26] The diffusion of cell phones in SSA provides an opportunity to address this knowledge gap, although women are less likely to benefit from cell phone expansion than men in this context.

Although this dissertation has many strengths, there are limitations. The first limitation is the short questionnaire length in CATI compared to the FTF survey, resulting in limited information collected during the cell phone surveys. The lack of sociodemographic information limited the scope of post-stratification weighting, which only accounted for age, education and area of residence, although sample size also constrains post-stratification to a limited number of variables.

Although we were interested in fielding cell phone surveys in low-resources settings, Burkina Faso presented a particularly challenging environment, which has implications for the generalizability of the study. First, the infrastructure of the country hindered the RDD study in particular, with frequent electricity and internet cuts and difficulty with mobile network operators, making implementation of the study quite challenging. In other West African countries with stronger [153] and better functioning mobile network operators, phone surveys using RDD would be more feasible. Second, the low literacy and high language fractionalization among women in Burkina Faso increases non-response.

### *7.3 Implications*

*Policy*

Governments track fertility to anticipate population dynamics and allocate resources accordingly. The ability to use data for real-time decision making is hindered by infrequent collection of population-based estimates, among other factors. Although PMA2020 has addressed this problem by providing a cost- effective platform for modern contraceptive use data collection, their reach is limited to eleven countries. The conclusions from this dissertation can be used to gauge when a country could feasibly use cell phone surveys to track SRH indicators. Specifically, this dissertation illustrates the feasibility and limitations of cell phone surveys in a West African setting. Although we would not suggest supplementing FTF surveys in West Africa with RDD phone surveys, mainly due to low phone ownership, (frame bias), follow-up phone surveys of FTF samples could be worthwhile for research questions that benefit from monthly or bi-monthly data collection, which is not the case for many SRH indicators.

*Research*

Several aspects of this dissertation contribute to the literature on collecting data via phone calls in LMICs. Our results establish more specific research areas for cell phone surveys in SSA. The results in Aim 1 showed that phone ownership among women of reproductive age was not prevalent enough to generate valid estimates of modern contraceptive use in Burkina Faso. We suggest a similar study in a country with higher levels of female phone ownership and higher literacy to identify a threshold of cell phone diffusion in the population which would reduce sample distortion and produce more valid estimates of SRH outcomes. Future studies should also try to collect more data on RDD respondents, to better identify differences between RDD and FTF samples and to use this information for post-stratification weights. Future research should also be sure to stratify age groups by gender: female respondents in both the follow-up and RDD study were more likely to be 25- 34 than ages 15-24 whereas 56% of male and female respondents in a recent RDD survey in Ghana were ages 15-24. [72] However, two-thirds of the Ghana RDD sample was male. The contrast in age distribution by gender underscores the need for researchers to dis-aggregate their population by gender, especially when reporting age. Finally, future follow-up phone surveys that want to reduce attrition would need to take place more quickly after baseline. [154]

The RDD study illustrates the difficulty of randomly calling women in settings with low cell phone ownership. Further research should explore less time consuming ways to survey women via RDD in SSA, such as having a man pass the phone to a female household member or experimenting with other approaches such as Redirected Inbound Call Sampling (RICS), a non-probability sampling approach that re-routes misdialed phone calls to a survey; an approach that has recently gained popularity in the United States. [74] Not only was data collection logistically taxing, the RDD results were not comparable to the FTF results, even after weighting. The inability to successfully weight the RDD data to the FTF data indicates there are currently unmeasurable factors causing a difference in modern contraceptive use between the two samples. To better assess measurement error, the follow-up study design could be replicated by conducting the phone follow-up sooner in order to assess the reliability of SRH responses and to assess the factors contributing to differences in response.

## 8. References

[1] Model-based Estimates and Projections of Family Planning Indicators 2018. New York United Nations; 2018.

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

[3] Bongaarts J. Modeling the fertility impact of the proximate determinants: Time for a tune-up. Demographic Research. 2015;33:25.

[4] Ahmed S, Li Q, Liu L, Tsui AO. Maternal deaths averted by contraceptive use: an analysis of 172 countries. The Lancet. 2012;380:111-25.

[5] Cleland J, Bernstein S, Ezeh A, Faundes A, Glasier A, Innis J. Family planning: the unfinished agenda. The Lancet. 2006;368:1810-27.

[6] Bloom DE, Kuhn M, Prettner K. Africa's Prospects for Enjoying a Demographic Dividend. Journal of Demographic Economics. 2017;83:63-76.

[7] ADDING IT UP: Investing in Contraception and Maternal and Newborn Health, 2017. New York, NY: Guttmacher Institute 2017. p. 4.

[8] Family Planning 2020. 2016.

[9] 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.

[10] Biemer P. Overview of Design Issues: Total Survey Error. In: Marsden PV WJ, editor. Handbook of Survey Research. 2 ed. Bingley, United Kingdom: Emerald Group; 2010. p. 886.

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

[12] 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.

[13] Total Survey Error in Practice. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2017.

[14] Stover J, Ross J. How increased contraceptive use has reduced maternal mortality. Matern Child Health J. 2010;14:687-95.

[15] Bongaarts J. Trends in unwanted childbearing in the developing world. Stud Fam Plann. 1997;28:267-77.

[16] Speidel J, Sindig S, Gillespie D, Maguire E, Neuse M. Making the Case for U.S. International Family Planning Assistance 2009. p. 20.

[17] Global Reproductive Health Indicators. In: Evaluation M, editor. Global Reproductive HEalth Indicators Rockville, MD.

[18] Bryant J. Theories of Fertility Decline and the Evidence from Development Indicators. Population and Development Review. 2007;33:27.

[19] Institute IS. The World Fertility Survey: An International Programme of Fertility Research. International Statistical Review / Revue Internationale de Statistique. 1973;41:12.

[20] Jain A. Measuring the Effect of Fertility Decline on the Maternal Mortality Ratio. Studies in Family Planning.42:14.

[21] Soares RR. On the Determinants of Mortality Reductions in the Developing World. Population and Development Review. 2007;33:40.

[22] Bongaarts J, Casterline J. Fertility Transition: Is sub-Saharan African Different? Population and Development Review. 2013;38:16.

[23] Reher DS. Economic and Social Implications of the Demographic Transition. Population and Development Review. 2011;37:11-33.

[24] Family Planning: Francophone West Africa on the Move - A Call to Action. Washington, DC: Population Reference Bureau; 2012. p. 23.

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

[26] Sustainable Development Goals. Goal 3: Ensure healthy lives and promote well-being for all at all ages. In: Nations U, editor.2016.

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

[28] About PMA2020. In: PMA2020, editor. Baltimore, MD 2016.

[29] Hawes M, Safi S, Greenleaf A, et al. Response patterns on behavioral outcomes in relation to use of resident enumerators over multiple survey rounds. Performance Monitoring and Accountability 2020 (PMA2020) Methodological Reports. Baltimore, Maryand, USA: Bill & Melinda Gates Institute for Population and Reproductive Health, Johns Hopkins University; 2017.

[30] Gibson D, Pereira A, Farrenkopf B, Pariyo G, Labrique A, Hyder A. A systematic review of the literature for population representative mobile phone surveys in LMICs. Journal of Medical Internet Research. 2017.

[31] Hazel E, Amouzou A, Park L, et al. Real-time assessments of the strength of program implementation for community case management of childhood illness: validation of a mobile phone-based method in Malawi. Am J Trop Med Hyg. 2015;92:660-5.

[32] 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.

[33] Ferreira A, Cesar C, de Souza AD, Proietti C, Bernal F, Caiaffa R. Validity of data collected by telephone survey: a comparison of VIGITEL 2008 and the 'Saúde em Beagá' survey. Rev Bras Epidemiol. 2011;14:16-30.

[34] Gibson DG, Pariyo GW, Wosu AC, et al. Evaluation of Mechanisms to Improve Performance of Mobile Phone Surveys in Low- and Middle-Income Countries: Research Protocol. JMIR Res Protoc. 2017;6:e81.

[35] Poushter J, Oates R. Cell Phones in Africa: Communication Lifeline. Washington, DC: Pew Research Center; 2015.

[36] Sub-Saharan Africa: Ericsson Mobility Report. Stockholm, Sweden: Ericsson.

[37] 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.

[38] Measuring mobile penetration. 5 New Street Square, New Fetter Lane, London EC4A 3BF: GSMA; 2014. p. 3.

[39] The Mobile Economy: Sub-Saharan Africa 2015. The Walbrook Building 25 Walbrook London EC4N 8AF: GSMA 2015. p. 60.

[40] Abaza H, Marschollek M. mHealth Application Areas and Technology Combinations*. A Comparison of Literature from High and Low/Middle Income Countries. Methods Inf Med. 2017;56:e105-e22.

[41] 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.

[42] Brinkel J, Kramer A, Krumkamp R, May J, Fobil J. Mobile phone-based mHealth approaches for public health surveillance in sub-Saharan Africa: a systematic review. Int J Environ Res Public Health. 2014;11:11559-82.

[43] Gentry S, van-Velthoven MH, Tudor Car L, Car J. Telephone delivered interventions for reducing morbidity and mortality in people with HIV infection. Cochrane Database Syst Rev. 2013:CD009189.

[44] Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. J Health Commun. 2012;17 Suppl 1:82-104.

[45] Du X, Wang W, van Velthoven HM, et al. mHealth Series: Text messaging data collection of infant and young child feeding practice in rural China - a feasibility study. J Glob Health. 2013;3:020403.

[46] 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.

[47] Hox JJ, de Leeus ED, Dillman DA. The cornerstones of survey research. In: de Leeuw EH, J; Dillman, D editor. International Handbook of Survey Methodology. New York: L'Erlbaum Associates; 2008. p. 549.

[48] Labrique A, Blynn E, Ahmed S, Gibson D, Pariyo G, Hyder AA. Health Surveys Using Mobile Phones in Developing Countries: Automated Active Strata Monitoring and Other Statistical Considerations for Improving Precision and Reducing Biases. J Med Internet Res. 2017;19:e121.

[49] Mulry MH. Coverage Error. In: Lavrakas PJ, editor. Encyclopedia of Survey Research Methods. Thousand Oaks: Sage Publications, Inc.; 2011. p. 162-6.

[50] Johnson TP, Wislar JS. Response Rates and Nonresponse Errors in Surveys. Journal of American Medical Association. 2012;307:2.

[51] Groves RM. Nonresponse Rates and Nonresponse Bias in Household Surveys. Public Opinion Quarterly. 2006;70:20.

[52] Sudman S, Bradburn N, Schwarz N. Thinking about answers: The application of cognitive processes to survey methodology. . 1 ed. New York, NY1996.

[53] Carley-Baxter LR. Respondent-Interviewer Rapport. In: Lavrakas PJ, editor. Encyclopedia of Survey Research Methods. Thousand Oaks, CA: Sage Publications, Inc.; 2011.

[54] Biemer PP. Total Survey Error: Design, Implementation, and Evaluation. Public Opinion Quarterly. 2011;74:817-48.

[55] Cronbach LJ, Meehl PE. Construct Validity in Psychological Tests. Psychol Bulletin. 1955;52:12.

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

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

[58] de Leeuw ED. Mode effects in survey research: a comparison of mail, telephone and face. BMS: Bulletin of Sociological Methodology. 1993;41:17.

[59] Randall S, Coast E, Compaore N, Antoine P. The power of the interviewer. Demographic Research. 2013;28:763-92.

[60] Kreuter F, Preser S, Tourangeau R. Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity. The Public Opinion Quarterly. 2008;72:847-65.

[61] Tourangeau R, T.W. S. Asking Sensitive Questions: The Imapct of Data Collection Mode, Question Format, and Question Context. Public Opinion Quarterly. 1996;60:30.

[62] van de Wijgert J, Padian N, Shiboski S, Turner C. Is audio computer-assisted self-interviewing a feasible method of surveying in Zimbabwe? International Journal of Epidemiology. 2000;29:6.

[63] Midanik LT, Greenfield TK. Interactive voice response versus computer-assisted telephone interviewing (CATI) surveys and sensitive questions: the 2005 National Alcohol Survey. J Stud Alcohol Drugs. 2008;69:580-8.

[64] Midanik LT, Greenfield TK. Reports of alcohol-related problems and alcohol dependence for demographic subgroups using interactive voice response versus telephone surveys: the 2005 US National Alcohol Survey. Drug and alcohol review. 2010;29:392-8.

[65] Corkrey R, Parkinson L. A comparison of four computer-based telephone interviewing methods: getting answers to sensitive questions. Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc. 2002;34:354-63.

[66] Brooks MJ, Bear T, Hacker K, et al. An Alternative Technique for Youth Risk Surveillance Outside of the School System. J Adolesc Health. 2017;60:72-8.

[67] Kepple NJ, Freisthler B, Johnson-Motoyama M. Bias in child maltreatment self-reports using interactive voice response (IVR). Child Abuse & Neglect. 2014;38:1694-705.

[68] Lynn P, Kaminska O. The Impact of Mobile Phones on Survey Measurement Error. Public Opinion Quarterly. 2012;77:586-605.

[69] Francisco PMSB, de Azevedo Barros MB, Segri NJ, Porto Alves MCG, Cesar CLG, Malta DC. Comparison of estimates for the self-reported chronic conditions among household survey and telephone survey – Campinas (SP), Brazil. Rev Bras Epidemiol. 2011;14:5-15.

[70] Leo B, Morello R, Mellon J, Peixoto T, Davenport S. Do Mobile Surveys Work in Poor Countries. CGD Working Paper 398. Washington, DC: Center for Global Development 2014.

[71] Larmarange J, Kassoum O, Kakou E, Fradier Y, Lazare S. Feasibility and Representativeness of a Random Sample Mobile Phone Survey in Côte d'Ivoire. Population. 2016;71:10.

[72] L'Engle K, Sefa E, Adimazoya EA, et al. Survey research with a random digit dial national mobile phone sample in Ghana: Methods and sample quality. PLoS One. 2018;13:e0190902.

[73] Lavrakas PJ, Blumberg S, Battaglia M, et al. New Considerations for Survey Researchers When Planning and Conducting RDD Telephone Surveys in the U.S. With Respondents Reached via Cell Phone Numbers. In: Report CPTF, editor.: American Association for Public Opinion Research; 2010. p. 136.

[74] Lavrakas PJ, Benson G, Blumberg S, et al. The Future of the U.S. General Population Telephone Survey Research. In: Research AAfPO, editor.2017. p. 74.

[75] Khout A, Keeter S, Doherty C, Dimcock M, Christian L. Assessing the Representativeness of Public Opinion Surveys. Washington, DC.: Pew Research Center; 2012. p. 52.

[76] Richard J-B, Andler R, Gautier A, Guignard R, Leon C, Beck F. Effects of using an Overlapping Dual-Frame Design on Estimates of Health Behaviors: A French General Population Telephone Survey. Journal of Survey Statistics and Methodology. 2016.

[77] Slavec A, Toninelli D. An Overview of Mobile CATI Issues in Europe. In: Toninelli D PR, de Pedraza, P, editor. Mobile Research Methods: Opportunities and challenges of mobile research methodologies: Ubiquiety Press; 2015. p. 168.

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

[79] 2016 World Population Data Sheet. Washington, DC: Population Reference Bureau; 2016.

[80] RECENSEMENT GENERAL DE LA POPULATION ET DE L'HABITATION DE 2006. In: FINANCES MDLEED, editor. Ouagadougou, Burkina Faso2008. p. 52.

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

[82] Plan National d'Accélération de Planification Familiale du Burkina Faso 2017-2020. Ouagadougou, Burkina Faso: Ministère de la Santé; 2017. p. 108.

[83] 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.

[84] PMA2020 - Burkina Faso - Round 4. In: PMA2020, editor. Key Indicator Briefs - Family Planning. Baltimore, MD: Johns Hopkins University 2017. p. 2.

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

[87] Enquête sur les Indicateurs du Paludisme (EIPBF) Rockville, Maryland: Institut National de la Statistique et de la Démographie, Programme National de Lutte contre le Paludisme, ICF International; 2014. p. 170.

[88] Survey Methodology. In: PMA2020, editor. Baltimore, MD 2018.

[89] 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.

[90] DHS Model Questionnaires. In: ICF, editor. Gaither Road, Suite 500, Rockville, MD2018.

[91] Zimmerman L, Ahmed S. PMA2020 Weight Creation Memo - v4. In: PMA2020, editor. Baltimore, MD2017.

[92] Family planning/Contraception. In: Organization WH, editor. Fact Sheet Geneva, Switzerland2017.

[93] Long- Acting Reversible Contraceptives (LARCs). In: k4Health, editor.2018.

[94] Khatun F, Heywood AE, Hanifi SMA, et al. Gender differentials in readiness and use of mHealth services in a rural area of Bangladesh. BMC Health Serv Res. 2017;17:573.

[95] Velthausz D, Donco R, Skelly H, Eichleay M. Mozambique Mobile Access and Usage Study: Computer-Assisted Telephone Interview (CATI) Survey Results. In: FHI360, editor.2016. p. 66.

[96] GSMA. Bridging the gender gap: Mobile access and usage in low and middle-income countries. In: GSMA, editor. Connected Women2015. p. 130.

[97] Fry K, Firestone R, Chakraborty N. Measuring Equity with Nationally Representative Wealth Quintiles. In: PSI, editor. Washington, DC: PSI; 2014.

[98] Salmerón R, García CB, García J. Variance Inflation Factor and Condition Number in multiple linear regression. Journal of Statistical Computation and Simulation. 2018;88:2365-84.

[99] Method mix. In: Evaluation M, editor. Family Planning and Reproductive Health Indicators Database2018.

[100] Stone H. Getting your program out there: how to recruit rural women. Kumasi, Ghana: VotoMobile; 2015.

[101] Guidelines for Best Practice in Cross-Cultural Surveys. 4th ed. Survey Research Center, Institute for Social Research: University of Michigan; 2016. p. 842.

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

[103] Research TAAfPO. Standard Definitions: Final Dispositions of Case Codes. Standard Definitions. 9th ed. Washington, DC The American Association for Public Opinion Research 2016. p. 80.

[104] Minnis AM, Muchini A, Shiboski S, et al. Audio computer-assisted self-interviewing in reproductive health research: reliability assessment among women in Harare, Zimbabwe. Contraception. 2007;75:59-65.

[105] Langhaug LF, Cheung YB, Pascoe SJ, et al. How you ask really matters: randomised comparison of four sexual behaviour questionnaire delivery modes in Zimbabwean youth. Sex Transm Infect. 2011;87:165-73.

[106] Langhaug LF, Cheung YB, Pascoe S, Hayes R, Cowan FM. Difference in prevalence of common mental disorder as measured using four questionnaire delivery methods among young people in rural Zimbabwe. J Affect Disord. 2009;118:220-3.

[107] VotoMobile. What is Voto? 2013.

[108] Kish L. Weighting for Unequal Pi. 1992;8:183-200.

[109] Areski B. Hang Up Cause Code Table FreeSwitch; 2014.

[110] Andridge RR, Little RJ. A Review of Hot Deck Imputation for Survey Non-response. Int Stat Rev. 2010;78:40-64.

[111] Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. J Gen Intern Med. 2011;26:192-6.

[112] The Mobile Economy in West Africa 2018. London: GSMA 2018. p. 36.

[113] Etang-Ndip A, Hoogeveen J, Lendorfer J. Socioeconomic Impact of the Crisis in North Mali on Displaced People. Policy Research Working Paper. Washington, DC The World Bank; 2015. p. 30.

[114] Dillon B. Using Mobile Phones to Conduct Research in Developing Countries. Journal of International Development. 2011;24:11.

[115] Lessler JT, Kalsbeek WD. Nonsampling Error in Surveys. USA: John Wiley & Sons, Inc.; 1992.

[116] Pearson JM. Frame In: Lavrakas PJ, editor. Encyclopedia of Survey Research Methods. Thousand Oaks: Sage Publications, Inc.; 2011. p. 291-2.

[117] McNabb DE. Frame Error. Nonsampling Error in Social Surveys. Thousand Oaks: SAGE Publications, Inc.; 2014. p. 81-96.

[118] Measure DHS. Calverton: ICF International; 2016.

[119] Afrobarometer. Afrobarometer: About 2017.

[120] Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO. Heterogeneous mobile phone ownership and usage patterns in Kenya. PLoS One. 2012;7:e35319.

[121] Zurovac D, Otieno G, Kigen S, et al. Ownership and use of mobile phones among health workers, caregivers of sick children and adult patients in Kenya: cross-sectional national survey. Globalization and Health. 2013;9:7.

[122] Blumenstock J. Divided We Call: Disparities in Access and Use of Mobile Phones in Rwanda. Information Technologies & International Development. 2012;8:17.

[123] Jennings L, Omoni A, Akerele A, Ibrahim Y, Ekanem E. Disparities in mobile phone access and maternal health service utilization in Nigeria: a population-based survey. Int J Med Inform. 2015;84:341-8.

[124] 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.

[125] Maiga A, Hounton S, Amouzou A, et al. Trends and patterns of modern contraceptive use and relationships with high-risk births and child mortality in Burkina Faso. Glob Health Action. 2015;8:29736.

[126] PMA2020, démographie BFIndlsedl. PMA2016/Burkina Faso-R4. In: PMA2020, editor. Ouagadougou, Burkina Faso and Baltimore, Maryland, USA: The Bill & Melinda Gates Institute for

Population and Reproductive Health at The Johns Hopkins Bloomberg School of Public Health; 2016.

[127] Eckman S, O'Muircheartaigh C. Performance of the Half-Open Interval Missed Housing Unit Procedure. Survey Research Methods 2011;5:125-31.

[128] Kennedy C, McGeeney K, Keeter S, et al. Implications of Moving Public Opinion Surveys to a Single-Frame Cell-Phone Random-Digit-Dial Design. Public Opinion Quarterly. 2018;82:279-99.
[129] Stat Compiler Demographic and Health Surveys 2018.

[130] Bethlehem J, Bakker B. The impact of nonresponse on survey quality. Statistical Journal of the IAOS. 2014;30:5.

[131] Dillman DA, Eltinge JL, Groves RM, Little RJA. Survey Nonresponse in Design, Data Collection, and Analysis. In: Groves RM, Dillman DA, Eltinge JL, Little RJA, editors. Survey Nonresponse. New York NY: John Wiley & Sonsi, INc.; 2002. p. 500.

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

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

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

[135] 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.

[136] Keeter S, Kennedy C, Dimock M, Best J, Craighill P. Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey. The Public Opinion Quarterly. 2006;70.

[137] Marston M, Harriss K, Slaymaker E. Non-response bias in estimates of HIV prevalence due to the mobility of absentees in national population-based surveys: a study of nine national surveys. Sex Transm Infect. 2008;84 Suppl 1:i71-i7.

[138] Reniers G, Eaton J. Refusal bias in HIV prevalence estimates from nationally representative seroprevalence surveys. AIDS. 2009;23:621-9.

[139] Guiella G, Turke S, Coulibaly H, Radloff S, Choi Y. Rapid Uptake of the Subcutaneous Injectable in Burkina Faso: Evidence From PMA2020 Cross-Sectional Surveys. Glob Health Sci Pract. 2018;6:73-81.

[140] Demombynes G, Gubbins PR. Challenges and Opportunities of Mobile Phone-Based Data Collection Evidence from South Sudan. In: Bank TW, editor. Policy Research Working Paper. Washington, DC: The World Bank; 2013.

[141] Cecatti JG, Camargo RP, R.D. P, et al. Computer-assisted telephone interviewing (CATI): using the telephone for obtaining information on reproductive healt. Cad Saúde Pública. 2011;27:8.

[142] Short Fabic M, Choi Y, Bird S. A systematic review of Demographic and Health Surveys: data availability and utilization for research. Bull World Health Organ. 2012;90:604-12.

[143] World Contraceptive Use by Marital Status and Age - Methodology. New York, NY: United Nations, Department of Economic of Economic and Social Affairs; 2018. p. 5.

[144] Choi Y, Short Fabic M. Monitoring Progress in Equality for the Sustainable Development Goals: A Case Study of Meeting Demand for Family Planning. Glob Health Sci Pract. 2018;6:387-98.

[145] Chandy L, Zhang C. Stuffing data gaps with dollars: What will it cost to close the data deficit in poor countries? : Brookings Institute; 2017.

[146] The Mobile Economy Sub-Saharan Africa 2018. In: GSMA, editor. London 2018.

[147] Segri NJ, Cesar CL, Barros MB, Alves MC, Carandina L, Goldbaum M. Health survey: comparison of interviewees according to ownership of a residential telephone line. Rev Saude Publica. 2010;44:503-12.

[148] Tran MC, Labrique AB, Mehra S, et al. Analyzing the mobile "digital divide": changing determinants of household phone ownership over time in rural bangladesh. JMIR Mhealth Uhealth. 2015;3:e24.

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

[150] Stecklov G, Weinreb AA, Sana M. Family Planning for Strangers: An Experiment on the Validity of Reported Contraceptive Use. Plos One. 2015;10.

[151] Bernal RT, Malta DC, Araujo TS, Silva NN. Telephone survey: post-stratification adjustments to compensate non-coverage bias in city of Rio Branco, Northern Brazil. Rev Saude Publica. 2013;47:316-25.

[152] Schreiner M. There's No Place Like Home (How the Interview Method Affects Results with the Progress out of Poverty Index). In: India GF, editor.: Grameen Foundation; 2015. p. 51.

[153] Dahir A. Afircan governments are withholding millions of dollars meant to increase internet access. Quartz Africa; 2018.

[154] Dabalen A, Etang A, Hoogeveen J, Mushi E, Schipper Y, von Engelhardt J. Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection. Directions in Development. Washington, DC World Bank 2016.

## 9. CV

# Abigail Greenleaf, MPH

abgreenleaf@gmail.com
*New York, NY*

## Education

2019 — Doctor of Philosophy, Department of Population Family & Reproductive Health
**Johns Hopkins University Bloomberg School of Public Health**

2012 — Master of Public Health, *magna cum laude*
**Columbia University Mailman School of Public Health**

2007 — Bachelor of Science in Public Health, minor Anthropology, *cum laude*
**The George Washington University**

## Employment

**Performance Monitoring and Accountability (PMA) 2020, Johns Hopkins University, Baltimore, MD**

12/2017 – Present — Program Officer

- Design and conduct random digit dial (RDD) reproductive health computer-assisted telephone interview (CATI) study of 2,400 women in Burkina Faso. Tasks include writing and submitting IRB protocol, preparing survey tools, organizing pilots, training staff (6 supervisors & 20 interviewers), call center oversight and report writing.
- Designed and managed CATI & Interactive Voice Response (IVR) randomized cross-over follow-up study of 1,800 women in Burkina Faso. Tasks included creating IRB protocol, translating survey tools, recording IVR questions and conducting user testing to ensure accurate translations, managing survey pilots, training staff and managing day-to-day operations of 15 interviewers and their supervisors in call center.

9/2015 – 11/2017 — Research Assistant, Data Management Team

- Created study protocol and implemented qualitative research in Burkina Faso to explore the impact of interviewer-respondent familiarity on data quality.
- Cleaned large data sets collected in STATA for public release.
- Supervised three Masters students working on data cleaning for public release.

**Data 4 Health, Johns Hopkins University, Baltimore, MD**

5/2016 – 10/2017 — Research Assistant

- Led data management and assisted with survey operations of 15 IVR mini-studies in three countries.
- Managed two Master students conducting data cleaning.

**VotoMobile, Niamey, Niger**

5/2016 – 09/2016 — One Health – Animal and Human Health Surveillance Project

- Collaborated with Ministry of Health and Ministry of Livestock & Agriculture to design and implement IVR active surveillance hotline for two human diseases (measles and polio) and two animal diseases.

**American Schools and Programs of Public Health, Allan Rosenfield Global Health Fellow**

9/2014 – 8/2015 — CDC Cameroon - PEPFAR

- Coordinated a serial cross-sectional study of sexually active women living with HIV/AIDS at two HIV treatment clinics in southwest Cameroon. Managed interviewer training, pilot and baseline data collection via interviews and chart abstractions.
- Conducted maternal death surveillance and response situation analysis with the objective of improving maternal death reporting and use of data for action in Cameroon.
- Collaborated with implementing partner to write qualitative research proposal to examine facilitators and barriers to Option B+ uptake among HIV+ women in Southwest Region of Cameroon.

9/2013 – 9/2014 CDC Ethiopia - Field Epidemiology Training Program, (*1 year fellowship limit*)

- Created protocol and submitted to Addis Ababa University IRB to examine experiences of Ethiopia FETP graduates during classroom training and in the field by gender. Created online questionnaire and analyzed data.
- Created standard measles and dengue fever case-control questionnaires for outbreak investigations.
- Managed proposal development and the implementation of ten mini-grants for planned malaria studies.
- Participated in measles outbreak investigation and case-control study in rural Oromia Region, Ethiopia.
- Compiled geospatial data for polio and pastoralist landscape analysis to improve polio micro-planning.

**Columbia University, Mailman School of Public Health, New York, NY**

3/2013 – 8/2013 ICAP – Health System Division, Program Assistant

- Assisted with health systems strengthening training for over 50 CDC, USAID and Ministry of Health staff from seven countries in Southern Africa.

4/2013 – 9/2013 Office of Educational Programs, Teaching Assistant

- Reviewed all department practicum evaluations to launch school-wide core practicum evaluation.
- Assisted with the expansion of qualitative foundation curriculum from four to six lectures.

**New York City Department of Health & Mental Hygiene, Bureau of HIV/AIDS Prevention & Control**

4/2012 – 11/2012 Intern**,** New York, NY

- Conducted interviews on housing status, HIV treatment, sexual behavior, and drug use for a randomized controlled trial that assessed the health outcomes of HIV-positive homeless persons when stably housed.

**Harlem Health Promotion Center, New York, NY**

9/2011 – 5/2012 Mobile Health Educator, Project STAY

- Presented sexual health education to and conducted sexual health risk assessments with high school students in New York City in order to reduce unintended pregnancy and encourage STI testing.

**United States Agency for International Development, President's Malaria Initiative**

7/2010 – 9/2011 Program Assistant, Washington, DC

- Managed day-to-day workflow and deliverables, including coordinating technical briefings and communication for US Global Malaria Coordinator and Deputy.
- Organized workshops with external partners such as The Global Fund, The World Bank, DFID, Roll Back Malaria and multiple African Ministries of Health.
- *Benin***:** Provided technical assistance to PMI Benin team. Prepared Malaria Operational Plan and edited monitoring and evaluation documents. Worked with the Health Team to finalize strategic plan for family planning and maternal and child health programs.

- *Mali*: Prepared the annual Performance Plan Report and contributed to the Global Health Initiative strategic document. Organized and attended maternal and newborn health conference jointly sponsored by US Government and Organization of Islamic Conference.

**Peace Corps, Adamawa Province, Cameroon**

9/2007 – 12/2009 Health Extension Volunteer

- Organized five health education summer camps in four towns for 120 youth. Managed funding, trained and supervised 9 counselors.
- Taught 14 women quilting skills to generate income. Wrote legalization papers and taught business skills.
- Conducted needs assessments with community groups including the Association of Handicapped Persons, SALAMA Muslim Women's group, and a teenage girl's literacy group to identify priorities and organize trainings and educational programs.

## Teaching & Training Experience

**Instructor**

| Year | Semester | Location | Course |
|---|---|---|---|
| 2016 | Summer | Webinar | Co-instructor, *Introduction to Open Data Kit, 2 hours,* PMA2020 |
| 2014 | Winter | Ethiopia | Instructor, Research Design, 4 days, 18 Addis Ababa Univ. students |
| 2013 | Winter | Ethiopia | Instructor, Research Design, 2 days, 16 Addis Ababa Univ. students |

**Teaching Assistant**

| Year | Semester | Number | Format | Course Name |
|---|---|---|---|---|
| **Johns Hopkins University** | | | | |
| 2018 | Spring | 380.840 | In-class | Racism and Sexual and Reproductive Health* |
| 2017 | Winter | 380.640 | In-class | Children in Crisis** |
| 2017 | Spring | 380.840 | In-class | Children in Crisis Practicum |
| 2017 | Spring | 380.749 | In-class | Adolescent Sexual and Reproductive Health |
| 2016 | Winter | 380.640 | In-class | Children in Crisis |
| 2016 | Spring | 380.840 | In-class | Children in Crisis Practicum |

*proposed course to instructor, Dr. Anne Burke, and led syllabus creation

** assisted with updating syllabus and created new assignment

**Centers for Disease Control**

2013 Winter In-class ArcGIS for Polio Microplanning (*two week ArcGIS training sponsored by CDC's Global Immunization Division for 30 FETP graduates and residents*)

**Columbia University**

| Year | Semester | Number | Format | Course Name |
|---|---|---|---|---|
| 2012 | Fall | P6031 | In-class | Qualitative Foundations *(lead TA, managed 7 TAs)* |
| 2012 | Fall | P6052 | In-class | Globalization and Global Health |

**Trainer**

| Year | Length | Location | Training |
|---|---|---|---|
| 2018 | 7 days | Burkina Faso | Supervisor and interviewer training for CATI surveys, PMA2020* |
| 2017 | 4 days | Burkina Faso | Interviewer training for CATI & Hybrid surveys, PMA2020* |
| 2017 | 5 days | Nigeria | In-depth Interviewer (IDI) training, PMA2020 |
| 2017 | 3 days | Niger | Training of trainers, IVR health surveillance, VotoMobile* |

| | | | |
|---|---|---|---|
| 2016 | 5 days | Burkina Faso | Co-trainer, IDI and focus group discussions (FGD), PMA2020* |
| 2016 | 1 day | Niger | End-user (female community health worker) training, VotoMobile* |
| 2015 | 4 days | Cameroon | Interviewer training, family planning and HIV integration, CDC |
| 2015 | 0.5 day | Cameroon | Site Improvement Monitoring System tablet data collection, CDC |
| 2014 | 4 days | Ethiopia | Co-trainer, In-depth interview & Focus Group Discussions for HIV study, CDC |
| 2013 | 1 day | Ethiopia | Open Data Kit training for 16 Ethiopia FETP students and graduates |

* denotes training was in French

# Publications

**2018**

**Greenleaf AR**, Ahmed S, Moreau C, Guiella G, Choi Y. 2018. Cell Phone Ownership and Modern Contraceptive use in Burkina Faso: Implications for Research and Interventions using Mobile Technology. Contraception, 2018. https://doi.org/10.1016/j.contraception.2018.11.006

**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.

**Greenleaf AR**, Gadiaga A, Turke S, Battle N, Ahmed S, Moreau C, Choi Y. 2018. Comparison of remote data collection modes to monitor family planning progress in Burkina Faso: representativeness, data quality, and cost. Performance Monitoring and Accountability 2020 Methodological Reports No. 4. Baltimore, Maryland, USA: Bill & Melinda Gates Institute for Population and Reproductive Health, Johns Hopkins University Bloomberg School of Public Health and *Institute Supérieur des Sciences de la Population. (Forthcoming, Dec 2018)*

**2017**

**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(5): e140.

Gibson DG, Pariyo GW, Wosu AC, **Greenleaf AR**, Ali J, Ahmed S, et al. *Evaluation of Mechanisms to Improve Performance of Mobile Phone Surveys in Low- and Middle-Income Countries: Research Protocol.* JMIR Res Protoc. 2017;6(5): e81.

Hawes M, Safi S, **Greenleaf A,** Tsui A, Guiella G, Shiferaw S, Seme A, Otupiri E, Gicahngi P, Makumbi F. 2017. *Response patterns on behavioral outcomes in relation to use of resident enumerators over multiple survey rounds*. Performance Monitoring and Accountability 2020 (PMA2020) Methodological Reports No. 1. Baltimore, Maryland, USA: Bill & Melinda Gates Institute for Population and Reproductive Health, Johns Hopkins University.

# Presentations

**2018**

Mobile phone surveys for family planning: Comparison of data quality between two collection modes in a cross-over randomized study in Burkina Faso. **Presenter,** with Guiella G, Gadiaga A, Turke S, Battle N, Ahmed S, Moreau C, Choi Y. International Conference on Family Planning. Kigali, Rwanda. Nov 12-15, 2018.

Non-response among women of reproductive age in Burkina Faso contacted for IVR or CATI cell phone survey. **Presenter**, with Ahmed S, Moreau C, Guiella G, Gadiaga A, Choi Y. American Association for Public Opinion Research Annual Conference. Denver, Colorado. May 17.

**2017**
*Exploring mobile phone surveys for population health: scientific, implementation and ethical considerations.* **Panelist**, with Ali J, Gibson D, Labrique A. Global Digital Health Forum 2017. Washington, DC. December 5.

*Optimizing interactive voice response data quality: Lessons learned from Bangladesh, Tanzania & Uganda.* **Presenter**, with Gibson D, Labrique A, Pariyo G, Hyder A. Comparative Survey Design and Implementation Workshop. Mannheim, Germany. March 16.

*Resident Interviewers and Repeat Surveys: Effects on Measures of Reproductive Health*. Presentation. Safi S (presenter), **Greenleaf A**, Hawes M, Gichangi P, Guiella G, Makumbi F, Otupiri E, Shiferaw S, Tsui A. Population Association of America. Chicago, Illinois. April 27.

**2016**
*Maternal Death Surveillance and Response District Level Situation Analysis Abong Mbang, Bafia and Sa'a Districts Cameroon, May – June 2015.* Presentation. Bita G (presenter), **Greenleaf A**. 2016 CityMatCH Leadership & MCH Epidemiology Conference. Philadelphia, Pennsylvania. September 16.

# Awards & Honors

**2018** **American Association for Public Opinion Research (AAPOR) – DC Chapter**

- DC-AAPOR Student Paper Competition Winner "Comparability of modern contraceptive use estimates between a face-to-face survey and a cellphone survey among women of reproductive age in Burkina Faso"

**2017** **Johns Hopkins Bloomberg School of Public Health**

- Department of Population, Family and Reproductive Health General Scholarship Recipient

**2016** **Johns Hopkins Bloomberg School of Public Health**

- Fund in Recognition of Laurie Schwab Zabin for Population and Family Planning Students, Receipt
- Edward J. Dehne Award in Population Dynamics, Recipient

**2013** **Presidential Management Fellows Program Finalist**

**2012** **Mailman School of Public Health, Columbia University**

- Global Health Initiative funding for summer internship with UNICEF West/Central Africa in Dakar, Senegal