Title
Self-Managed Abortion: Addressing Methodological Challenges in Measuring Safety, Incidence, and Effectiveness

Author
Jayaweera, Ruvani Tharanga

Publication Date
2021

Peer reviewed|Thesis/dissertation

Self-Managed Abortion: Addressing Methodological Challenges in Measuring Safety, Incidence, and Effectiveness

By

Ruvani Tharanga Jayaweera

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Epidemiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jennifer Ahern, Chair
Professor Patrick Bradshaw
Professor Cassondra Marshall
Professor William McFarland

Spring 2021

Abstract

Self-Managed Abortion: Addressing Methodological Challenges in Measuring Safety, Incidence, and Effectiveness

By

Ruvani Tharanga Jayaweera

Doctor of Philosophy in Epidemiology

University of California, Berkeley

Professor Jennifer Ahern, Chair

Barriers to safe, effective, and affordable methods of abortion include restrictive laws, unwilling or untrained providers, long wait times, high costs, lack of services in the public sector, and abortion stigma. These barriers prevent individuals from exercising their fundamental human right to choose if and when to have a child, and threaten bodily and reproductive autonomy. Self-managed abortion with medication—defined as ending one's own pregnancy outside of a formal healthcare setting using misoprostol alone or in combination with mifepristone—can be safe and effective; however, there is a lack of research on its incidence, safety, and effectiveness, and there are considerable methodological challenges in its assessment.

This dissertation aims to address key methodological challenges in the measurement of self-managed abortion via the following approaches: development of a conceptual framework for measuring abortion complications; assessment of a novel application of respondent-driven sampling (RDS) to study the incidence of abortion; and estimation of the effectiveness of self-managed medication abortion in probabilistic bias-adjusted models.

Chapter 1 describes the development of a framework for measuring complications and outcomes from medication abortion based on self-report, informed by thematic analysis of in-depth interviews from people with various abortion experiences in Argentina, Indonesia, Mexico, Nepal, Nigeria, and Sierra Leone. Findings from this analysis demonstrated that individuals describe and quantify their experiences with bleeding and cramping in varied ways, and highlights the need for a person-centered framework that emphasizes the individual's preferences around medical care seeking. Chapter 2 is a methodological assessment of an RDS study on the incidence of abortion in Soweto, South Africa, and explores the implications of potential violations of RDS assumptions on incidence estimation. In this study, several key assumptions of RDS were not met, yielding potentially biased estimates of abortion incidence. Chapter 3 utilizes data from a prospective observational study assessing the effectiveness of self-managed medication abortion among callers to accompaniment groups and safe abortion hotlines in Argentina and Nigeria, and demonstrates a Monte Carlo Sensitivity Analysis approach to adjusting for

misclassification and selection bias. After adjusting for potential misclassification, selection bias, and enrollment of ineligible participants, self-managed medication abortion remains highly effective, conditional on our assumptions around the chosen bias parameters.

Findings from this dissertation will contribute to the development of a self-report questionnaire to measure complications and outcomes from medication abortion; highlights the pitfalls of respondent-driven sampling and offers potential remedies; and can be used to inform future analyses of the effectiveness of self-managed medication abortion based on observational data.

TABLE OF CONTENTS

I. DEDICATION

For Ammi and Thaththi.

II. ACKNOWLEDGEMENTS

I am extremely grateful to everyone in my life who has given me so much love, support, encouragement, and guidance throughout this journey.

First and foremost, I would like to thank my dissertation committee for their thoughtful support and guidance. Jennifer Ahern, my chair and advisor, has been a continual source of encouragement and wisdom throughout the doctoral program. I am grateful for her willingness to help me think through any study design or analytic problem—whether related to a project we are working together or not—and her understanding and support as I navigated my role as a student and as an employee at Ibis. I have been fortunate to learn from Cassie Marshall, who pushed me to consider how to incorporate a reproductive justice lens not only in this dissertation, but in my overall approach to research. I am also thankful for Patrick Bradshaw's guidance. I have learned so much from him about bias analysis and epidemiologic methods, and I am so grateful for Patrick's support in thinking through the bias framework in this dissertation. I would also like to thank Willi McFarland, whose insightful and direct feedback and interpretation of my research findings helped to center the implications this work in the broader public health field. I feel incredibly grateful to have had the opportunity to work alongside these amazing researchers, mentors, and teachers, and look forward to continuing to collaborate with them as I move forward in my career. In addition to my committee, I would like to thank my PhD cohort for their continual collaboration, support, and friendship.

I would also like to acknowledge my mentor, Caitlin Gerdts, for being such a strong advocate for me and my career development. From her guidance as my GSI when I was an undergrad at Berkeley and just starting to explore my academic interests in sexual and reproductive health, to her role as my supervisor for my MPH practicum where she entrusted me with leading a research study with a safe abortion hotline, to her continual support of me in my work at Ibis, I truly would not be where I am today without her. Thank you for always believing in me, for giving me so many countless opportunities to grow, for your mentorship, and for your friendship.

I'd also like to thank Barbara Abrams and Mahasin Mujahid, who inspired me to pursue a career in Epidemiology. Barbara, thank you for guiding me into public health, and for always being one of my staunchest advocates. Mahasin, thank you for encouraging me to spread my wings and leave the Bay Area for my MPH – and for reminding me that I can always come back.

I have been lucky to work with a truly exceptional team of researchers at Ibis Reproductive Health throughout my doctoral career. Thank you to all of Ibis Oakland for always cheering me on, and for your understanding and patience as I juggled my role as a student and as an employee. In particular, I'd like to thank Heidi Moseson for her thought leadership, intellectual brilliance, and warmth. I am eternally grateful how she is always available to provide insightful feedback, to help me think through analytical problems, and to provide life advice. I'd also like to thank Anna Katz, Katie Key, Relebohile Motana, Sarah Baum, Sofia Filippa, Chiara Bercu, Brianna Keefe-Oates, and Sofia Carbone, without whom this

dissertation would not have been possible. They are all such brilliant researchers and amazing human beings, and I am so grateful to have had the opportunity to work alongside them.

I would also like to thank my colleagues at accompaniment group and safe abortion hotlines who were partners in this work. Ika Ayu, Sybil Nmezi, Ruth Zurbriggen, and Ijeoma Egwuatu – thank you for all that you do for people in need of abortion. I am continually inspired by your courage, strength, and generosity. I have learned so much from all you, perhaps most importantly that activists can be researchers, and researchers can be activists.

To my partner, Anoop Muniyappa—I am forever grateful for your continuous encouragement, love, and support throughout this journey. Thank you for letting me explain statistical concepts to you, for supporting me through the challenges of studying, research, and travel, and for reminding me to find balance and humor amidst it all.

I am so grateful to have had the support of my family. Thank you to my brother, for helping troubleshoot when my code wouldn't run. And of course, a huge thank you to my parents, who instilled in me a love of science and a belief in the importance of working to improve the lives of those around us in whatever way I can. Ammi and Thaththi, you inspired me to get a doctoral degree; thank you for your encouragement and unwavering support in whatever path I decided to follow.

And finally, I would like to thank the research participants for their generosity in sharing their time and their experiences.

III. INTRODUCTION

All individuals have the fundamental human right to maintain personal bodily autonomy, have children, not have children, and parent the children they have in safe and sustainable communities.(1) Achievement of these rights is the core principle of reproductive justice, a framework developed in 1994 by Black women activists. Unfortunately, these rights are not universally enjoyed or accessible to all people, and reproductive oppression—the control and exploitation of women, girls, and individuals through their bodies, sexuality, labor, and reproduction—persists globally. The lack of equitable access to safe and effective methods of abortion is one way in which reproductive oppression operates in limiting reproductive justice. Furthermore, the consequences of this lack of access—which include forced childbearing and its consequent morbidity and mortality, as well as morbidity and mortality from the use of unsafe abortion methods—are distributed inequitably. Poor women[1], women of color, and women from countries recovering from centuries of colonialism bear the brunt of this lack of access. The World Health Organization (WHO) estimates that 97% of the 25.1 million unsafe abortions that occur each year occur in low and middle income regions, and unsafe abortion is responsible for an estimated 8-13% of maternal deaths globally.(2, 3)

WHO-recommended safe methods of abortion include: medication abortion (misoprostol alone, or in combination with mifepristone), manual vacuum aspiration, or dilation and evacuation.(4) Ample evidence has demonstrated that abortions performed by trained providers using these methods are extremely safe with low rates of complications.(5) Unsafe abortions are differentiated as *less safe* and *least safe;* less safe abortions are performed by trained providers using a non-recommended or outdated method (e.g. dilation and curettage), or by a non-trained provider using a recommended method.(3, 4) Least safe abortions are provided by untrained individuals using dangerous,(3, 4) such as ingestion of caustic substances, or insertion of foreign bodies into the uterus.(5) Least safe abortions carry the highest risk for complications such as incomplete abortion, hemorrhage, infection, uterine perforation, life-long morbidities, or death.(5)

Despite the existence of safe methods of abortion, differential access to these methods persist. To date, efforts to expand abortion access have been dominated by discussions around "choice" or legality, with the implication that if abortion is made legal, access will follow. Nearly a quarter of the world's population reside in one of the 66 countries where abortion is prohibited outright or only allowed when the woman's life is at risk; in 72 countries (35% of the world's population), abortion is only permitted under certain conditions, such as rape, incest, mental health exceptions, and socioeconomic grounds.(6) The proportion of abortions that are unsafe in restrictive countries is substantially higher

[1] This dissertation recognizes the limitations of binary definitions of gender, and recognizes that "woman" does not encapsulate the identities of all people who can and will get pregnant, which may also include transgender men and those with other gender identities; furthermore, not all people who define themselves as "women" can become pregnant. The terms "women" and "people" are used throughout this prospectus, with the aim of using the more inclusive term "people" or "people who can get pregnant" when referring to the implications of this work. The use of "women" is not intended to mask the experiences of those who do not use that identity to describe themselves, but is used when describing research that was conducted exclusively on cisgender women (or presumed to be cisgender women), laws or policies that explicitly target women, and in recognition of "woman" as a self-defined category that many choose.

than in countries with less restrictive laws.(3, 5) However, while removal of restrictive abortion laws is a necessary component to reproductive autonomy, it is not sufficient. Even in settings where abortion is legal, laws that permit providers to object to abortion provision,(7) paucity of clinics,(8) lack of willing or trained provides, long wait times,(9, 10) high cost for services, lack of information about where to access safe services, and abortion stigma serve as additional barriers.(11, 12) These barriers are compounded in settings where abortion is legally restricted.(13) Abortion stigma—pervasive in many settings—can discourage people from seeking care in the formal health sector due to fear of judgment and mistreatment.(14, 15) This stigma can be amplified by other forms of stigma, often related to gender, race, and socioeconomic status. As a result of these barriers, equitable access to safe abortion is not a reality for most people, and these barriers are differentially experienced across racial, national, and socioeconomic lines. Consequently, strategies that seek to expand abortion access solely through legal means or via expanding access in public sector services will perpetuate existing inequalities. Women, specifically women of color, have historically been denied equal protection under the law; institutional racism and histories of mistreatment in the medical system highlight the limitations of legal strategies alone. For example, the Hyde Amendment, first introduced in 1977 and incorporated into every Senate Appropriations bill since, restricts the use of United States federal funding to cover abortions; effectively banning those who have publicly-funded health insurance, such as Medicare, Medicaid, Indian Health Services, and Tri-Care (military) from insurance coverage for their procedure.(1) Thus, while Roe v. Wade protects the right to abortion *vis-à-vis* privacy rights, the government has imposed abortion restrictions through other means that disproportionately target low-income people, who, due to centuries of disenfranchisement and oppression, are disproportionately people of color.

Self-managed abortion

Abortion has been practiced in every society around the world for thousands of years. While the medicalization and professionalization of abortion provision occurred during a time of advancement in the safety and effectiveness of abortion methods, it had the negative impact of shifting control and access to fertility control methods from midwives and women themselves into the realm of the medical community and under the control of the state.(1) Prior to the advent of medication abortion and safe surgical methods of abortion, people in need of abortion utilized methods such as herbs, teas, ingestion of caustic substances, insertion of sharp objects into the uterus via the cervix, and abdominal massage, among others.(1, 5) While these methods are still used around the world today, particularly in places where WHO-recommended methods are inaccessible, many of these methods are ineffective and/or carry higher risks of complications such as incomplete abortion, hemorrhage, infection, uterine perforation, life-long morbidities, or death.(5) However, the discovery of the use of misoprostol—a medication developed as treatment for gastric and duodenal ulcers—as a safe and effective abortifacient medication by women in Brazil,(16-18) highlighted the potential of medication abortion to provide people with a safe and effective method of abortion that shifts power and reproductive autonomy back into their own hands.(19)

Misoprostol, one of the two WHO-recommended medications for abortion, is widely available even in countries with restrictive abortion laws at pharmacies or on the black market, given its other indications for use, which have been expanded to include arthritis and treatment for post-partum hemorrhage.(5, 19) Misoprostol can be used alone or in combination with the mifepristone to safely and effectively induce an abortion; both regimens have a well-established evidence base with high safety and effectiveness profiles.(20, 21) As the off-label use of misoprostol has increased, various strategies have emerged to further promote access. For example, in Uruguay, influential physicians began counselling women on how to safely use misoprostol, including the dosing regimen and potential warning signs of complication, and provided them with post abortion counseling to confirm abortion completion. Physicians would not provide any information on where the pills could be obtained. This "harm reduction model" was adopted as official health policy in 2004, and the consequent declines in maternal mortality and morbidity from unsafe abortion (22, 23) are often cited as a leading factor in the successful push to legalize abortion on request in 2012.(24)

Safe abortion hotlines and organizations have expanded on the harm reduction framework, and have emerged as the primary sources of access to information on self-managed abortion (25, 26). Globally, more than 40 feminist grassroots organizations provide people with evidence-based counseling and support through the medication abortion process, largely in settings where abortion is legally restricted and access to abortion within the formal health sector is poor. These safe abortion hotlines, sometimes called "accompaniment models," vary in their modes of operation, but all provide step-by-step protocols for how to use medication to safely induce abortion based on evidence-based protocols. Organizations may also provide information on where to obtain quality medications, how the drugs function, how to manage pain, how to recognize complication signs, how to prepare for potential interactions with medical personnel in case of emergency treatment seeking, how to confirm abortion completion, what to expect after the abortion, and prevention of future unwanted pregnancy.

Research gaps

The safety and effectiveness of medication abortion in clinical settings is well established: mifepristone in combination with misoprostol is 95% effective up to 63 days gestation with extremely low (<0.5%) rates of complications,(27) and misoprostol alone is 75% - 85% effective with similarly low rates of complications.(28-30) Given restrictions on abortion access, research has focused on shifting certain aspects of the abortion process out of the clinic as a means to increase the range of providers who are deemed "qualified" abortion providers. To date, much of the research has focused on comparing safety and effectiveness between doctors and advanced practice clinicians, the role of ancillary providers in assessing contraindications for medication abortion, and assessing the need for follow-up care.(31)

However, there is growing awareness of these accompaniment models and increasing recognition of the role of the individual in safely managing their own abortion. In 2015, the WHO released guidelines that outlined task-shifted roles for health workers in the

provision of safe abortion. For the first time, these guidelines included women as actors in their own abortion process, and acknowledged that the experience of self-management of abortion can be empowering and could lead to a more optimal use of scarce health resources.(1) Preliminary evidence suggests that the practice of self-managed medication abortion may be safe and acceptable,(19, 22, 23, 32, 33) though the lack of prospectively collected data, as well concerns around the quality of medications people are able to obtain outside of the formal healthcare system, their ability to self-assess their gestational age and possible contraindications to medication abortion, as well as their ability to follow the medication abortion protocol, have precluded the definitive evaluation of the individual as a safe provider of one's own abortion. While decades of experience from accompaniment models suggest that self-managed abortion is safe and effective, there is a lack of peer-reviewed evidence in the scientific literature about the incidence, safety, and effectiveness of self-managed medication abortion administered completely outside of the formal healthcare system. As a result, WHO task-shifting guidelines do not yet recognize the full independent management of medication abortion. Key challenges in measuring incidence, safety, and effectiveness are outlined below.

Measuring incidence

The methodologies most commonly used to estimate abortion incidence have widely acknowledged limitations.(34-37) Direct measurement methods, such as asking people directly about their abortion experience(s), results in under-reporting due to social desirability bias, stigma, and privacy concerns. Under-reporting is likely magnified in contexts where abortion is highly stigmatized and/or legally restricted.(34-36, 38) Indirect measurement methods, such as the Abortion Incidence Complications Methods (AICM) and the Anonymous Third Party Reporting (ATPR) method,(39, 40) aim to improve accuracy by turning to sources of information other than the individual, such as health facility data and/or healthcare provider perspectives. With the AICM, a senior health provider at a particular facility is asked to estimate the number of women who presented for post-abortion care (PAC) during the past month; this number is multiplied by estimated ratios of the number of women who had a complication but did not seek care and the number of women who had an abortion but did not have a complication in order to generate an estimate of the total incidence of abortions in a particular country.(41) While indirect methods generate country-level estimates in settings where abortion is restricted or its practice is unsafe, the assumptions and extrapolations of such methods are difficult to test, time-intensive, and complex, and often rely on incomplete data.(42-46) As a result, while recent evidence suggests that nearly half of all abortions that occur globally are performed in illegal or unsafe conditions,(3) challenges in data collection likely lead to underestimates of abortion incidence, as well as biased data on the characteristics and outcomes of abortion in such contexts.

Due to the stigma and secrecy associated with informal sector abortion and abortion in general, new data collection and estimation strategies are needed. Respondent-driven sampling (RDS) is a sampling methodology that has been used to more accurately estimate the prevalence and incidence of sensitive and illegal behaviors among hidden populations such as injection drug users, sex workers, and men who have sex with men.(47-53) RDS

leverages a small non-random sample of initial participants (known as *seeds*) within social networks engaging in hidden or stigmatized behaviors to recruit others within the same social networks (i.e. the target population). Each individual seed is given a set number of coupons with which they can recruit their social network peers. Once a participant with a valid coupon presents to the study site, they are provided with the same number of coupons with which to enroll other members of the social network, thus resulting in a lengthy chain of participants representing the target population.(49, 50, 54) RDS samples are adjusted for potential selection bias in analyses by weighting participants with more contacts in the target population inversely proportional to the number of contacts in the network itself.(49, 50, 55) The validity of RDS as a sampling methodology relies on several key assumptions.(48, 49, 56) First, the population being recruited must be able to identify those in their social network as members of the target population, and form social ties on the basis of this shared characteristic. Second, the referral process should result in a series of overlapping networks (networks of networks), rather than isolated referral chains. Third, sampling should replicate sampling with replacement. Additional assumptions are that participants can accurately report their network size and are randomly sampling their recruits from within this personal network. Studies employing RDS to study informal sector abortion should aim to evaluate these assumptions.

Measuring safety

Though there is a pressing need for accurate estimates of complications from unsafe abortion, our best and most up-to-date estimates rely on facility-based studies and national health statistics in relatively few countries and likely suffer from multiple sources of systematic error, primarily misclassification and selection bias.(57-59) A recent systematic review found 70 facility-based studies of abortion complications in settings where unsafe abortion persists; however, the authors were unable to combine results across studies given substantial heterogeneity in study design and measurement of abortion complications.(60)

While facility-based estimates of abortion complications can tell us about those who present for post-abortion care services,(59) a lack of available data on the total number of people having abortions makes it difficult to accurately estimate the proportion of overall abortions that lead to complications, an issue that is compounded in settings where the majority of abortions occur outside of the formal healthcare system. The AICM(61) attempts to re-create this missing denominator by applying a multiplier of the ratio of those who seek care to those who do not; however, these estimation techniques only yield aggregate counts and proportions of the number of individuals who have abortion complications, and not individual-level information about experiences. Furthermore, people may choose to not seek care because of a lack of perceived severity of abortion symptoms, or may face insurmountable barriers to care such as a lack of financial resources, worries about perceived mistreatment at healthcare facilities, as well as fears of legal repercussions in settings where abortion is criminalized,(62) possibly leading to an undercounting of the number of true complications. Alternative methodologies for addressing this source of selection bias such as robust follow-up for in and out of clinic models in order to document the outcomes of clients and better understand their

trajectories and experiences, and/or innovative methods of recruiting people who have had induced abortions outside of the formal healthcare sector are sorely needed.

In addition to issues with selection bias, there are substantial limitations in our current ability to accurately measure abortion complications (misclassification bias). This bias can operate in two possible directions, with the underreporting of abortion complications in some instances and the over-reporting of complications in others. Standard practice in facility-based estimates is to count all people who present for post-abortion care as an abortion complication. However, it is possible that some people who come to clinics or hospitals during their abortion process are experiencing abortion symptoms (e.g. bleeding, cramping), and are misclassified as having an abortion complication for the sole reason that they presented for care, leading to over-reporting of abortion complications. This situation is likely to arise in settings where abortion is legally restricted, but misoprostol is widely available outside of the formal healthcare setting. It is likely that many such cases are, in fact, not complications but rather abortions following a normal process, but for which people sought care due to a lack of preparedness for the symptoms or having been told to present for care when bleeding began by a lay provider or other confidant. Conversely, people presenting at facilities with true abortion complications may be misclassified as miscarriage complications, leading to an underreporting of abortion complications.(59) This type of underreporting may be especially prevalent in settings where people are prosecuted for suspected abortion and reports of abortion complications or suspicion of induced abortion would place the person at legal risk.(5, 63) While current estimates of abortion complications do attempt to adjust for these sources of misclassification, adjustment factors used in current methodologies do not take into account misclassification due to individuals presenting for care in the course of a medication abortion that is proceeding normally; the possibility of this type of misclassification is likely increasing given the growing availability of misoprostol. Instead of being treated as the root cause of systematic misclassification, availability of misoprostol has historically been presented in the literature as a potential *cause* of documented increases in abortion complications due to low quality medications, untrained providers, or incorrect use. Furthermore, there is a lack of data parsing out the severity of complications from abortion, which has wide ranging implications for how we understand and measure abortion safety.

While classifying abortions correctly from the clinical record alone is challenging enough, self-report of complications is potentially even more fraught, with the difficulty of mapping individual's descriptions about their experiences to clinical definitions of complications. For out of clinic or telemedicine models of abortion provision, the lack of a standardized framework and definition of abortion complications makes it impossible to differentiate self-reported symptoms from complications.(64) An important first step in developing a standardized and validated self-report measurement tool is high-quality qualitative research that furthers our understanding of how people experience and describe abortion symptoms and potential complications, as well as their reasons for seeking or not seeking medical care.

Measuring effectiveness

Given the high percentage of individuals choosing to terminate their pregnancies outside of the healthcare system, rigorous well-designed studies are needed to assess the effectiveness of these alternative models of abortion care. Indeed, a recently proposed research agenda identified the issue of evaluating the effectiveness of self-managed medication abortion as one of three priority research gaps to address.(65) In a recent scoping review of the literature on self-managed abortion, the authors identified 13 studies that described the experiences of people using mifepristone and misoprostol to self-manage their abortion, and 35 studies that described the experience of people using misoprostol alone.(66) For some studies, the study population were women who sought care for abortion complications, and thus only contributes to our understanding of effectiveness among a select group of individuals who sought additional medical care. Among the remaining studies, there was substantial heterogeneity in how effectiveness was defined, precluding the ability to pool or compare findings across studies. Furthermore, the authors highlight the lack of prospectively collected data on a non-hospital-based sample of people self-managing their abortion, limiting our understanding of the safety and effectiveness of self-managed abortion among a more representative sample of abortions.

However, measuring effectiveness of self-managed abortion outside of facility settings carries a host of unique challenges. Clinical studies have established the effectiveness of misoprostol alone and mifepristone in combination with misoprostol at 75-85%(28-30) and 90-95%(27) respectively, with effectiveness defined as complete abortion without the need for surgical intervention. However, complete abortion in these studies was typically ascertained by ultrasound or negative pregnancy test confirmed by a provider, which is not possible in studies where abortion is not legal, occurring outside of the formal healthcare system, and where care-seeking may place people at legal risk. Furthermore, counselors at safe abortion hotlines and accompaniment organizations may never meet the people they are supporting face-to-face; given legal risks and social stigma, those who are seeking abortion support do not often disclose their real names and use temporary phones, making robust follow-up challenging. Thus, studies that seek to measure the effectiveness of self-managed abortion in these contexts may be prone to misclassification and selection bias if loss to follow-up is non-differential with respect to the outcome. Analytic techniques that seek to quantify and correct for these sources of bias should be employed.

The overarching goal of this dissertation is to address specific limitations and gaps in research on self-managed abortion that arise because of methodological challenges in applying standard measurement approaches; namely, limitations in accurately measuring abortion **incidence** in settings where abortion is legally restricted, the lack of validated tools to measure abortion **safety** via self-report of complications, and the lack of prospective data measuring the **effectiveness** of self-managed medication abortion.

1 Chapter 1. Development of a new framework for conceptualizing complications from medication abortion

1.1 Abstract

To date, there is no standardized framework for measuring complications or outcomes from medication abortion based on self-report. To inform the development of a framework, we conducted in-depth interviews with 68 medication abortion users from a range of contexts. We found that individuals describe and quantify their experiences with bleeding and cramping in varied ways. Our findings highlight the potential discrepancies between individual's language and perceptions of their abortion experience with clinical signs and criteria, as well as the need for a person-centered framework that emphasizes the individual's preferences around medical care seeking. Additionally, we found the need for reframing medical treatment as a positive outcome (if the person receives the treatment they wanted in a timely manner), rather than considering it as an indicator of a "complication."

We propose a new framework that reserves the use of the term "complication" for moderate or severe unanticipated problems that arise following a treatment, procedure, or condition, and shifts our focus on measuring outcomes from medication abortion to measuring other dimensions of quality that center the needs and desired experiences of individuals.

1.2 Introduction

Medication abortion is the most commonly used method of abortion worldwide, and innovations in increasing access to medication abortion carry the greatest potential for expanding safe abortion care in a wide variety of settings and contexts. Medication abortion methods, as compared to surgical methods, require fewer resources to administer, can be used in a range of settings, are straightforward to use, and can be provided by a wider cadre of providers, including the pregnant person themselves. The World Health Organization's (WHO) two recommended regimens for medication abortion—misoprostol on its own, and mifepristone in combination with misoprostol(67)—are 80% - 95% effective, depending on regimen, route of medication administration (buccal, vaginal, sublingual), and pregnancy duration.(68, 69)

The use of medication abortion began outside of the formal healthcare system: in the late 1980s, women in Brazil, unable to obtain surgical abortions in the formal healthcare system, discovered the use of misoprostol (which was originally developed as a treatment for gastric and duodenal ulcers) as a safe and effective abortifacient.(16) Following clinical trials that demonstrated the safety and efficacy of misoprostol for abortion, misoprostol became a standard option for clinic-based abortion globally. As use of medication abortion within the formal healthcare system expanded, use of these medications outside of the formal healthcare systems also continued to rise, buttressed by growing availability of the medications and growing recognition of its' safety and effectiveness.(70) Self-managed abortion, defined here as when a person performs their own abortion using misoprostol

alone or in combination with mifepristone without clinical supervision, is credited with declines in maternal morbidity and mortality.(71) Over the past 40 years, online telemedicine services, safe abortion hotlines, feminist networks, and other community-based distribution models have emerged around the world. These organizations operate in contexts where facility-based abortion is most difficult to access, and provide people who are self-managing their abortion with evidence-based information about how to procure medications, accurate timing and dosing, what to expect, how to confirm completion, and how and when to seek necessary health care. These organizations have further facilitated the rise of self-managed abortion using both medication abortion regimens, and have expanded access to people in need of abortions around the world, regardless of legal context.(72) In addition to the rise of these de-medicalized models of care, there has been a recent shift in the formal healthcare system to service delivery models that do not require the person in need of an abortion to have multiple in-person visits to a health facility, such as "no-touch" and telemedicine models of care.(73, 74)

Approaches to assess the outcomes of self-managed or telemedicine abortion models are an important area of ongoing inquiry. In addition, accurate estimates of the rate and severity of abortion complications are key to evaluating innovative out-of-clinic models of abortion provision, such as safe abortion hotlines, medication abortion access at pharmacies, and telemedicine models. While WHO task-shifting guidelines currently recognize certain the role of nurses and pharmacists in the provision of medication abortion care, these guidelines for abortion provision do not yet include lay providers or the person themselves as recommended providers of medication abortion care due to a lack of published evidence on the safety of these models.(75) While we know from ample clinical evidence that medication abortion is exceedingly safe, and serious safety events are rare,(68, 69, 76) availability of misoprostol has historically been presented in the literature as a potential *cause* of documented increases in abortion complications due to low quality medications, untrained providers, or incorrect use.(59) As a result, efforts to increase access to models that support individuals in self-managed medication abortion are hampered by an inability to accurately assess outcomes from different models of care.

While there is a pressing need for accurate estimates of complications and outcomes from medication abortion, our best and most up-to-date estimates rely on facility-based studies and national health statistics in relatively few countries and likely suffer from multiple sources of systematic error, primarily misclassification and selection bias.(57-59) While facility-based estimates of abortion complications can tell us about those who present for post-abortion care services,(59) people may choose not to seek care because of a lack of perceived severity of abortion symptoms, or may face insurmountable barriers to care such as a lack of financial resources, worries about perceived mistreatment at healthcare facilities, as well as fears of legal repercussions in settings where abortion is criminalized,(62) possibly leading to an undercounting of the number of true complications.

In addition to issues with selection bias, there are substantial limitations in our current ability to accurately measure abortion complications (misclassification bias). This bias can operate in two possible directions, with the underreporting of abortion complications in

some instances and the over-reporting of complications in others. Standard practice in facility-based estimates is to count all people who present for post-abortion care as an abortion complication.(61) However, it is possible that some people who come to clinics or hospitals during their abortion process are experiencing abortion symptoms (e.g. bleeding or cramping),(77) and are misclassified as having an abortion complication for the sole reason that they presented for care, leading to over-reporting of abortion complications. This situation is likely to arise in settings where abortion is legally restricted, but misoprostol is widely available outside of the formal healthcare setting.(57) It is likely that many such cases are in fact not complications, but rather abortions following a normal process, but for which people sought care due to a lack of preparedness for the symptoms or having been told to present for care when bleeding began by a lay provider or other confidant.

Furthermore, there is a lack of research parsing out the severity of complications from abortion, which has wide ranging implications for how we understand and measure abortion safety. For example, existing frameworks for measuring abortion complications are not specific to medication abortion methods, which carry much lower risk of physical injury. Furthermore, these frameworks rely on clinical signs and symptoms (such as blood pressure or hemoglobin levels) or receipt of medical treatment as criteria for classifying abortion severity.(78, 79) These frameworks do not incorporate any information about the individuals' reasons or motivations for seeking care. As a result, they may misclassify those who seek care due to concerns around bleeding but who receive no medical treatment, as a "low severity" complication. Importantly, receipt of any surgical intervention after a medication abortion is used as an indication of a potentially adverse event, though it does not capture whether surgical intervention was indicated or requested by the individual themselves.

As a result of these limitations, alternative methodologies to document abortion outcomes, particularly for those who never seek medical care, are sorely needed. However, while correct classification of abortions from the clinical record is challenging enough, many individuals who self-manage may never seek care or have clinical confirmation of their experiences. For these individuals, standardized methods to capture their experiences via self-report is needed. Challenges around classifying complications and understanding safety from self-report arise given the nuances of individual language and perceptions, which may not have standardized understanding and may be difficult to align with clinical definitions of complications. As a result, for out of clinic or telemedicine models of abortion provision, the lack of a standardized framework and definition of abortion complications makes it difficult to systematically differentiate self-reported symptoms from complications.(64) An important first step to develop a standardized and validated person-centered self-report measurement tool is formative qualitative research that furthers our understanding of how people experience and describe abortion symptoms and potential complications in their own words, as well as their reasons for seeking or not seeking medical care.

In light of these measurement challenges, and to inform the development of a self-report measurement tool for measurement of abortion outcomes, including complications, we

conducted a qualitative study with participants from six different countries. The aim of this paper is to describe individual's perceptions and descriptions of their abortion symptoms, specifically bleeding and cramping, to understand how these factors contribute to post-abortion care seeking. Based on these findings, we proposed a measurement framework for quantitatively measuring and classifying outcomes from medication abortion via self-report.

1.3 Methods

Participants and study sites

Between August 2018 and October 2019, we conducted semi-structured in-depth interviews with 68 participants. This study was embedded as part of several separate studies on abortion experiences; eligibility, recruitment, and study procedures varied slightly across the different sites. Participants were recruited from private non-profit clinics, feminist accompaniment groups, and safe abortion hotlines from Argentina, Indonesia, Mexico, Nepal, Nigeria, and Sierra Leone. These locations and models of care were chosen in order to capture a broad range of abortion experiences.

In Mexico City and Nepal, abortion is legal, but a substantial proportion of abortions occur outside of the formal healthcare sector; in Argentina, Indonesia, Nigeria, and Sierra Leone, abortion is only legal to save the life of the pregnant woman, but medications for abortion are widely available via informal sellers or from pharmacies under other indications.

Participants were eligible to participate in the qualitative interviews if they 1) were at least 18 years of age; 2) sought safe abortion (SA) or post-abortion care (PAC) services at participating clinic or received information about self-managing their abortion with medications from a participating hotline or accompaniment group during the study period; 3) were able to grant informed consent; 4) consented to being interviewed and audio-recorded; 5) spoke English or primary local language (Bahasa Indonesia, Hausa, Igbo, Nepali, Spanish, Swahili, or Yoruba) at each site; and 6) had completed their PAC or SA treatment. Staff at each participating site identified potentially eligible clients for participation in this study after their care provision was complete.

The larger study included 12 participants who had surgical abortion experiences in clinic settings; data from this study restricted to those who reported using medication abortion for their initial abortion attempt.

Study procedures

Interested participants were purposively selected for interview based on age, service received (surgical abortion, medication abortion, PAC, self-managed abortion information and support), and health care seeking experiences to ensure that a range of experiences were represented. Sample size ranges for each country site were determined *a priori* based on the goals of the larger studies. The lead investigator and local investigator made a determination when to collect additional data up to predetermined maximum sample size

based on resource constraints if informational redundancy from the completed interviews was not met. Depending on the specific study site context, interviews took place via telephone or in-person at a private interview location. Interviews were conducted by local interviewers trained in qualitative research data collection methods. All participants provided their verbal or written informed consent to participate and to have their interview audio-recorded prior to the start of the interview. Participants received an incentive between $5 - $25 USD for participation in cash, gift card, mobile money, or transportation voucher, depending on the study site.

The interview guide included questions about participants' experiences with their abortion, experiences and preparedness for side effects and warning signs of potential complications, and care seeking behavior for any potential complications that were experienced. Questions related to experiences of complications were the same across all sites; however, depending on site, additional questions may have been included as this study was embedded in several larger studies; those data are not presented here. The portion of the interview guide related to experiences with abortion was developed by the lead investigator, with input from clinic research staff and hotline staff at each site, as well as a review of existing complications frameworks.(78, 79) All interviews were audio-recorded, transcribed in the language they were conducted in, and translated into English.

Ethical approval for this study was obtained from the Marie Stopes Ethical Review Committee (London, United Kingdom), the Nepal Health Research Council (Kathmandu, Nepal), the Allendale Investigative Review Board (Old Lyme, CT) and the Comité de Bioética de Fundación Huésped (Buenos Aires, Argentina).

Qualitative analysis of in-depth interviews

As the overall aim of the study was to identify the breadth of ways in which people describe their abortion experiences in order to identify key method, transcripts were thematically analyzed using a phenomenological approach, which seeks to understand the "essence" of an experience or phenomena as experienced by the individual.(80, 81) In the case of this study, we sought to understand how participants, in their own words, describe their experiences with bleeding, cramping, and other experiences related to their abortion, as well as the complexity of the decision-making process of whether to seek post-abortion care. A draft codebook was developed *a priori* based on the qualitative interview guide and aims of the overall study. Two coders, the lead investigator and the second author, both based in the United States and trained in qualitative research methods, reviewed four transcripts to identify additional emergent codes; these additional codes were incorporated into the codebook. The same two coders applied the codebook to two additional transcripts, discussed discrepancies in coding application, and modified the codebook accordingly. One coder (second author) proceeded to code all transcripts were coded in MAX QDA qualitative analysis software;(82) coding was reviewed by the lead investigator for consistency and completion.

The second author summarized codes in analytical memos, illustrated using direct quotes from the participants. The lead investigator and the second author reviewed the codebook

and memos, and discussed emergent themes. Salient themes, along with illustrative quotes are included below.

Conceptual framework development

Findings from the in-depth interviews and implications for measuring medication abortion outcomes were discussed individually with six leading clinical experts in abortion research and provision. Based on these discussions and the results from the qualitative research, we developed a conceptual framework of understanding what constitutes a "high quality" abortion and key domains for measurement of abortion complications.

1.4 Results

A total of 68 participants were interviewed for this study. The majority of participants were between the ages of 25 – 34 (Table 1). The duration of pregnancy at the time of their abortion ranged from 1 month to 4 months, though pregnancy duration was unknown for 19 participants. A total of 36 participants were medication abortion clients recruited from a private clinic or mobile clinic, 26 were recruited from a safe abortion hotline or accompaniment group, and 6 were post-abortion care clients who started their abortion process outside the formal healthcare system and were recruited from a private clinic after receiving post-abortion care (surgical intervention). Other demographic characteristics (education, number of children) are described in Table 1. All participants ultimately had a complete abortion; 11 had a surgical intervention after taking medication abortion.

We explore participants' experiences with bleeding, experiences with pain and cramping, and reasons for seeking care below. Salient quotes include the following participant descriptors: participant age, pregnancy duration at time of abortion, model of abortion access and support, whether or not the participant sought any additional medical care, and treatment received. Models of medication abortion (MA) access and support include: MA from mobile clinic, MA from private clinic, MA with hotline support, MA from unknown provider.

Experiences with bleeding

Experiences with bleeding were described by participants across the following domains: amount or intensity of bleeding, duration of bleeding, consistency of bleeding, management of bleeding, and concerns around bleeding.

There were two primary typologies for how participants described the *amount or intensity* of bleeding: in comparison to the amount of bleeding they experience in their normal menstrual period, and/or bleeding that was normal, heavy, or severe.

A total of 15 participants described their bleeding as similar to what they experienced on their period: "*But it wasn't like I was bleeding badly.* ***It was normal****, like if you had your period*." (34 years old, 12 weeks, MA with hotline support, went to gynecologist to confirm completion, had ultrasound). Three participants described their bleeding by saying they got their "menses", rather than referring to having bleeding in distinct terms related to the

procedure: *"**I did not bleed I just saw my menses."*** (29 years old, 8 weeks, MA from mobile clinic, did not seek additional care). Eight participants described their bleeding as "heavier" than they typically experienced on their period.

A total of 23 participants described their bleeding as "not too much" or "normal." However, most (n = 42) participants tended to describe the bleeding they experienced as "abundant," "heavy," or "a lot": "***Sometimes the blood was only a little. But some other time, it was like a flood***."(Unknown age and pregnancy duration, MA with hotline support, did not seek follow-up care).

Overall, participants typically reported bleeding that began within hours or a couple days after taking the medication, and lasted anywhere from a couple days to more than a month. Often, participants described initial heavy bleeding that tapered off after several days or weeks: *"**It flowed very heavy** at the first day and it became lighter afterwards at the second… the day following it became lighter **but it was very heavy and in clots on that first day**."* (23 years old, 5 weeks, MA with hotline support, did not seek follow-up care).

Participants frequently described the consistency of their bleeding, with words such as "chunks," "clots," or "chewy," when describing bleeding leading up to passing the products of conception: "*At first my bleeding would continuously happen and after two days it would stop. And then chunks would come out*." (26 years old, unknown pregnancy duration, MA from unknown provider, sought care at clinic due to ongoing bleeding and received a surgical intervention) and "*The bleeding was **hugely clotted**… the clots that was coming out was heavy."* (18 years old, 6 weeks, MA from mobile clinic, did not seek care).

<u>Management of bleeding</u>

Participants in all contexts reported using sanitary napkins (pads) or cloth to manage their bleeding; the types of pads reported varied within and across contexts. Participants frequently described needing to find and purchase pads that were different from the pads they normally wore, such as pads with wings, or thicker/heavier pads for nighttime use that could withstand heavier bleeding and helped clients avoid staining their clothing. They generally had pads on hand from their monthly periods, had purchased the pads from a medical shop in anticipation of bleeding, or had been given pads by their provider. The amount of pads or cloths used, type of pads used, and frequency with which they changed their pads varied depending on how much they were bleeding:

> "*I changed the pads after an hour, or it was less than an hour, the blood was already a lot. So I didn't count… I had just changed (the pads), then only after a few minutes, I felt like the blood was going to come out again, and it was a lot, like having a period, but it came right out like it was leaking. I immediately rushed to the bathroom and changed the pad again, like that*." (Unknown age and pregnancy duration, MA with hotline support, did not seek follow-up care*)*.

Participants overwhelmingly used the number and frequency of pads to quantify their bleeding; their perceptions of whether they were bleeding "a lot" or "not a lot" was often tied to the amount of pads they were using, and how it aligned with the number of pads

they might typically use during their menstrual period. Participants reported changing their pads anywhere from every 20 minutes to every 2 – 3 hours when their bleeding was the heaviest.

Participants reported different thresholds for when they would decide to change their pads, related to the consistency of their bleeding, the quantity, and the amount of resources that they had. Some participants reported changing their pads frequently, even if the pad wasn't soaked, particularly if the bleeding was "clotted" or "thick": "*I changed the pads 3-4 times a day, and it wasn't because the sanitary napkins were full, but because I wanted to keep my hygiene, just like regular menstruation*."(Unknown age and pregnancy duration, MA with hotline support, sought care to confirm completion, ultrasound failed abortion, repeated MA procedure successfully with hotline support and did not seek additional care). Others reported changing their pads only after their pads had been "soaked": "*I changed as often as [every] 30 minutes … [the pads] were soaked when I changed it*."(28 years old, 4 weeks, MA with hotline support, sought care at laboratory to confirm completion).

For some participants, management of bleeding was related to worries about disclosure and others finding out about their abortion:

> *"Normally during the day my husband is not at home so I had no problem in managing [the bleeding], but* ***at night I had to sneak into the bathroom*** *and change [cloths/pads] secretly because I did not want my husband to notice me."* (27 years old, 8 weeks, MA with mobile clinic, did not seek follow-up care)

For some, bleeding was often an emotional ordeal. A few participants described feeling worried about the amount of blood, often stemming from a sense of confusion about what was happening to them and whether it was normal: "***I was scared that my blood was finished.*** *When I lost that much blood I was worried that my blood was finished."* (26 years old, unknown pregnancy duration, MA from unknown provider, sought care at clinic due to ongoing bleeding and received a surgical intervention).

<u>Bleeding expectations and preparedness</u>

Participants' levels of preparedness and expectations around the bleeding they experienced varied. Factors that helped participants felt prepared included making sure they had purchased enough pads or gathered enough cloths before their abortion to help absorb the blood flow: "*I prepared a lot of pads, three packs; it helps because* ***it will be troublesome if I have to look for it [later].****"* (Unknown age and pregnancy duration, MA with hotline support, sought care to confirm completion, ultrasound showed they were still pregnant, repeated MA procedure successfully with hotline support and did not seek additional care). Other factors that helped participants feel prepared were support from family and friends, childcare arrangements, and plans to seek care if they needed additional support.

Participants' feelings of preparedness also seemed to be tied to the information they received from their providers and expectations that they had for their bleeding, whether

they came from clinical or non-clinical sources. For example, when participants were informed about the level of bleeding to expect from a provider or accompaniment group, and when this information aligned with their experience, they often reported feeling prepared to manage their bleeding, even while also simultaneously feeling "worried":

> *"**I was worried** because I have not had such bleeding before, for six hours I was just going to the toilet, changing my pad, doing replacements… **[but] that was what [the hotline counselor] said I should expect**, she talked about heavy bleeding, when I asked her how much I will bleeding she said… it will not be normal like the like the monthly flow."* (21 years old, 5 weeks, MA with hotline support, sought care at laboratory to confirm completion, no other treatment received).

Some participants described feeling unprepared for the bleeding that they experienced. Participants spoke about not having enough information about what might happen or how they should deal with it, and what was normal:

> "***I did not realize it would be so difficult**… I did not think such bad days would come. I thought it would be easy…I got the medicine. He also made it sound easy. I took it thinking it will be alright. Later I faced such problems."* (20 years old, unknown pregnancy duration, SMA with pills from pharmacy, experienced heavy bleeding and sought care at clinic due to concerns around completion, received dilation and evacuation procedure).

Experiences with pain and cramping

Participants frequently described the pain they experienced with words like "cramping," "twisting," and "squeezing," Some tied their experience of pain directly to experiences of contractions or cramping based on the timing of their medication doses and the passage of products of conception. Others additionally described pain felt from "bloating" or diarrhea, likely due to side effects of misoprostol. Participants were frequently unable to distinguish between stomach and uterine pain, particularly as diarrhea was often experienced in tandem with heavy bleeding and passage of blood clots.

As with bleeding experiences, participants described a range of pain experiences. For many, pain and cramping came in waves prior to passing the products of conception or each set of blood clots. A few (n = 5) participants described it as similar to "labor pains" or their experiences giving birth. Some participants (n = 12) explicitly described their pain as worse than menstrual cramping, at a level they had never felt before: "*I've always gotten period cramps, every month, but no, that doesn't compare to this at all….I didn't know just how bad it would be, and it was even worse than I expected.*" (23 years old, 15 weeks, MA with hotline support, sought care because concerned about cramping, had ultrasound and was told to continue taking ibuprofen). Six participants, however, found that some of their experiences with pain and cramping was aligned with their experiences during their menstrual period: "***The pain was like when I have my menstruation**. I have the same cramps when I have the menstruation and it was more or less the same, only perhaps it was one level higher so it wasn't that painful*." (Unknown age or pregnancy duration, MA with

hotline support, sought medical care due to heavy bleeding, received dilation and curettage procedure and blood transfusion). Another said, "*The pain was like an 8, almost 9, but in any case yeah, I pretty much had intense menstrual cramps, really intense, but they were just menstrual cramps*," (26 years old, 6 weeks, MA with hotline support, didn't seek follow-up care).

Only five reported not experiencing any pain or experiencing minimal pain; however, two of these individuals ultimately received a surgical intervention due to an incomplete or failed abortion after using medications. Nearly half (n = 29) reported experiencing severe or extreme pain, and used words such as "severe," "sharp," or "stabbing" to describe their pain.

Almost all participants described intensive cramping and pain immediately prior to the start of bleeding, usually within 10 – 30 minutes of taking the medication, and tied to the passage of products of conception or blood clots: "*"**Whenever the product wanted to come out I feel the pain**. But when the product finished coming out, I did not feel the pain again."* (18 years old, 8 weeks, MA from mobile clinic, sought care because of pain and to confirm completion, received pain medication).

Pain management

Participants described a variety of strategies for dealing with the pain they experienced from cramping. Many participants in all contexts reported using painkillers, such as paracetamol or ibuprofen, hot water compresses, resting or lying down, or trying to stay distracted.

Most participants who took pain medication reported taking them once they started experiencing painful contractions after each dose of medication; after taking medication, the pain would subside briefly but would often immediately resume after the next passage of blood clots/bleeding. Most participants reported being told by clinic or accompaniment group staff to purchase painkillers in preparation for their abortion experience, or had been provided painkillers by their service provider in preparation.

A minority of participants described not doing anything to manage their pain—their decisions were overwhelmingly related to worries around disclosure, as well as fears about the effectiveness of medication being affected if they took something else: "*I was able to manage the pain because **I did not want my secret to be exposed, so I tried to be strong**."* (26 years old, 4 weeks, MA from mobile clinic, sought care at pharmacy for vaginal discharge, received oral antibiotics).

Expectations of pain

Participants frequently reported that while they knew to expect pain and cramping as part of their process and had pain-relieving medications or strategies at the ready, they were not prepared for the intensity of the pain they experienced during their process, and feeling fearful or scared: "*I did not feel prepared for the cramping, the pain was too much….I didn't*

know it would be too tense, I didn't know it was going to last for that hour, ***I didn't know it was going to pain me like that.****"* (35 years old, 5 weeks, MA with hotline support, did not seek follow-up care). As another stated, "*I never experienced this type of cramping before*." (25 years old, unknown pregnancy duration, MA from clinic, sought follow-up care for IUD placement, told she had an infection and retained tissue and given additional dose of misoprostol and oral antibiotics).

As with bleeding experiences, some participants cited concerns about whether the level of pain they were experiencing was normal:

> "*But because I also didn't know…I end up worrying every time I felt a massive pain. It was only in the beginning when the pain went more; I would think 'Should I go to the hospital? Should I go?' But I held up. So, I was confused, '****Is it normal? Why is this so painful?*** *But what if I go to the hospital and it turns out that this is nothing?' So I just held up."* (Unknown age and pregnancy duration, MA with hotline support, sought care to confirm completion, ultrasound showed they were still pregnant, repeated MA procedure successfully with hotline support and did not seek additional care).

However, for the few participants who described their experience as "not painful" or similar to menstrual pain; these participants described their experiences as in line with their expectations, or even less than what they had anticipated.

Experiences with side effects

Fever, nausea, and diarrhea were the most commonly reported other side effects experienced by participants. Most of those who experienced fever described it as their body feeling "hot," "shivering," or feeling "cold;" for most individuals, the timing of this symptom began a few hours after taking their first misoprostol dose, and continued for around one day. Most participants did not use a thermometer to measure their temperature, and managed this symptom with paracetamol or other fever reducers, typically in concert with pain relief medication.

Nausea was reported by nearly half all participants, also typically starting a few hours after the misoprostol dose; some participants described vomiting due to the nausea, while many mentioned loss of appetite as a result, often for several days.

A total of 24 participants described experiences with diarrhea. These experiences were closely tied to participants' experiences with cramping and heavy bleeding, as many participants described the period in which the cramping was the most intense as triggering both gastro-intestinal distress along with the passage of blood clots. Among those who experienced more severe diarrhea, it was difficult to quantify their blood loss in the time periods when they were using the toilet.

Vaginal discharge was reported by half of participants. Participants typically reported discharge that lasted a few days to one week after their abortion, and described the discharge as "slippery," "dirty water," "white," and often malodorous.

Deciding whether to seek medical care

A total of 32 participants reported seeking follow-up care, six of whom had been recruited based on presenting for post-abortion care at one of the study sites.

Participants who sought additional medical care did so for three primary reasons: to confirm abortion completion (n = 12), worries about the level of bleeding, pain, or discharge they were experiencing (n = 14), and/or a lack of abortion symptoms (n = 7). Underlying all reasons for seeking care was a desire for reassurance that their process was complete and their abortion experience was typical, health-related concerns, or guidance on what next steps to take. A total of 11 participants received a surgical abortion at the health facility to treat a failed or incomplete abortion.

One participant who sought care due to extreme bleeding described:

> "*Because at that time the bleeding was too heavy, I was not afraid of the abortion but I was afraid what if I ran out of the blood. And my head was very dizzy, so at 10 or 11 PM I decided to go to the hospital, and yes at that time I already used the new sanitary pad and then one hour after that when I reached the hospital, it was already full*." (Unknown age or pregnancy duration, MA with hotline support, sought medical care due to heavy bleeding, received dilation and curettage procedure and blood transfusion).

However, many participants who sought care did so to confirm abortion completion or for management of less severe abortion symptoms; some did not disclose their abortion to the provider they saw. For example, some participants sought care at labs/diagnostic centers to confirm they were no longer pregnant via either blood test or ultrasound without disclosing anything to the provider/technician, and presented as "just wanting to know if they are pregnant." Others were also able to access the care or treatment they needed without disclosing their abortion: they describe obtaining blood booster tablets, pain medications, antibiotics, or allergy medications to treat allergic reactions to misoprostol directly from pharmacists, often after phone consultation from an outreach team. Among the majority of participants who did not have any form of follow-up care, some described not seeking care because they were confident in the information they had received from their medical provider or lay accompanier/hotline counselor:

> *"**If got panicked, I would refer to the information that was given by the provider**… So, if I suddenly got panic, I remember, I just read counsellor's chat, to find out whether what's happening is normal or not, like that."* (Unknown age and pregnancy duration, MA with hotline support, didn't seek follow-up care)

For most, they did not seek medical care because "***there was no serious symptoms****, there was no serious issue that came up that would warrant me seeking for care."* (29 years old, MA with hotline support, didn't seek care).

However, a few reported not seeking medical care because of fears due to restrictive legal contexts, fear of disclosure, even when experiencing bleeding or pain that concerned them. One participant describes waiting for her symptoms to subside:

> "*I couldn't stand it anymore, because the pain was unbearable, the bleeding was pretty heavy, so I really wanted to go to the hospital. I tried to stand up; my companion wanted to take me to the hospital. Then, when I was standing, that biggest lump came out. The blood was pouring out, the bleeding was so heavy. Then the pain was gradually diminished,* ***so I didn't go to the hospital, even though I wanted to.****"* (Unknown age and pregnancy duration, MA with hotline support, sought care to confirm completion, ultrasound showed they were still pregnant, repeated MA procedure successfully with hotline support and did not seek additional care).

Another participant described how her decision to not seek care was related to fears around disclosure of their abortion: "*"Because if I sought any additional care,* ***they will ask me what is wrong with me*** *and they expect me to explain everything*,"(26 years old, 18 weeks, MA from mobile clinic, did not seek care), while another described her fear that additional treatment doing something that would interfere with the abortion process: "*"I endured all the symptoms, because* ***I don't want to take any medicine and the abortion stops."*** (35 years old, 8 weeks, MA from mobile clinic, did not seek care).

1.5 Discussion

Our study found a wide variation in the way that participants described their experiences with bleeding and pain. Though many participants described intense experiences with bleeding or pain, these experiences were not necessarily tied to whether or not an individual sought medical care. There was a range of reasons why people sought care, including concern about the level of bleeding or pain or wanting confirmation that their abortion was complete, or management of side effects. Participants who arrived at healthcare facilities knowing exactly what medical treatment that they wanted (i.e., antibiotics, a pregnancy test, blood boosters, pain killers, antihistamines) and who were prepared to make a specific request, more commonly received the services that they needed without a need to disclose the abortion to the provider. Our findings highlight the challenge in objective measurement of these experiences based on self-report; we discuss implications for measurement below and possible solutions below.

Our analysis suggests that what participants may consider to be "normal" bleeding may be directly related to their expectations around the amount of bleeding, the preparedness and information they have received from clinicians or other sources of support, and their typical menstrual cycle. Participants had different perceptions of what constituted "heavy," "light," or "normal," bleeding, often related to their normal menstruation experiences, the

different types of materials used for absorption/management strategies, as well as the intervals in which they assessed their bleeding. Perceptions were also related to their expectations around the amount of bleeding, and the preparedness and information they have received from clinicians or other sources of support. As a result, using descriptors such as "heavy," "light," or "normal" may not have particular utility in a self-report instrument, clinical chart, or assessment from a call center staff or accompaniment support person in determining whether an individual needs follow-up care. Rather, asking about how frequently patients or callers are changing their pads/menstrual hygiene products, the type/thickness of these products, and the state of these products when they are changed ("soaked") may be a better alternative to identify instances of heavy bleeding. Additionally, directly asking how experiences compare to typical menstrual periods may offer additional details on bleeding experiences.

Similarly, participants' descriptions of pain and cramping also differed based on individual past experiences, with participants benchmarking their pain to either typical menstrual pain/cramping, or labor pains for those who have had an experience of giving birth. Participants who used medication abortions to end their pregnancies often reported experiencing extreme pain; often directly tied to the timing of the passage of blood clots or products of conception; many cited pain as a reason for why they contemplated or sought clinical or additional care. Most participants described their experiences with both bleeding and cramping during a medication abortion as substantially greater than what they experience during a typical period; thus, descriptions or preparation for the process as "like a heavy period", may understate individual's true experiences and leave them underprepared for the abortion process.

Furthermore, particularly in contexts where abortion is legally restricted or highly stigmatized, participants commonly reported being afraid to seek care, or experiencing fear when seeking care. Call center staff, accompaniment groups, as well as mobile outreach services are uniquely situated to be able to prepare and support those having abortions in this type of care seeking, particularly in contexts or situations where they are unable to access a trusted provider. For example, providing assurances that they do not need to disclose their abortion to a provider, particularly if they are seeking treatment for pain, mild infection or vaginal discharge, or confirmation of abortion completion; reminding patients or callers that symptoms of medication abortion are identical to symptoms of miscarriage; and helping them assess the severity of their complications in order to provide them options for sources of medical care. Future studies should seek to assess whether these strategies improve individual's abortion experiences particularly as it relates to receipt of high-quality, timely, and desired medical treatment.

As with most qualitative research, this study is not intended to be representative of all medication abortion experiences, but rather sought to capture the breadth of ways in which people describe their experiences with medication abortion in order to identify important domains for measurement of potential complications. As participants were purposively recruited into the study based on their experience with health care seeking and report of certain physical abortion symptoms, we cannot draw any conclusions about how common the experiences of bleeding, cramping, and health care seeking experiences

may be among medication abortion users outside of this study. Future research should explore how descriptions of bleeding are related to different cultural/contextual factors, and how duration of pregnancy, prior pregnancy experience may be related to amount of bleeding.

1.6 Implications for measurement

Based on our analysis of the in-depth interviews and individual discussions with experts, three broad domains that encapsulate the overall medication abortion experience were identified: individual experiences related to the abortion process, experiences in the healthcare system, and the final abortion outcome (Figure 1). We discuss each domain and implications for measurement below.

Individual experiences

Within the domain of "individual experiences," we identified the following measurement sub-domains for common abortion symptoms (for example, bleeding, cramping, fever) that would be important to include in a quantitative self-report instrument: 1) amount or intensity, 2) management, 3) duration, 4) type or features, 5) worries or concerns, 6) physical impact, and 7) support.

Proposed measurement sub-domains of amount or intensity, management, and duration directly stemmed from our findings that participants used different types of materials for absorption or management of their bleeding, and described their amount of bleeding in different increments (for example, on an hourly basis, on a daily basis, by number of pads used, by the strength of the flow. Based on discussion with experts on potential strategies for reconciling these different management strategies and descriptions of bleeding into a potentially standardized, comparable self-report measure, we developed the idea to link blood loss to symptoms that might be indicative of a complication. For example, rather than having a self-report instrument that focuses exclusively on quantifying blood loss, we might seek to measure the effects of sustained blood loss – such as lightheadedness, fainting, malaise, or other symptoms consistent with the consequences of sustained blood loss—this is reflected in the proposed measurement domain as "Physical impact." However, we note that experiencing light-headedness may also be a side effect of seeing blood or being concerned about bleeding; this should be further explored.

From these discussions, we also developed the idea to link complications, medical problems, and/or outcomes that are of greatest concern, such as infection due to retained products of conception, failed abortion, incomplete abortion, to specific symptoms that would differentiate these experiences from pain, cramping, and other symptoms that are commonly experienced in a medication abortion process. For example, fever is a commonly experienced side effect of misoprostol, so rather than asking if they experienced a high fever, fever indicative of an infection that would need further treatment might start more than 4 hours after the last misoprostol dose, and be sustained over time (reflected in measurement sub-domains "Duration"). Similar challenges in differentiating worrisome discharge from "typical" discharge were discussed, especially since even discharge without

cause for concern may carry an odor, or be of a different consistency that the individual typically experiences. Given that most individuals who have a medication abortion report experiencing pain, typically tied to the cramping or contractions leading up to and following expulsion of the products of conception, understanding and differentiating pain due to cramping from pain due to a possible infection is important in any self-report instrument. A recommendation was to focusing on pain that was localized, tender, and did not abate with passing products of conception (reflected in measurement sub-domain "Type/features").

Finally, given that many of these symptoms are non-specific, an additional recommendation was to consider markers of possible infection (pain, discharge, fever) together, to try and understand what constellation of symptoms are an indication for needing antibiotic treatment or additional medical care.

Health care seeking behaviors

Based on our analysis of the in-depth interviews, we identified several factors influencing whether individual may seek health care during their abortion process. While reasons for seeking care are largely driven by their individual experiences with bleeding, cramping, and other symptoms, factors such as support or information received from others, availability of providers, cost, worries about legal risk, stigma, or judgement, and worries about medical procedures all play a role in the decision or ability to seek health care. Discussions with experts further emphasized the multitude of drivers for care seeking, and there was universal agreement that care-seeking or receiving additional medical treatment during the course of a medication abortion was not necessarily indicative of a complication. Proposed sub-domains of measurement around health-seeking experiences involved linking the type of treatment received to the reason why an individual was seeking care: whether the treatment the individual received was related to confirmation abortion completion, treatment related to a true medical complication, and treatment related to symptom management.

Abortion outcomes

"Incomplete abortion," broadly defined as the partial loss of the products of conception was identified by experts as an outcome classification that deserved further scrutiny. Incomplete abortion, or treatment for incomplete abortion, is often used as an indication of an "adverse event" or "abortion morbidity" in much of the abortion complications literature.(78, 79, 83) However, given the known and expected level of effectiveness of medication abortion (defined in clinical studies as complete abortion without surgical intervention) is less than 100%,(68, 69) treatment or management for incomplete abortion or failed abortion should be considered part of the overall abortion process, rather than an indication of an adverse event, particularly since the counterfactual (if, without intervention, the abortion would resolve on its own or lead to further problems) is unknown. Rather, a suggestion was to reframe these experiences as "outcomes" of medication abortion. However, these experiences may progress to a "complication" if left untreated, such as infection from retained products of conception.

There was some discussion around difficulties in identifying whether or not a treatment or intervention was "necessary," with emphasis on the importance of remembering that the ultimate goal in abortion care is that the person is no longer pregnant, experienced no morbidity or long-term health impacts, and they received the medical care that they wanted or needed in a timely manner (Figure 1). Experts encouraged the use of a framework that centered the person having the abortion, rather than one that exclusively focused on identifying whether treatment was "necessary" from a clinical standpoint, to a person/patient-centered standpoint. Even if a surgical intervention may not be strictly "necessary,"—for example, even if a person would have a complete abortion if they continued to wait, or if they had an additional misoprostol dose—if the person found their bleeding and pain problematic enough to want to seek care, and wished for their process to be over, we may view their abortion as a "success" if they received a surgical intervention when they wanted it. For these individuals, while waiting for the bleeding to resolve on its own or taking additional doses of misoprostol may result in a complete abortion without the need for additional medical treatment, the overall experience might be a not preferred experience given the level or duration of bleeding or pain.

What is a complication?

Our analysis of the in-depth interviews and discussions with experts elucidated the inherent challenge of defining a complication. Experts raised concerns with studies or frameworks that use the term, "complications from unsafe abortion," due to the importance of disentangling the idea of what is a "complication" from conceptualizations of safety, which, based on the World Health Organization safety framework, is defined by the method and the setting, rather than the outcome. While the inherent link between measuring "complications" and expanding the definition of what type of providers fall under the "safe" versus "less safe" categories was clear, we determined that a new framework is needed for thinking about care-seeking, treatment, and outcomes of medication abortion in a context where people self-manage their abortion but seek follow-up care in the formal healthcare system.

Overall, this work demonstrates the inherent challenges in measuring or defining complications from medication abortion, and suggests focusing efforts towards a patient-centered framework that prioritizes assessing desired outcomes and experiences. Furthermore, there is a need to consider failed or incomplete abortion outside of the complication framework, particularly in the context of medication abortion, and along with the importance of considering the individual's desires, treatment, and context in measuring outcomes from abortion. A patient-centered approach to measuring individual experiences might triangulate life-preserving treatment, the individual's desired treatment, and the medical treatment they actually received when classifying an experience as a potential "adverse event" or "complication." Figure 2 summarizes this perspective.

1.7 Conclusion

Participants describe a range of experiences in bleeding, cramping, pain, and reasons for seeking medical care. A self-report instrument is needed that centers the individual's

perspective and needs, and shifts away from a strict biomedical frame focused on complication "severity" to one that centers abortion "quality" and outcomes. Complications from medication abortion (hemorrhage, infection) have clearly defined clinical signs as demonstrated by the abortion complications literature. However, these classification schema conflate other undesired outcomes, such as ongoing pregnancy, or incomplete abortion, with true medical complications, and often use health care seeking as an indication of a complication. Reframing medical treatment as a positive outcome (if the person receives the treatment they wanted in a timely manner), rather than considering it as an indicator of a "complication," may help us shift to a more patient-centered consideration of abortion outcomes and care.

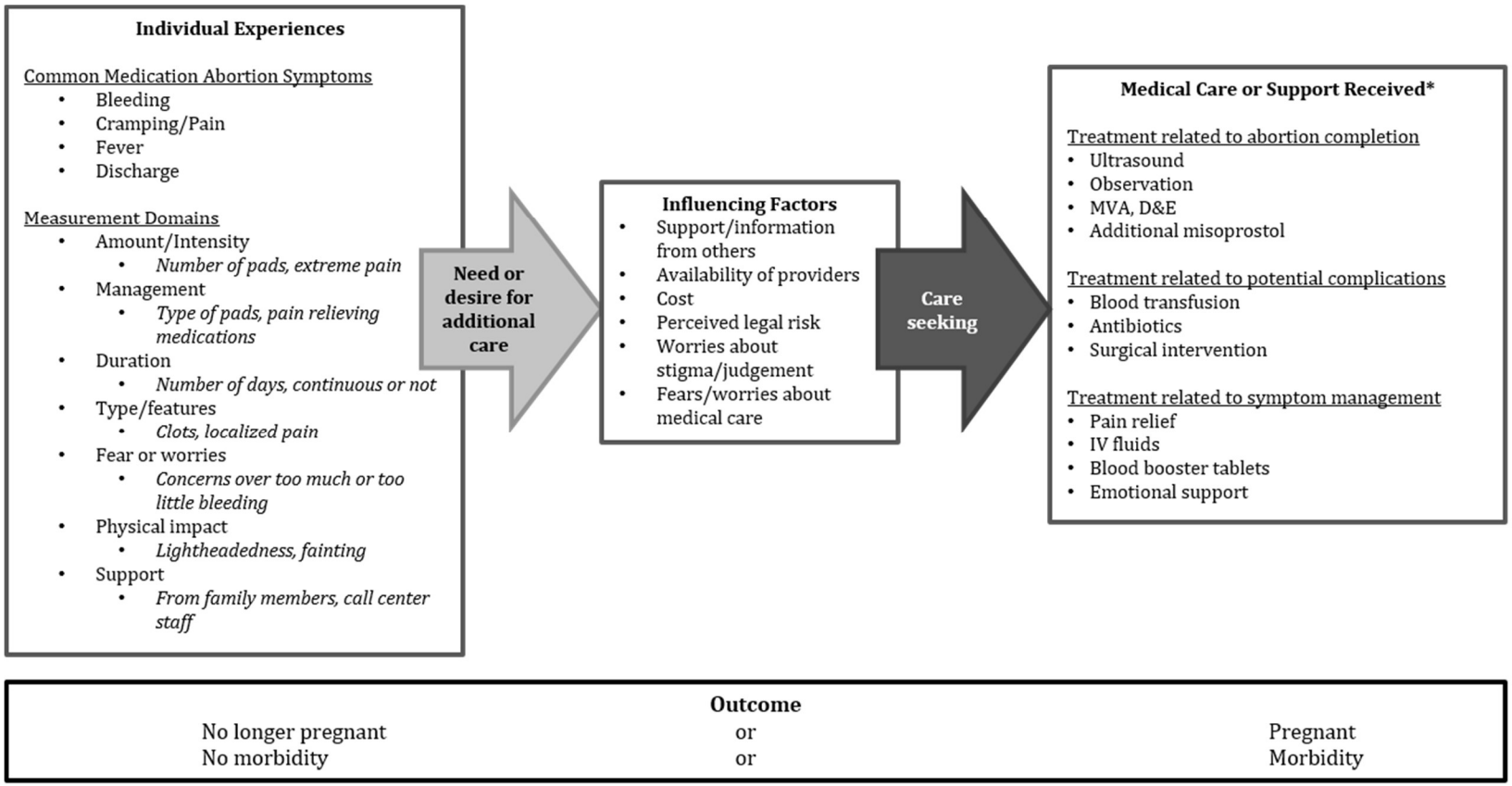
We propose a new framework that reserves the use of the term "complication" for moderate or severe unanticipated problems that arise following a treatment, procedure, or condition, and shifts our focus on measuring outcomes from medication abortion to measuring other dimensions of quality that center the needs and desired experiences of individuals.

1.8 Tables and Figures

Table 1. Characteristics of participants in a qualitative study on medication abortion experiences, N = 68.

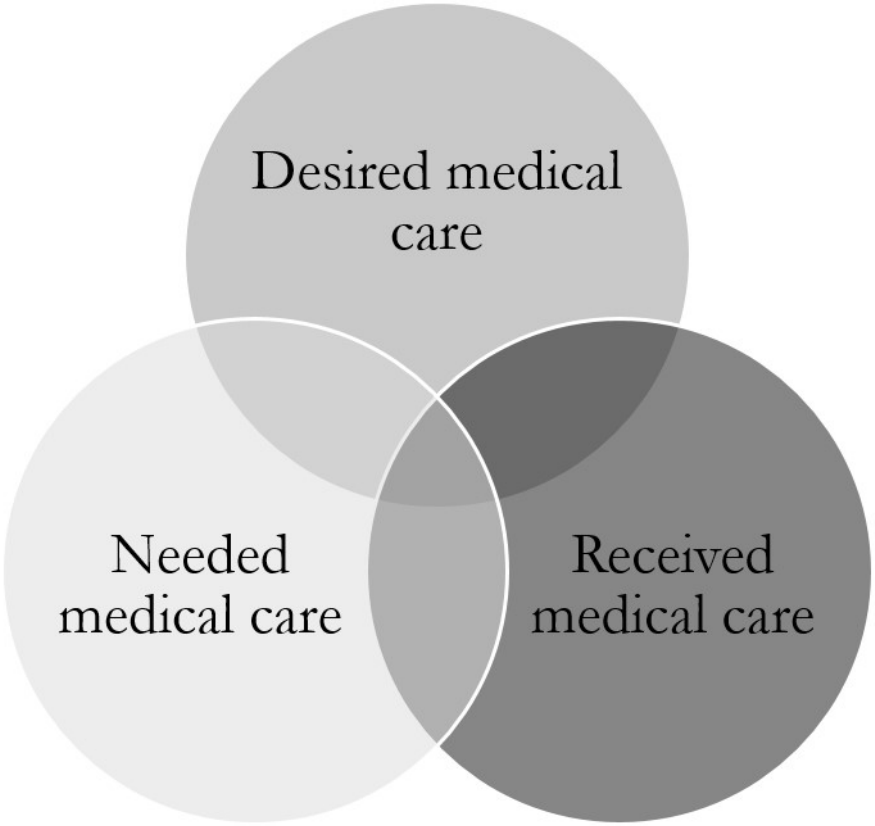
Characteristics	N
Initial abortion experience	
Clinic-based medication abortion	36
Self-managed medication abortion	32
Method of recruitment	
Private clinic/mobile clinic	36
Hotline callers	26
Private clinic after post-abortion care	6
Age	
18 - 24	18
25 - 34	36
35 - 44	5
Unknown	7
Pregnancy Duration	
< 7 weeks	25
7 - 9 weeks	14
10 - 12 weeks	4
12 - 15 weeks	4
15 + weeks	2
Unknown	19
Reported any previous experience with abortion	8
Abortion Outcome	
Complete without surgical intervention	57
Complete with surgical intervention	11
Reason for any care-seeking	
Didn't seek care	34
Sought care to confirm completion	12
Sought care for symptoms	21
Unknown/other reason	1

Figure 1. Dimensions of measuring medication abortion experiences, follow-up care, and outcomes.



*Type of treatment under each category are examples of each domain and not a comprehensive list

Figure 2. Challenges in conceptualizing medication abortion follow-up care.



One component of a "good" abortion is one where the individual's desired medical care, needed medical care, and the medical care received are in alignment.

Desired medical care that is received, even if not life-preserving or medically "necessary" (for example: receiving pain medications, or a MVA to treat an abortion in process with prolonged bleeding) may also be considered a core component of "good" or "high-quality" abortion care.

Any medically necessary care that is not received (regardless of whether or not it is desired) may increase risk of abortion morbidity; conversely, receiving medical care that is not needed nor desired (for example, a D&E for an abortion in process, when the individual would have preferred additional misoprostol) indicates a lack of patient-centered care.

Additionally, it is possible that some may receive medically necessary care that was not desired, highlighting the challenges and nuance in conceptualizing and measuring individual decision making and preferences regarding abortion care.

2 Chapter 2. Respondent- Driven Sampling to Assess Cumulative Lifetime Incidence of Abortion in Soweto, South Africa: A Methodological Assessment

2.1 Abstract

There are many challenges to accurate estimation of abortion incidence. Respondent-driven sampling (RDS), a methodology that relies on peer-to-peer recruitment, may be a potential strategy for abortion incidence estimation. However, the applicability of RDS estimates as representative estimates rely on several, often untested, assumptions. We conducted a RDS study to estimate the cumulative lifetime incidence of abortion among women of reproductive age in Soweto, South Africa. We recruited 849 eligible participants from 11 seeds (initial recruiters) from April 2018 – December 2018 in Soweto, South Africa. All participants completed a baseline questionnaire and were given 3 recruitment coupons to distribute to eligible participants in their social network. A subset of 289 participants completed a follow-up survey about their recruitment experiences. We assessed several key RDS assumptions, and compared RDS estimates of abortion incidence to estimated abortion incidence adjusted for employment and age, based on available census data.

We found that the following RDS assumptions were not met: accurate reporting of degree, and random recruitment within an individual's social network. Failure to meet these assumptions yielded a sample of participants with demographic characteristics that were substantially different from the target population in terms of employment status, and was not resolved by standard RDS weighting methods. As abortion experiences are different across employment status, the RDS estimate of abortion incidence from this study is likely biased. This study presents a unique and rare opportunity to rigorously assess the assumptions of RDS in a non-hidden population. We caution researchers in the use of RDS as a method to generate representative estimates of abortion incidence; use of post-survey weights to adjust for differences in demographic characteristics between the sample and target population rather than RDS weights based on network size may generate more representative results. Future research should explore the estimation of post-survey weights in populations where census data do not exist. However, peer-to-peer recruitment methods remain a promising strategy to generate an inclusive convenience sample of participants who may be excluded from research.

2.2 Introduction

Unsafe abortion is a completely preventable cause of maternal mortality and morbidity worldwide,(5) yet is responsible for an estimated 8-13% of global maternal deaths.(3) There are substantial challenges to the accurate estimation of abortion incidence. Reporting of abortions that occur within the formal healthcare system is often incomplete or inaccurate, and there is a lack of individual-level data on abortions that occur outside of the formal healthcare system. Indirect methods for estimating the incidence of abortions rely on assumptions and extrapolations that are often difficult to test.(39, 40) Due to social and cultural stigma, and fear of legal consequences, people who have had abortions outside of the formal healthcare system may be reluctant to seek care in the event of complications,

and are thus excluded from facility-based estimates of abortion incidence based on the rate of care-seeking. If they do seek care, people who have had abortions may also be reluctant to disclose their abortion—or may intentionally report a miscarriage-- and their experiences may be recorded as such.(34-36, 38) Recent evidence suggests that nearly half of all abortions that occur globally are performed outside the formal healthcare sector.(3) However, study recruitment of only those who seek medical care, along with under-reporting, likely lead to underestimates of abortion incidence, as well as biased data on the characteristics and outcomes of abortion. This bias may be amplified in settings where the majority of abortions occur outside of facility-based settings. In settings where abortion is legally available but barriers to access exist, accurately measuring the incidence of abortion, and understanding the proportion of abortions that happen outside of facility settings can help to identify gaps in the accessibility of clinic-based abortion services, and help facilitate the development of interventions to address issues around access to and quality of care.

Respondent-driven sampling (RDS) has been proposed as a potential solution to the above challenges in abortion incidence estimation.(84) RDS is a methodology that was developed to more accurately estimate the prevalence and incidence of sensitive and illegal behaviors among hidden populations such as injection drug users, sex workers, and men who have sex with men.(47, 49-54) RDS leverages a small non-random sample of initial participants (known as *seeds*) within social networks engaging in hidden or stigmatized behaviors to recruit others within their social network from a specific target population.(49, 50, 54) As the initial seeds are not randomly selected, participants who are well-connected may have more success in recruiting, and individuals are more likely to form social ties with those who have similar characteristics, RDS samples are likely not representative. However, RDS estimation methods account for this potential selection bias by weighting participants with more contacts in the target population inversely proportional to the number of contacts in the network itself.(49, 50, 55)

The validity of these estimation methods rely on several key assumptions around recruitment dynamics.(49, 54, 85) However, despite the rapid proliferation of the use of RDS as a sampling approach, these assumptions are rarely, if ever, rigorously assessed in RDS studies.(86) Furthermore, given that these studies are typically conducted among hidden population for which no sampling frame exists, validation of whether these estimation methods yield a representative sample are often impossible to assess.

To address this research gap as well as address some of the above-identified measurement and recruitment challenges in abortion incidence estimation, we conducted a respondent-driven sampling (RDS) study to estimate the cumulative lifetime incidence of abortion among women of reproductive age in Soweto, South Africa. While RDS has traditionally been used to measure outcomes among a stigmatized population, to our knowledge, this is the first study that uses RDS to measure abortion (a stigmatized outcome) among a general population. However, we believe that this question is well-suited to RDS for several reasons. Population-representative surveys, such as household surveys, may exclude young women, those who are living in informal settlements or inconsistent/variable living conditions, or refugees. Furthermore, those who participate in any study may be likely to

underreport their abortion experiences.(36, 87) RDS, by leveraging peer to peer recruitment, may reach a broader population than traditional research methods. Additionally, the process of being recruited into the study by someone known to the participants may generate trust between the recruiter and the researcher and encourage disclosure of sensitive experiences, though this has not been assessed.

RDS holds potential as a strategy for more accurate estimation of abortion incidence in a general population. However, rigorous evaluation of whether the assumptions of RDS have been met is warranted. The aim of this paper is to assess whether the above assumptions of RDS have been met in a study of abortion incidence in Soweto, South Africa, and provides generalized information on what the impact failing to meet underlying assumptions has on RDS estimates.

2.3 Methods

Study participants

The study was conducted from April 2018 – December 2018 in Soweto, South Africa. Participants were eligible if they were between the ages of 15-49, lived in Soweto, spoke one of the study languages (English, Tswana, isiZulu, Sotho, or Xhosa), had not already participated in the study, and had a valid recruitment coupon.

Procedures

With the assistance of community partners, we identified eleven initial seeds across a range of ages, socioeconomic statuses, and prior experience with abortion. After obtaining verbal consent to participate, seeds completed an interviewer-administered baseline questionnaire at the study site with questions on sociodemographic characteristics, social network size and composition, health behaviors, knowledge about sexual and reproductive health, and abortion experience. After completing the survey, each seed was provided three recruitment coupons and instructed to provide them to eligible participants in their social network. Seeds comprise the first study *wave*, those recruited by seeds make up the second study wave**,** and those recruited by the seed's recruits comprise the third study wave, and so on.

Interested participants contacted the study phone number; study staff confirmed eligibility and scheduled an interview time. Participants provided verbal consent at the study site prior to completing the baseline questionnaire, and were provided with three recruitment coupons. Participants and seeds received an incentive of R75 (USD 6) for completing the baseline questionnaire, and a secondary reimbursement of R50 (USD 4) for each successfully recruited study participant.

All participants were instructed to contact study staff within 4 weeks of participation to schedule a follow-up interview and receive their secondary incentive; those who reported an abortion in the baseline questionnaire were contacted and asked to complete an additional abortion follow-up questionnaire about more detailed abortion experiences. The

recruitment follow-up included questions about experience with recruiting, including details about recruitment attempts, reasons for coupon refusal, reported relationship to those they distributed coupons to, and additional questions about network size. Study questionnaires were developed based on findings from formative qualitative work,(88) recommendations from the literature,(84, 85) and piloted prior to implementation.

Ethical approval for the study was obtained from the Human Sciences Research Council (Pretoria, South Africa).

Sample size

We determined the minimum required sample size based on the method proposed by Salganik.(89) Given that we lack accurate estimates of the prevalence of informal sector abortion in South Africa, we chose a maximally conservative prevalence estimate of 50%. To be able to detect a 50% lifetime prevalence of informal sector abortion, with 80% power, with absolute precision of 3%, 95% confidence intervals, and assuming a design effect of 3, we aimed to recruit a sample size of 900 women to participate in the study, conditional on the distribution of selected socio-demographic characteristics becoming similar across waves (equilibrium).

Analysis

We assessed the following five assumptions of RDS based on methods proposed by Gile et al.(85) First, the population being recruited must be able to identify those in their social network as members of the target population (in this study, women of reproductive age), and form reciprocal social ties on the basis of this shared characteristic. Second, the final composition of the sample should be independent of the initial seeds. Third, sampling should replicate sampling with replacement. Fourth, participants can accurately report their network size. Fifth, participants should randomly sample their recruits from within this personal network.

To assess whether the population being recruited was able to identify those in their social network as members of the target population, we assessed the reported relationship between participants and their recruiters, and whether the participant reported they knew their recruiter, and would have recruited them (reciprocity of network ties). If participants report being recruited by their friends, family members, or others in their social network, this would indicate that individuals are able to identify members of their social network who are eligible for participation in the study. Participants reporting being recruited by strangers or by those who they wouldn't have recruited would indicate that this assumption was not met.

To assess whether the final sample was independent of the initial selection of seeds, we assessed homophily and bottleneck and convergence plots for key sociodemographic characteristics and abortion incidence. Homophily is a measure of preference for recruiting others who have a similar shared characteristic (i.e., age).(54) Homophily values of 1 indicate that the number of recruiter-recruit pairs with the same characteristic are similar

to what we would expect due to change; a homophily value of 1.6 would indicate that there are 60% more homophilius pairs than we would expect due to chance. The presence of homophily may indicate that the initial selection of seeds could play a role in the final sample composition. Bottleneck plots assess whether there are large differences in the characteristics of the sample by seed. Convergence plots assess whether sociodemographic characteristics are converging on a stable estimate. The lack of convergence may indicate that the initial selection of seeds may still be influencing the estimate, and would indicate the need for additional data collection.

To assess whether sampling approximated sampling with replacement, we assessed failed recruitment attempts due to previous participation in the study (i.e., whether potential recruits declined a coupon because they are already participated). By design, repeat participation was not permitted in the study. We also assessed whether the number of people who have already participated in the study that a participant reports knowing increased over time or study wave, which would indicate potential depletion of the target population (i.e., less of the target population within an individual's social network are available to be recruited for subsequent waves of the study).

While it is impossible to validate the accuracy of reported network size, we assessed *consistency* of reported network size by asking participants to report their network size at baseline and during the recruitment follow-up, as well as plausibility of responses by identifying large outliers or peaks in the distribution of reported network size to. We examined whether self-reported degree was correlated with success in recruiting, as an assumed feature of RDS is that individuals with a larger social network will be over-represented in the sample, and thus degree is used to correct for this over-sampling. Given the non-normality of reported degree, we used the more robust Spearman's rank correlation to measure the association between self-reported degree at baseline and with the following individual measures: estimated potential number of recruits, actual number of recruits, and self-reported degree at recruitment follow-up.

Finally, we assessed whether participants randomly sampled recruits from their personal network by comparing the proportion of the sample employed to the average proportion contacts in network who are employed as reported by participants. We chose to measure employment status based on findings from a previous RDS study among a non-hidden target population that found those of lower socio-economic status were over-represented in the RDS sample as compared to census data on the overall population.(90) Given that socio-economic status of an individuals in a participant's social network may be difficult to report, we chose employment status as a proxy to assess whether those who are unemployed had greater incentive to participate in the study.

We then assessed the potential impact of failing to meet these assumptions on the estimated cumulative lifetime incidence of abortion. We calculated the following cumulative lifetime incidence of abortion: 1) unadjusted cumulative lifetime incidence of abortion, 2) RDS-II estimator (definition below), 3) RDS-II estimator excluding those without reciprocal network ties, 4) sample proportion with post-estimation weights for

selected socio-demographic characteristics (based on results from homophily and random recruitment assessment).

The RDS-II estimator(91) is calculated as the proportion of respondents who report ever having had an informal sector abortion, weighted by the inverse of their network degree size (Equation 1). 95% confidence intervals are calculated over 500 bootstrapped samples.

$$\hat{p} = \frac{\sum_{j \in I} \frac{1}{d_j}}{\sum_{j \in S} \frac{1}{d_j}}$$

Equation 1: RDS-2 Estimator; where *j* indexes the respondent, ***S*** is the set of the full sample, ***I*** is the set of respondents who have ever had an abortion, d_j is the degree.

We used two different measures of degree for the RDS-II estimator: self-reported network size was measured as the number of eligible contacts that the respondent saw in the past week, and *visibility*(92) as an alternative to self-reported degree, particularly in the presence of outliers or when there are concerns around the accuracy of self-reported degree. Visibility was imputed based on participant self-reported degree, number of recruits, and time spent recruiting, using the *impute.visibility_mle* function RDS Analyst: Software for the Analysis of Respondent-Driven Sampling Data 0.65.(93) Differences in estimates were assessed using a two proportion Z-test.

All analyses were conducted using Stata 15(94) and RDS Analyst: Software for the Analysis of Respondent-Driven Sampling Data 0.65.(93)

2.4 Results

A total of 854 participants, including 11 seeds completed the baseline questionnaire. Participant demographics are in Table 1. Data from five respondents were excluded from analysis due to participant ineligibility (age older than 49 years, duplicate enrollment) or missing survey data, resulting in an analytic sample of 849. A total of 15.7% of the sample were between the ages of 15 – 19. Nearly two-thirds (64.0%) reporting having a romantic partner, and the majority (83.7%) were students or unemployed. The most commonly spoken home language was isiZulu (38.9%) and Sesotho (25.6%).

Among all participants in the analytic sample, 358 successfully recruited at least one recruit (42.2%). Of the eleven seeds, one did not recruit any participants. Among the remaining seeds, the maximum recruitment wave was 17 after 36 weeks of recruitment; the largest number of recruits were from wave 5 (116/849, 13.7%). Over half of the sample came from one seed. Recruitment chains are shown in Supplementary Material 1.

A total of 289 participants participated in the follow-up survey on recruitment, including 77% of participants who successfully recruited and 2.2% of those who did not successfully recruit other participants. Participants in the follow-up survey answered questions on their

experiences recruiting and their relationships to 822 individuals to whom they gave coupons (Table 2).

Assumption #1: Reciprocity of ties

In the baseline survey when asked, "Would you have recruited your recruiter," 92.2% of respondents reported yes, indicating reciprocity (Table 2). Among those whose reported "no" or "unsure," most reported they would not have recruited their recruiter because they don't see them very often, didn't think they would be interested in the study, or didn't know them very well. A minority (5, 0.6%) reported they would not have recruited their recruiter because they were recruited by a stranger. In the recruitment follow-up, among the 822 reported recruit – recruiter relationships, 84% of participants reported that their recruit would have given them a coupon, while 10.5% said they were unsure and 5.5% said no. Commonly reported reasons for why they believed their recruit would not have recruited them included not being socially close to them, not seeing them very often, or being helped by others to identify them. The presence of participants who reported being helped by others to identify recruits not only violates the assumption of reciprocity of ties (as they are recruiting from outside of their social network), but also violates the assumption of accurate reporting of degree and random recruitment from within their network.

In the baseline survey, most participants reported being recruited by a friend (47.5%), neighbor, community member, or church member (together 20.0%), or a female relative (24.8%). In the follow-up survey, participants commonly reported distributing their coupons to friends (53.7%), neighbors/community/church members (19.3%), or a female relative (15.7%). While we could not validate the concordance of reported relationships, overall, this suggests that the distribution of the relationships to their recruiters reported by participants is similar to the distribution of the relationships to their recruits reported by those in the follow-up survey.

Assumption #2: Seeds independent of final sample

In our assessment of recruitment patterns based on select socio-demographic characteristics, we found that there was a significant tendency for in-group recruitment (homophily) based on age and employment status (Table 1). In the sample, recruitment homophily was 1.11, indicating that participants were 11% more likely to have the same employment status as their recruiter, compared to if there was no differential recruitment based on employment status (Chi-square test for independence, $p < .01$). Participants were 68% more likely to be in the same age category as their recruiter, as compared to if recruitment was random (Chi-square test for independence, $p < .001$).

To assess whether the selection of seeds influenced the final estimate of cumulative lifetime incidence of abortion, we examined bottleneck and convergence plots. Figure 1 shows the convergence plot for the proportion of participants who reported ever having an abortion (cumulative lifetime incidence of abortion). The higher proportion earlier the sample shows that the cumulative lifetime incidence of abortion was higher among initial

participants, and converged on the stationary distribution after around 450 participants. Figure 2 shows the bottleneck plot for cumulative lifetime incidence which shows the cumulative lifetime incidence of abortion over time, by seed. The convergence plot for employment similarly reaches equilibrium around 450 participants (Figure 3). There is no evidence of bottlenecks for abortion, age, or employment status (Figure 4), indicating that there do not appear to be distinct sub-communities with respect to these characteristics within the RDS sample. However, while home language appears to converge in the overall sample (Figure 5), there do appear to be bottlenecks by home language (Figure 6), as the proportion of participants reporting each language do not converge on the same estimate across different seeds. This may suggest that the final sample is not independent of the choice of initial seeds, at least with respect to home language.

Assumption #3: With-replacement sampling

We assessed whether our study approximated sampling with replacement by assessing the following: failure to attain sample size, failed recruitment attempts, proportion of network who have participated. While we did not attain the proposed sample size in the study, it was due to the decision by the study team to end recruitment early given an overwhelming number of coupons in circulation and limited staffing to process interviews before we lost access to the study site due to the end of the year holidays in South Africa. Thus, failure to attain sample size was not due to a global finite population effect; furthermore, our sample size calculation was based on a conservative estimate of the cumulative lifetime incidence of abortion.

Among those who completed a follow-up survey, 93.4% reported that no one refused a coupon. However, those who successfully recruited were disproportionately represented among follow-up survey participants. In the baseline survey, when asked, "Not including the person who recruited you, how many people do you know who have already participated in the study?" nearly half (45.3%) reported not knowing any other study participants; most (89.3%) reported knowing less than 5 study participants. These data are informative insofar as participants disclosed their participation in the study to others in their social network. The number of contacts who have participated did not increase meaningfully over the study period (data not shown).

Additionally, given that the estimated target population size (~300,000) is much larger than participants mean network size (~70), it is unlikely that our study of ~900 faced issues with depletion of eligible participants.

Assumption #4: Respondents accurately report degree

In the baseline survey, the average degree (number of eligible participants in the respondent's social network who they have seen in the past week) was 72.3, with a range of 0 to 2500. The median degree was 20 (interquartile range (IQR) 10 – 50), 9.9% of participants reported a degree greater than 150, and 0.6% reported a degree greater than 1000. A substantial minority of participants (7.7%) reported a higher degree than their total reported network size. Additionally, examination of the distribution of reported

degree (Supplementary Material 2) highlights peaks for degrees that are multiples of 10, suggesting that participants may be rounding their responses. When asked, "How many coupons could you distribute by tomorrow if you were given unlimited coupons?" participants reported a mean of 32 coupons (median 10, IQR: 5-23, range: 0 – 2000). Rounding and extreme outliers that appear to be implausible are suggestive of inaccurate reporting of degree. Imputed visibility ranged from 1 – 13, with a mean of 6.3 (IQR 5 – 7). Self-reported degree was strongly correlated with estimated potential number of recruits (Spearman's R = 0.68) and self-reported degree at recruitment follow-up (R = 0.52). However, self-reported degree was not correlated with actual success in recruiting (R = 0.02).

The median difference between reported degree at baseline and follow-up was 0, suggesting there was no systematic differences between the two visits that may influence reporting of degree; however, most individual responses varied by more than 10% between visits.

Assumption #5: Respondents randomly refer within their network

On average, participants at baseline reported that 50.6% of eligible contacts who they have seen in the past week are currently working (employed). In the overall sample, only 16.3% reported they were currently working. In the recruitment follow-up, participants reported that only 11.2% of those they distributed contacts to were working. These results indicate over-recruitment of unemployed participants within individual's social networks. Furthermore, based on the most recent census figures, the estimated female workforce participation rate was 30%, suggesting that the network of participants in the sample may not represent the total target population as it pertains to employment status.(95)

We did not ask participants about the age composition of their social network or the age composition of their potential recruits. However, the age distribution of Soweto is approximately equal to the age distribution of the sample (Supplementary Material 3).

Effect of assumptions on estimate of cumulative lifetime incidence of abortion

We assessed the potential effect of failure to meet the above assumptions on the estimated cumulative lifetime incidence of abortion (Table 3). The crude estimate of cumulative lifetime incidence of abortion in the overall sample was 12.5% (95% CI: 10.4%, 14.9%). The RDS-II estimate of cumulative lifetime incidence was 12.1% (95% CI: 9.8%, 14.3%) using imputed visibility and 10.5% (95% CI: 5.6%, 15.5%) using self-reported degree at baseline. After excluding participants who reported they would not have recruited their recruiter (no reciprocity of ties), the RDS-II estimate of cumulative lifetime incidence was 11.7% (95% CI: 9.2%, 14.2%).

To assess the effect of self-reported degree (network size) on the estimate, we restricted to those who completed a follow-up and self-reported degree at two time points. Among this sub-sample of the population, the crude cumulative lifetime incidence of abortion was 18.3% (95% CI: 14.3%, 23.3%). The RDS-II point estimate varied based on whether the

weight variable was visibility, self-reported degree at baseline, and self-reported degree at follow-up, though differences were not significant (Table 3).

Finally, we estimated the cumulative lifetime incidence of abortion using post-survey weights to account for different probabilities of inclusion based on employment status and age (Table 3). The employment-adjusted point estimate using weights based on the target population employment proportions (16.9%, 95% CI 12.8%, 22.1%) was higher than all other adjusted point estimates in the full sample, though this difference was not significant. The age-adjusted proportion (13.1%, 95% CI 10.9%, 15.7%) was similar to the crude estimate (12.5%. 95% CI: 10.4%, 14.9%).

2.5 Discussion

This paper presents a rare and important opportunity to assess the theoretical assumptions of RDS in a study of reproductive age women, and finds that many core assumptions were not met. Many RDS studies do not collect follow-up data on participants' recruitment experiences, nor do they have census data to allow for comparison between the study sample and the target population. In this study, we were able to leverage census data and information collected from follow-up to rigorously assess these assumptions. Our findings suggest that RDS may not be an appropriate method for constructing a probability sample, at least among this particular target population (women of reproductive age living in Soweto, South Africa). Re-weighting the RDS sample based on population weights yielded higher estimates of cumulative lifetime incidence of abortion.

In this study, we found that the following RDS assumptions were met with strong certainty: Individuals can identify others as members of the target population, and sampling approximates sampling with replacement. There were minor violations of the assumption that seeds are independent of the final sample composition, and that social ties were reciprocal. The following assumptions were violated in this study: participants can accurately report their degree, and that participants randomly recruit from within their social network. The implications of violating these assumptions are discussed further below.

Assumptions

As peer-to-peer recruitment methods potentially oversample those who are well connected in a target population, accurate assessment of degree is key to adjustment in RDS studies in order to generate population-representative estimates.(96) However, we found that while reporting of degree (the number of people in their social network that the respondent has seen in the past week), was consistent across baseline and follow-up surveys, large and implausible reported network sizes, as well as evidence of rounding to multiples of ten when reporting network size suggest that degree may not be accurately measured. Furthermore, self-reported degree did not appear to correlate with actual success in recruiting; in other words, those with larger reported network sizes did not appear to recruit more potential participants than those with smaller reported network sizes. Visibility, an imputed measure based on reported degree and success in recruiting, and

time spent recruiting, has been suggested as an alternative as a measure of degree.(92) However, this measure still relies on self-reported degree, which we have shown does not correspond with actual recruitment success. Our findings suggests that RDS estimators that rely on self-reported degree (whether via reported social network size, or imputed visibility), may not adequately account for selection bias introduced in the study based on peer-to-peer recruitment due to measurement error of self-reported degree.

We also found that non-random recruitment within an individual social's network may have an important impact on RDS estimates, particularly if participants are more likely to be recruited or participate based on characteristic also associated with the outcome. In this study, we found that participants who were unemployed were over-represented in this sample, as compared recent estimates on female workforce participation in Gauteng province (where Soweto is located). Additionally, participants in the sample reported that their networks had a higher unemployment rate than the overall target population unemployment rate. RDS theory posits that adjustment based on degree should account for the over-representation of participants who have networks with higher unemployment than the target population.(49, 54, 91) However, RDS estimators do not account for differential recruitment within individual's social networks. Prior research on RDS in a general population for which census data was available similarly found that those of lower income and social status were more likely to participate in the study.(90) Unemployed participants in this study may be more motivated to accept a coupon and participate in the study due to the cash incentives, and likely had more time to travel to the study site and participate in a survey during traditional working hours. This may be of greater concern in settings with large wealth disparities.

Non-random recruitment and inaccurate degree have strong implications for the lack of validity of the RDS-II estimator. After adjustment using post-survey weights of estimated probabilities of selection based on employment status, we found that our cumulative lifetime incidence of abortion was higher than both the crude and RDS-II estimate.

Limitations

This study had several limitations. Importantly, given the lack of data on the true incidence of abortion in this sample, we are unable to assess which estimate in our study is closest to the "truth." However, based on global estimates of number of abortions in the Southern Africa region, and the low likelihood of over-reporting of abortion experiences, we believe that the true cumulative lifetime incidence of abortion is higher than any of the estimates from in this study. As such, we assume that the estimate generated by using post-survey weights based on census employment data are the least biased estimates. However, our ability to adjust for bias is limited by having province-level data (rather than city-level data) on employment status. Furthermore, reporting of stigmatized experiences could likely have been improved using data collection methods such as Audio Computer-Assisted Self-Interview (ACASI) that do not require participants to directly disclose sensitive items to study interviewers.

Furthermore, participants who successfully recruited were much more likely to have completed a follow-up survey on their recruitment experiences, because the survey was administered when participants returned to the study site to collect their secondary incentives. As a result, it is likely that our study did not capture the recruitment experiences of all participants, particularly those who did not recruit and thereby likely underrepresenting coupon refusal rates. Future studies should offer additional incentives for recruitment follow-up (separate from secondary incentives), to encourage participation even among those who did not successfully recruit.

Implications

This study has important implications for measurement of abortion incidence. We think that our estimates suffered from under-reporting of abortion due to stigma and social desirability bias, and have concerns about the utility of RDS to construct a population representative sample to measure abortion incidence in a general population of women of reproductive age. However, peer-to-peer recruitment of those who have had an abortion may be a powerful recruitment (but not sampling) strategy to gather important information from those who have had abortion experiences. While not representative, this strategy may reach a broader population of individuals that may be reached via convenience sampling from facilities, or who would be included in probability-based sample. For example, RDS studies among a target population of those who have ever had an abortion could yield information about the proportion of out of clinic abortions that are successful, information about the rate of complications, or care-seeking rates among this population. Future research should explore leveraging such data from RDS studies on abortion along with external data sources to use analytic approaches such as population size estimation(97) or the Abortion Incidence Complications Methodology(39) to generate estimates of abortion incidence.

RDS offers a mechanism to construct a sample of participants from a target population that are often hard to reach, and for whom additional research is critical. However, researchers employing this method should take caution in interpreting results as representative of the target population, as several key assumptions underlying RDS may not be met. Most importantly, the discrepancy between self-reported degree and successful recruiting, as well as non-random recruitment from within social networks, may have outsize impacts that threaten the representativeness of estimates under RDS assumptions. We believe that adjustment via post-survey weights holds greater promise in correcting for selection bias when probabilities of selection can be estimated, as was possible in this study among a general population of reproductive aged women. If census level estimates of the sociodemographic characteristics of the target population are not available, adjustment based on probability of selection within individuals' social network may correct for bias, to some extent. Future studies should explore the extent to which these alternative adjustments may correct for bias in a population where validation against census estimates is possible.

2.6 Conclusion

In conclusion, many of the key theoretical assumptions of RDS were not met in this population of reproductive age women. Post-survey weights based on the population characteristics may improve performance. However, many studies may not have access to census-level data to assess the representativeness of their RDS sample or to construct post-survey weights (given the lack of data on key populations often access by RDS methods that motivate these alternative sampling methods). Despite this challenge, possible remedies include the use of post-survey weights based on reported network sociodemographic characteristics. Studies that employ RDS methods should be preceded by rigorous formative qualitative work to identify key sociodemographic factors that might influence recruitment and the outcome of interest. Researchers can then collect information on important sociodemographic characteristics of participants' social networks, and construct post-survey weights using these estimates. Additionally, studies should plan for additional recruitment follow-up surveys, as recommended by Gile et al(85) to collect information on failed recruitment attempts, decisions on who to recruit, and additional demographics about those recruited. Finally, researchers should not discount the utility of convenience samples, particularly when they allow access to key populations for whom research is lacking. RDS may be able to increase inclusiveness and reach a broader population of individuals that may be reached via convenience sampling or would be included in probability-based sample, though may not be representative of the overall population. While RDS may not yield representative estimates of abortion incidence, peer to peer recruitment may be a possible method for recruiting convenience samples of people who have had abortions who may be missed in facility-based studies.

2.7 Tables and Figures

Table 1. Lifetime abortion incidence and socio demographic characteristics of women participating in a respondent-driven sampling survey, Soweto, South Africa, 2018 (N=849).

		Sample Proportion		RDS-II Estimated Population Proportion		Homophily†	Abortion incidence*
	N	%	95%CI	%	95% CI	Value	%
Total population	849	100%	-	100%	-	-	12.5%
Age (categorical)	**849**					1.68†	
15 - 19	133	15.7%	(13.4%, 18.3%)	16.1%	(14.2%, 18%)	1.52	0.0%
20 - 24	184	21.7%	(19%, 24.6%)	22.2%	(19.5%, 24.8%)	1.20	10.3%
25 - 29	166	19.6%	(17%, 22.4%)	19.2%	(17.5%, 20.8%)	1.00	17.5%
30 - 34	127	15.0%	(12.7%, 17.5%)	14.9%	(11.8%, 17.9%)	1.18	17.3%
35 - 39	109	12.8%	(10.7%, 15.3%)	12.6%	(9.8%, 15.4%)	1.03	11.0%
40 - 44	78	9.2%	(7.4%, 11.3%)	9.0%	(6.8%, 11.2%)	1.13	18.0%
45 - 49	52	6.1%	(4.7%, 8%)	6.1%	(3.6%, 8.7%)	1.27	19.2%
Marital Status	**843**					1.01†	
Living with partner	207	24.6%	(21.8%, 27.6%)	24.3%	(23.7%, 24.9%)	0.97	16.9%
Partner, not living together	332	39.4%	(36.1%, 42.7%)	39.6%	(36.1%, 43.1%)	0.99	10.8%
Separated/Divorced	14	1.7%	(1%, 2.8%)	1.6%	(0%, 4.9%)	0.94	21.4%
Single	290	34.4%	(31.3%, 37.7%)	34.5%	(31%, 37.9%)	1.04	10.7%
Employment Status	**848**					1.11†	
Employed	138	16.3%	(13.9%, 18.9%)	16.9%	(14.4%, 19.4%)	1.15	11.1%
Unemployed/student	710	83.7%	(81.1%, 86.1%)	83.1%	(80.6%, 85.7%)	1.15	19.6%
Home language	**845**					2.03†	
Afrikaans	2	0.2%	(0.1%, 0.9%)	0.2%	(0%, 2.1%)	0.80	0.0%
English	4	0.5%	(0.2%, 1.3%)	0.5%	(0.3%, 0.7%)	1.19	0.0%
IsiXhosa	66	7.8%	(6.2%, 9.8%)	7.7%	(7.7%, 7.8%)	1.21	7.6%
IsiZulu	329	38.9%	(35.7%, 42.3%)	39.9%	(38.9%, 40.8%)	1.48	15.5%

	Sample Proportion			RDS-II Estimated Population Proportion		Homophily†	Abortion incidence*
	N	%	95%CI	%	95% CI	Value	%
Sepedi	20	2.4%	(1.5%, 3.6%)	2.4%	(0.0%, 6.0%)	1.19	0.0%
Sesotho	216	25.6%	(22.7%, 28.6%)	24.5%	(21.3%, 27.8%)	1.41	13.4%
Setswana	59	7.0%	(5.4%, 8.9%)	7.2%	(6.9%, 7.5%)	1.08	11.9%
Tshivenda	35	4.1%	(3%, 5.7%)	4.0%	(2.5%, 5.5%)	1.77	8.6%
Xitsonga	100	11.8%	(9.8%, 14.2%)	12.1%	(7.1%, 17.1%)	2.13	10.0%
Shona	13	1.5%	(0.9%, 2.6%)	1.3%	(0.0%, 3.5%)	3.91	7.7%
Other	1	0.1%	(0%, 0.8%)	0.1%	(0.0%, 1.3%)	1.33	0.0%

† Values indicated with † are recruitment homophily values, indicating the tendency of recruits to have the same characteristic as their recruiter. All other values are population homophily values, which refers to whether there are more homophilius pairs within each level of a characteristic than would be expected due to chance.

*Cumulative lifetime incidence of abortion, row percentage (proportion of respondents in each category level who reported at least one abortion attempt during their lifetime).

Table 2. Recruitment dynamics among women participating a respondent-driven sampling survey, Soweto, South Africa (N = 849).

Recruitment Dynamics in Baseline	N	%
Completed Baseline	**849**	**100.0%**
Relationship to recruiter		
Friend	393	46.9%
Sister	46	5.5%
Cousin	61	7.3%
Daughter	13	1.6%
Mother	21	2.5%
Coworker/Colleague/Customer/Classmate	26	3.1%
Neighbor/Church member/Community member	169	20.2%
Friend of friend	26	3.1%
Other	4	0.5%
Stranger	5	0.6%
Other female relative	64	7.6%
Missing	10	1.2%
Would you have recruited your recruiter?		
No	45	5.4%
Yes	763	91.1%
Not sure	20	2.4%
Missing	10	1.2%

Recruitment Dynamics in Follow-Up	N	%
Total number of relationship reported on	822	100%
Relationship to potential recruit		
Friend	441	53.6%
Sister	54	6.6%
Cousin	54	6.6%
Daughter	13	1.6%
Mother	8	1.0%
Coworker/Colleague/Customer/Classmate	27	3.3%
Neighbor/Church member/Community member	159	19.3%
Friend of friend	5	0.6%
Other	61	7.4%
Would your recruit have recruited you?		
No	45	5.5%
Yes	691	84.1%
Not sure	86	10.5%

Table 3. Cumulative lifetime incidence of abortion among women participating in a respondent-driven sampling survey, Soweto, South Africa, 2018 (N=849).

Estimate of Cumulative Lifetime Incidence of Abortion	N	Weight Variable	Point Estimate	95% CI*
Crude Estimate	849	None	12.5%	(10.4%, 14.9%)
RDS-II	849	Imputed visibility	12.1%	(9.8%, 14.3%)
RDS-II	849	Degree at baseline	10.5%	(5.6%, 15.5%)
RDS-II (Excluding those with no reciprocal ties)	763	Imputed visibility	11.7%	(9.2%, 14.2%)
Employment-adjusted based on network structure	849	Employment	15.3%	(12.1%, 19.1%)
Employment-adjusted based on population structure	849	Employment	16.9%	(12.8%, 22.1%)
Age-adjusted based on population structure	849	Age	13.1%	(10.9%, 15.7%)
Among those who completed follow-up (N = 289)				
RDS-II	289	Imputed visibility	19.1%	(14.2%, 24.1%)
RDS-II	289	Degree at baseline	20.3%	(10.8%, 29.8%)
RDS-II	289	Degree at follow-up	21.2%	(11.2%, 31.1%)

* 95% Confidence Intervals for RDS-II estimates constructed via bootstrapping procedures.

Figure 1. Convergence plot for cumulative lifetime incidence of abortion from a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa (n = 849). The solid line shows the change in the cumulative sample proportion of abortion with increasing number of study participants; the dashed line shows the final overall cumulative lifetime incidence of abortion among all participants in the study. Equilibrium is reached when the cumulative proportion (solid line) converges on a stable estimate and does not change with successive waves of participants.

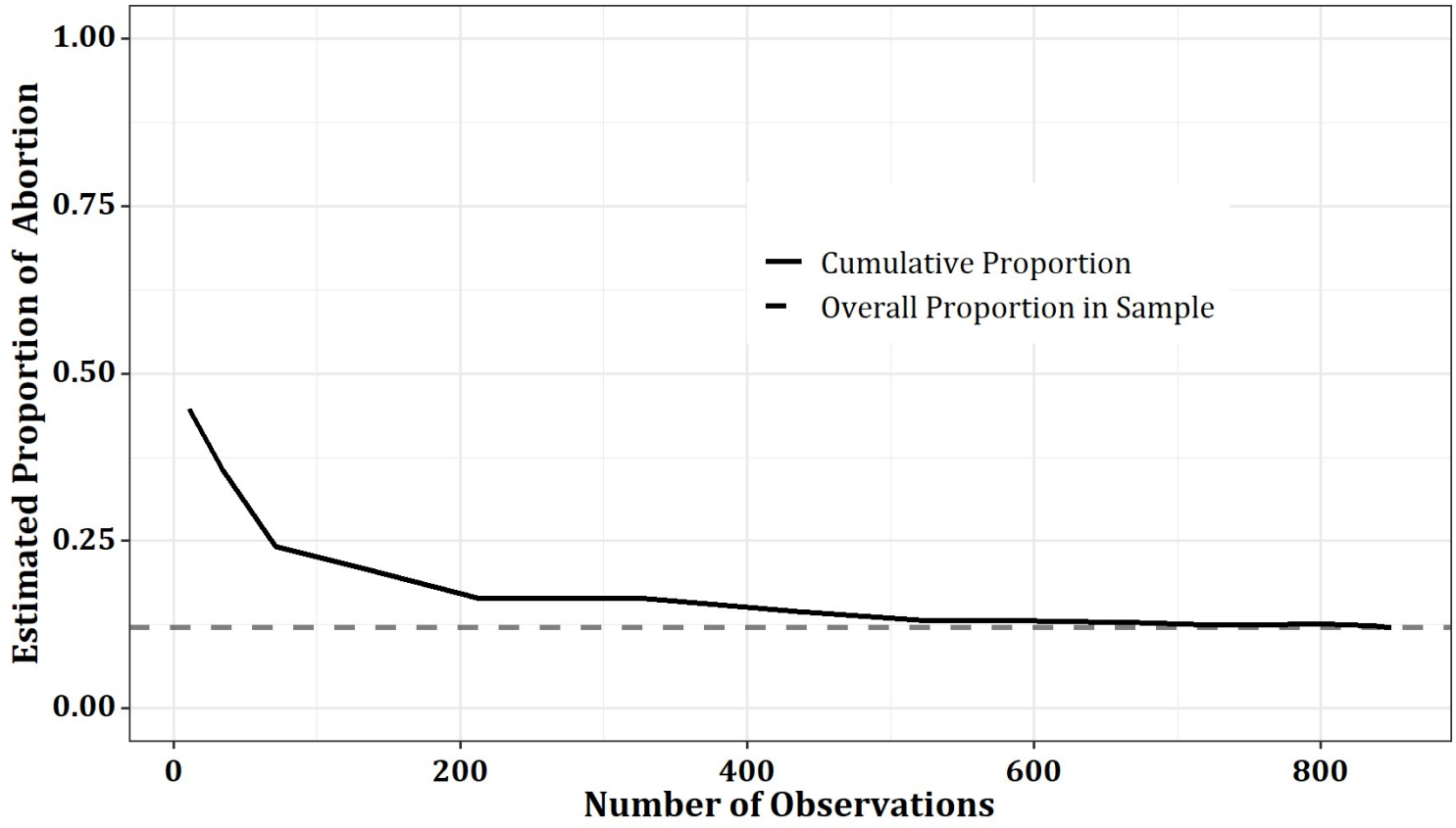


Figure 2. Bottleneck plot for cumulative lifetime incidence of abortion from a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa (n = 849). The solid lines show the cumulative proportion of abortion with increasing number of study participants, by seed (each separate line represents a different recruitment chain). The dashed line indicates the cumulative lifetime incidence of abortion in the overall sample at the end of the study. The solid dot indicates the cumulative lifetime incidence of abortion in each recruitment chain at the end of the study. Bottlenecks are present if the dots do not appear to converge on the dashed line, which would indicate that the overall estimate may be affected by the initial selection of seeds.

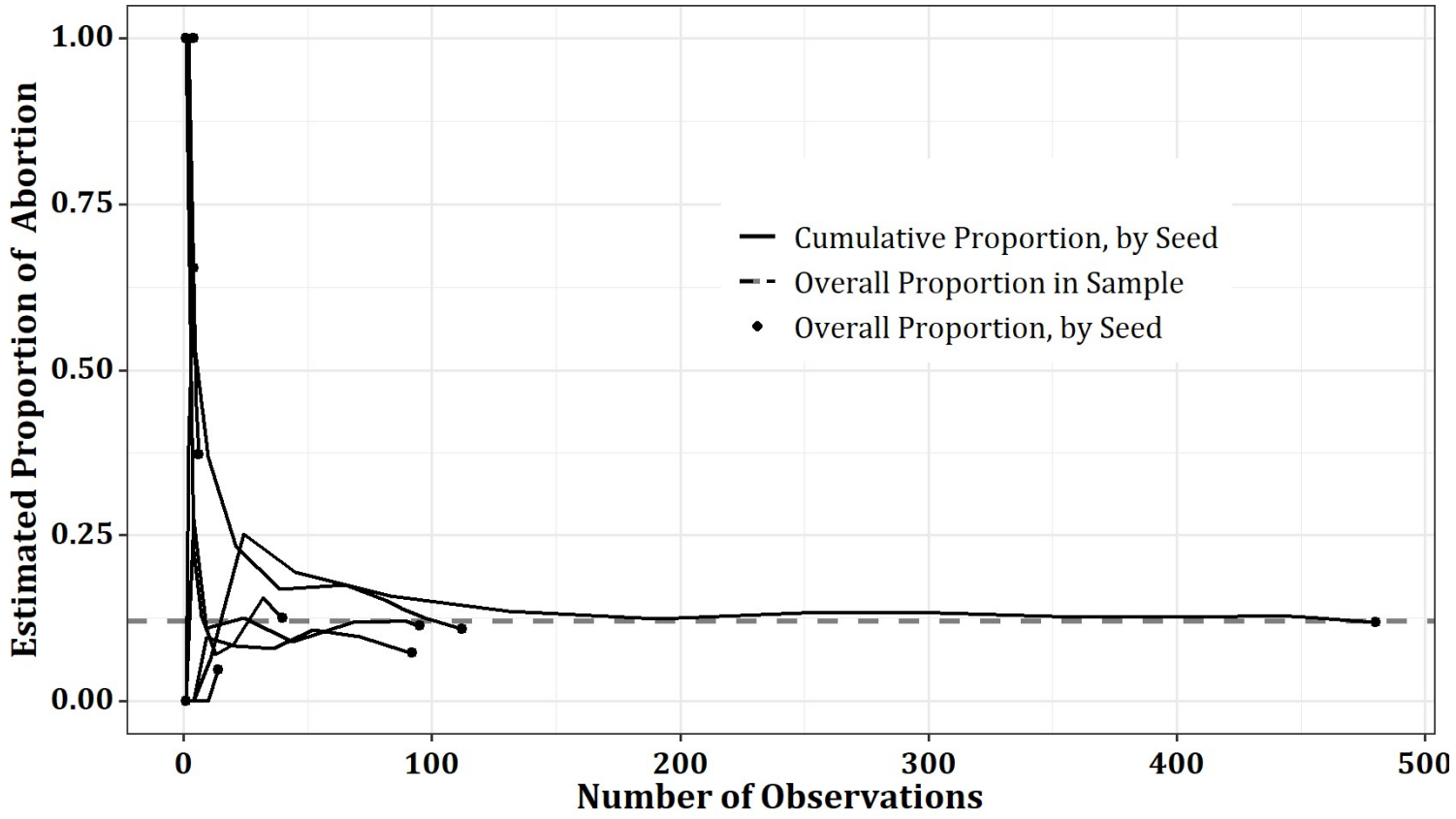


Figure 3. Convergence plot for employment status from a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa (n = 849)

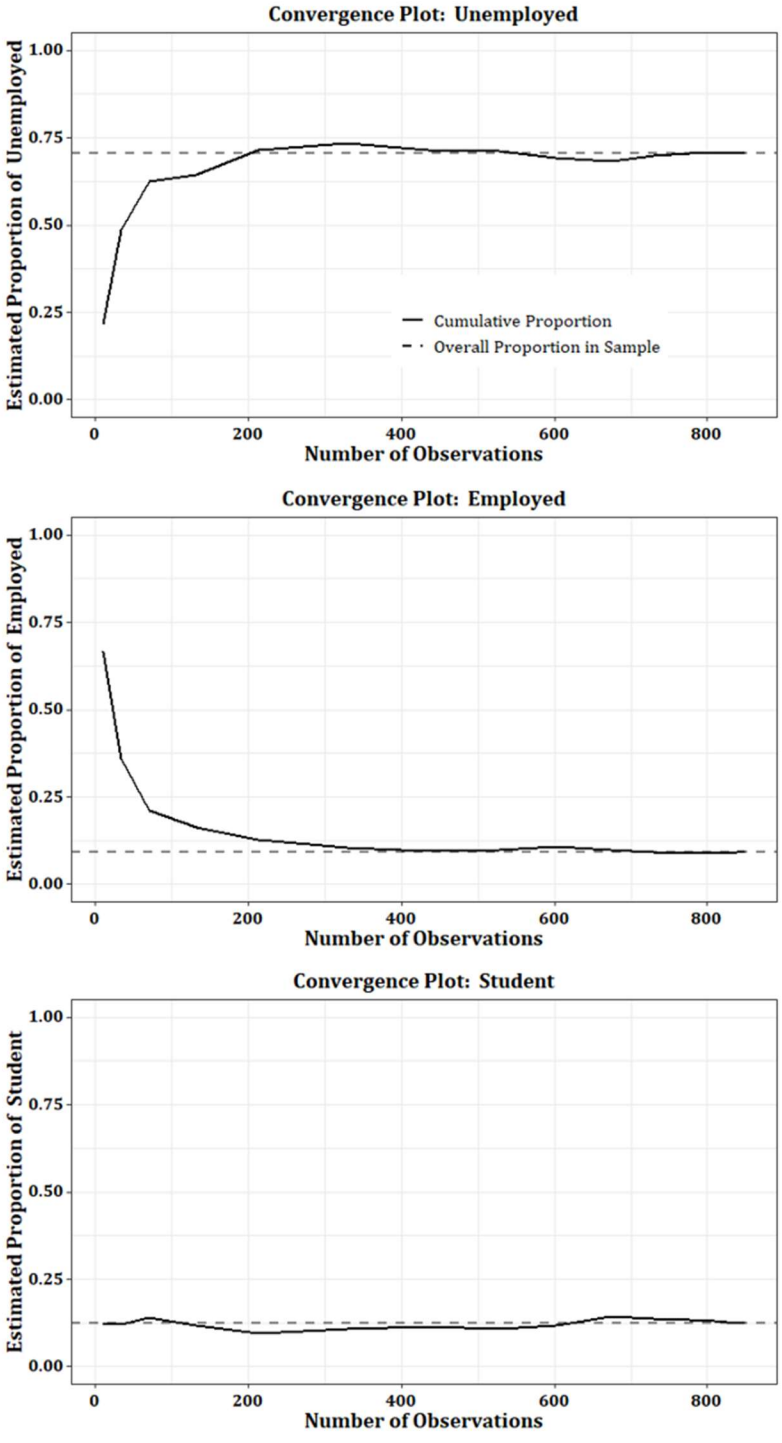


Figure 4. Bottleneck plot for employment status from a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa (n = 849)

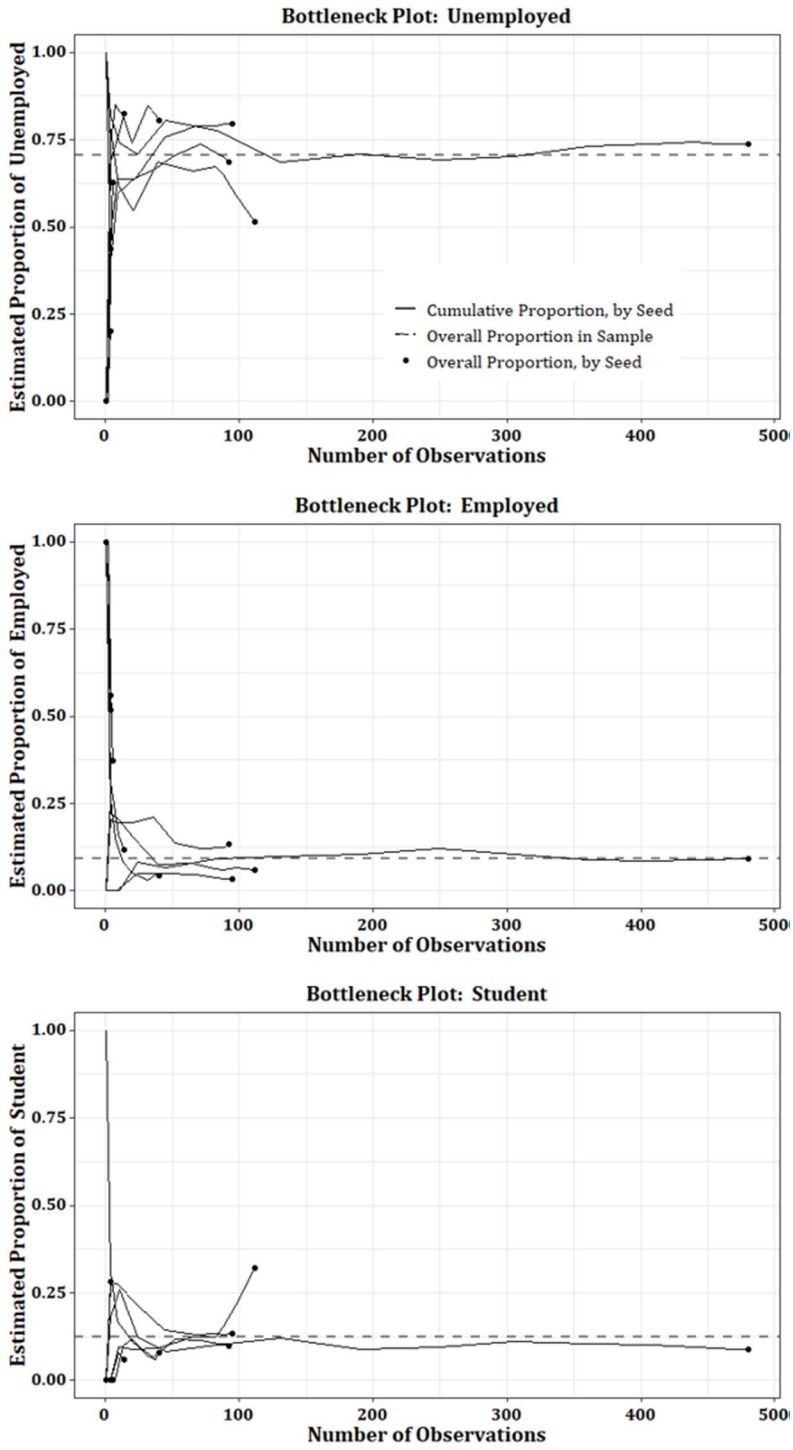


Figure 5. Convergence plot for home language from a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa (n = 849).

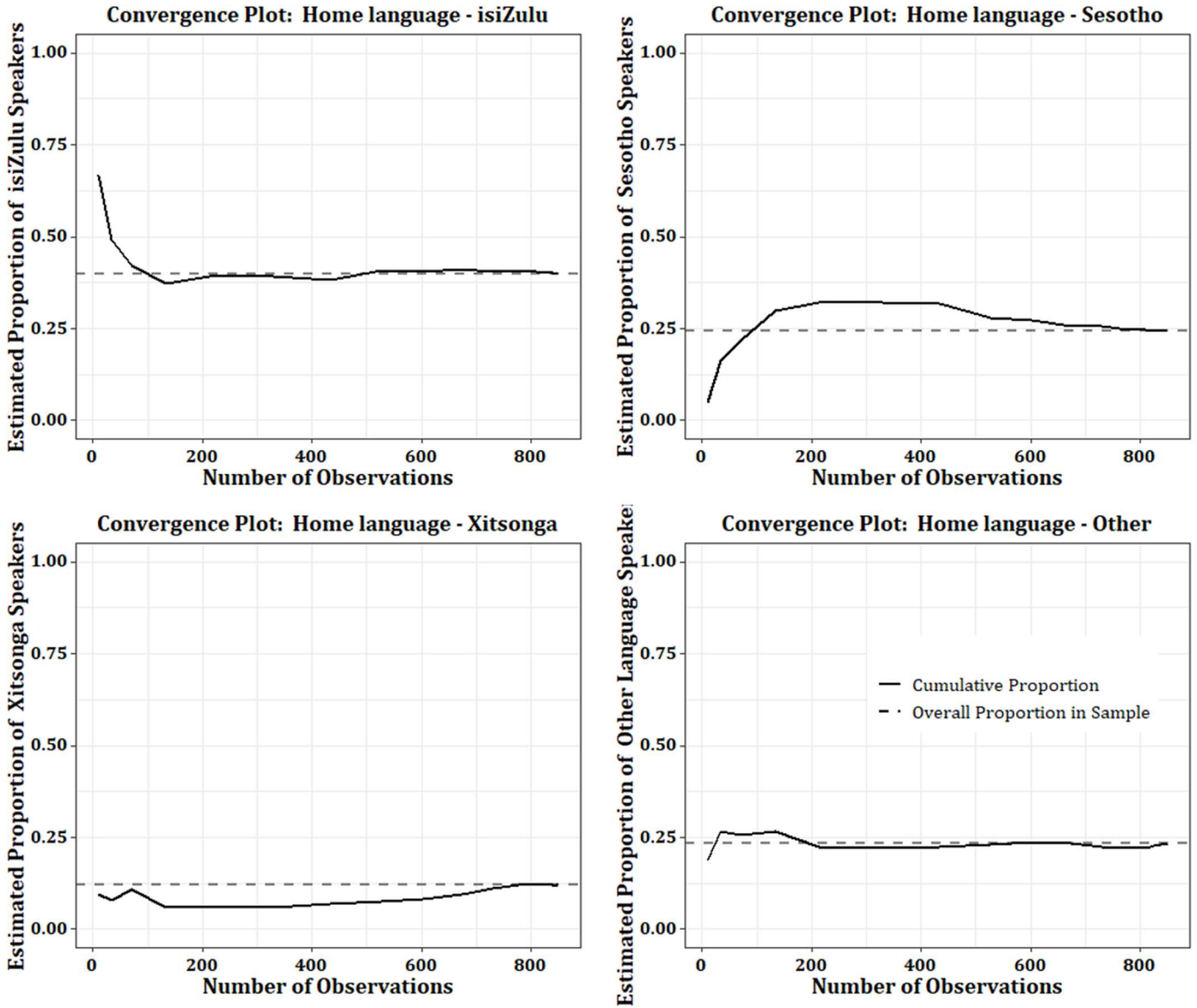
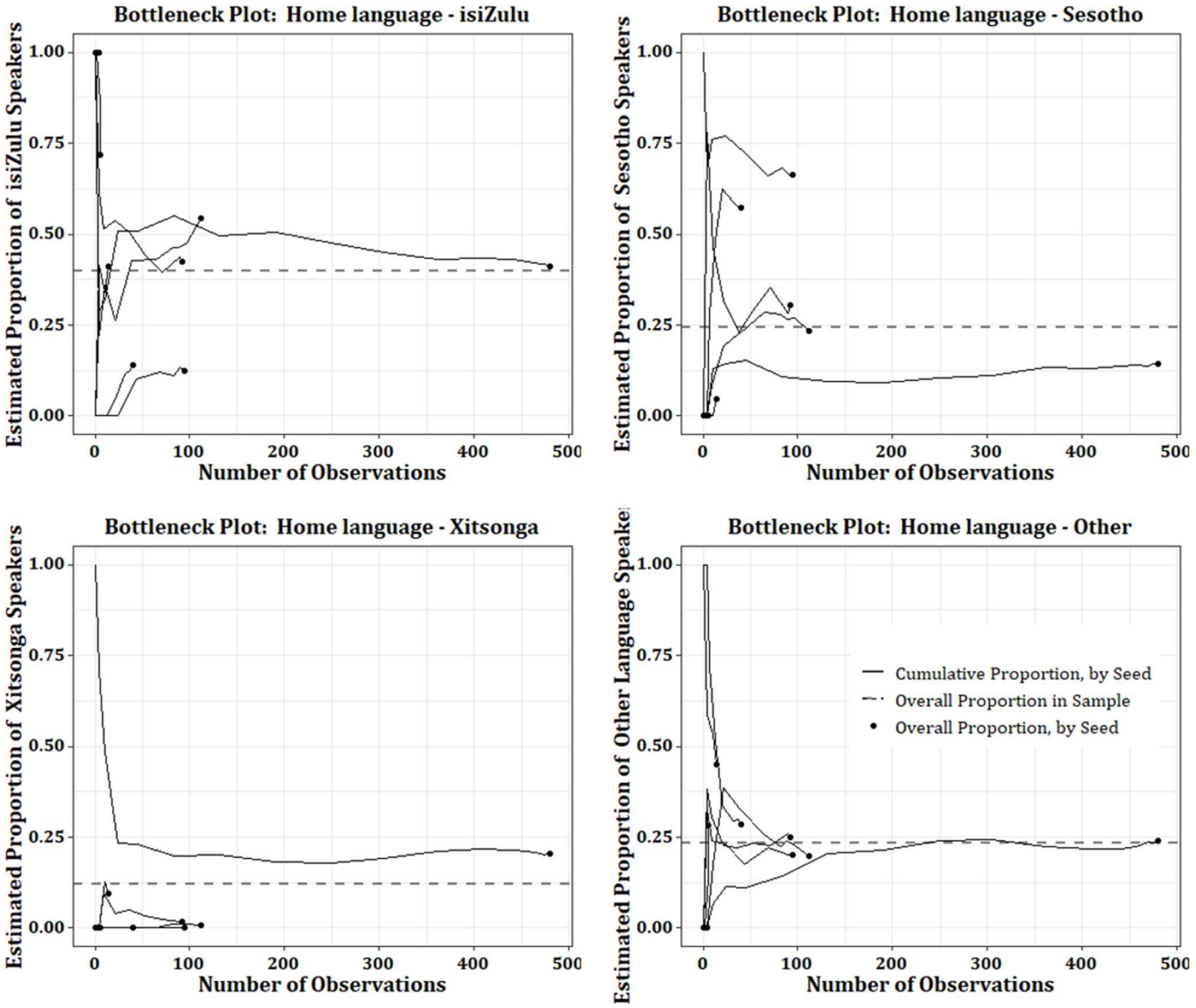


Figure 6. Bottleneck plot for home language from a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa (n = 849).

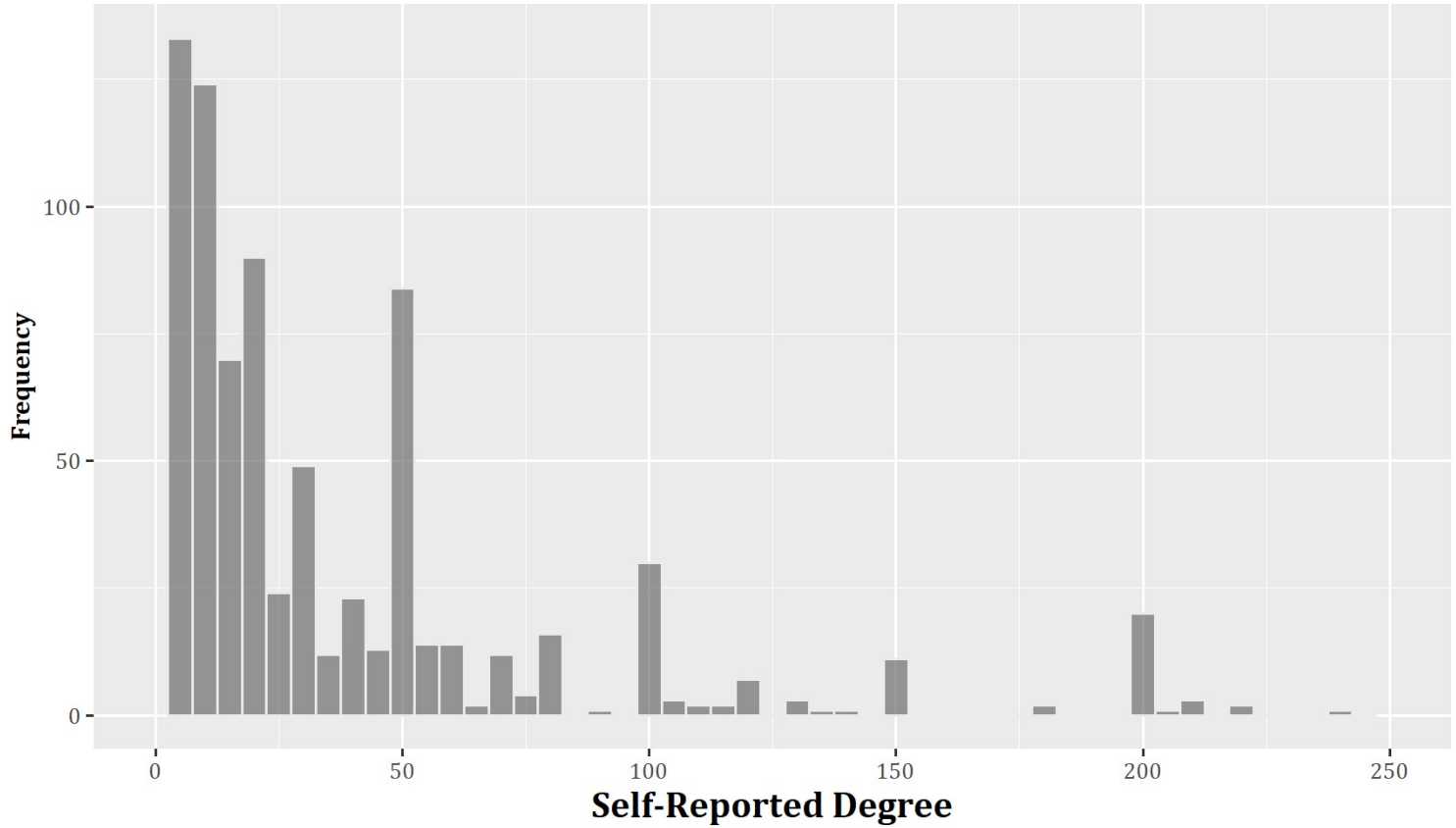


2.8 Supplementary Material

Supplementary Material 1. Recruitment tree from a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa (n = 849).[†]

[†]Each node represents a participant, connected to their recruits and recruiters. Nodes in black indicate a participant who reported any lifetime experience with abortion.

Supplementary Material 2. Histogram of self-reported degree in a respondent-driven sampling study of women aged 15 - 49, Soweto, South Africa, truncated to responses below 250. The histogram below shows the frequency of reported responses to the question, "How many women of reproductive age who live in Soweto who you know and who know you did you see in the past week?" Peaks in the graph around multiples of 10 (i.e. at 30 and 50) suggest that participants may be rounding their responses rather than providing an exact count, which has implications for the accuracy of self-reported degree.



Supplementary Material 3. Sample proportions, RDS-II estimated proportions, and estimated source population proportions of selected demographic characteristics for women of reproductive age in Soweto, South Africa

		Sample Proportion		RDS-II Estimated Population Proportion		Estimated Proportion in Source Population*
	N	%	95%CI	%	95% Bootstrap CI	%
Total population	849	100%	-	100%	-	-
Age (categorical)	**849**					
15 - 19	133	15.7%	(13.4%, 18.3%)	16.1%	(14.2%, 18%)	13.7%
20 - 24	184	21.7%	(19%, 24.6%)	22.2%	(19.5%, 24.8%)	18.8%
25 - 29	166	19.6%	(17%, 22.4%)	19.2%	(17.5%, 20.8%)	19.5%
30 - 34	127	15.0%	(12.7%, 17.5%)	14.9%	(11.8%, 17.9%)	15.4%
35 - 39	109	12.8%	(10.7%, 15.3%)	12.6%	(9.8%, 15.4%)	12.6%
40 - 44	78	9.2%	(7.4%, 11.3%)	9.0%	(6.8%, 11.2%)	10.6%
45 - 49	52	6.1%	(4.7%, 8%)	6.1%	(3.6%, 8.7%)	9.6%
Marital Status	**843**					
Living with partner	207	24.6%	(21.8%, 27.6%)	24.3%	(23.7%, 24.9%)	27.9%
Partner, not living together	332	39.4%	(36.1%, 42.7%)	39.6%	(36.1%, 43.1%)	-
Separated/Divorced	14	1.7%	(1.0%, 2.8%)	1.6%	(0%, 4.9%)	2.0%
Single	290	34.4%	(31.3%, 37.7%)	34.5%	(31%, 37.9%)	-
Employment Status	**848**					
Employed	138	16.3%	(13.9%, 18.9%)	16.9%	(14.4%, 19.4%)	70.6%
Unemployed/student	710	83.7%	(81.1%, 86.1%)	83.1%	(80.6%, 85.7%)	29.4%
Home language	**845**					
Afrikaans	2	0.2%	(0.1%, 0.9%)	0.2%	(0%, 2.1%)	1.3%
English	4	0.5%	(0.2%, 1.3%)	0.5%	(0.3%, 0.7%)	2.3%
IsiXhosa	66	7.8%	(6.2%, 9.8%)	7.7%	(7.7%, 7.8%)	8.7%
IsiZulu	329	38.9%	(35.7%, 42.3%)	39.9%	(38.9%, 40.8%)	37.1%

		Sample Proportion		RDS-II Estimated Population Proportion		Estimated Proportion in Source Population*
	N	%	95%CI	%	95% Bootstrap CI	%
Sepedi	20	2.4%	(1.5%, 3.6%)	2.4%	(0%, 6%)	5.1%
Sesotho	216	25.6%	(22.7%, 28.6%)	24.5%	(21.3%, 27.8%)	15.5%
Setswana	59	7.0%	(5.4%, 8.9%)	7.2%	(6.9%, 7.5%)	12.9%
Tshivenda	35	4.1%	(3.0%, 5.7%)	4.0%	(2.5%, 5.5%)	4.5%
Xitsonga	100	11.8%	(9.8%, 14.2%)	12.1%	(7.1%, 17.1%)	8.9%
Shona	13	1.5%	(0.9%, 2.6%)	1.3%	(0.0%, 3.5%)	0.0%
Other	1	0.1%	(0%, 0.8%)	0.1%	(0.0%, 1.3%)	3.7%

*Population proportions for age, marital status, and home language are from Statistics South Africa; data are from 2011 Census, localized to Soweto. Population proportions for employment status are from the *Quarterly Labour Force Survey* published by Statistics South Africa; data are female unemployment rates in South Africa from July – September 2018.

3 Chapter 3. A Monte Carlo Sensitivity Analysis to account for misclassification and selection bias in the estimation of the effectiveness of self-managed medication abortion

3.1 Abstract

Studies on the effectiveness of self-managed medication abortion may suffer from misclassification and selection bias, due to reliance on self-reported outcomes and loss to follow-up. In this paper, we present a Monte Carlo Sensitivity Analysis technique for estimation of self-managed abortion effectiveness for two different medication abortion regimens that adjusts for misclassification and selection bias and potential inclusion of non-pregnant participants. Data for this study is drawn from the Studying Accompaniment model Feasibility and Effectiveness Study (the SAFE Study). Between July 2019 and April 2020, a total of 1051 participants were enrolled in the study; 961 took abortion medications and completed at least one follow-up. We calculated measures of effectiveness adjusted for bias due to ineligibility, misclassification, and selection bias across 50,000 simulations with bias parameters drawn from the pre-specified beta distributions. The observed effectiveness of self-managed medication abortion was 93.7% for mifepristone in combination with misoprostol, and 99.3% for misoprostol alone. After accounting for the potential influence of various sources of bias, adjusted estimates remained consistent with a high effectiveness of self-managed abortion, conditional on our assumptions around the chosen bias parameters: 92.68% (95% simulation interval: 87.80%, 95.74%) for mifepristone in combination with misoprostol and 98.47% (95% simulation interval: 96.79%, 99.39%) for misoprostol alone.

3.2 Introduction

The World Health Organization (WHO) recommends two regimens for medication abortion: mifepristone in combination with misoprostol, and misoprostol alone.(67) These medications are inexpensive, shelf-stable, and easy to administer. However, legal restrictions, limited providers willing to provide abortion care, and fear of stigma or mistreatment serve as barriers to facility-based abortion care all around the world.(5) The rise of self-managed medication abortion, defined as when a person procures medications for abortion and uses them without clinical supervision, has been credited with declines in maternal morbidity and mortality and expanding access to abortion, particularly in contexts where abortion is legally restricted.(72) The WHO has acknowledged that individuals can provide a central role in their own abortion care, and has highlighted the importance of interventions that support people in safely self-managing.(98) However, WHO task-shifting guidelines do not yet include lay providers or the individual themselves as "recommended" providers of abortion care, citing lack of published evidence.(31)

There are a variety of models that support people in safely self-managing. One model is safe abortion hotlines or "accompaniment models," where counselors, most of whom are not clinically trained and volunteer their time, provide people with evidence-based counseling and support through the medication abortion process. These organizations vary in their modes of operation, but all provide step-by-step protocols for how to use medication to

safely induce abortion based on evidence-based protocols.(72) Globally, there are more than 40 feminist grassroots organizations that operate these services, and provide support to individuals over the phone, via chat, or in person. Accurate estimation of the effectiveness, safety, and other outcomes of those who self-manage their abortion is key to evaluate these models of support for those who self-manage, and to build the evidence base on the effectiveness of self-managed abortion.

Clinical studies have established the effectiveness of misoprostol alone and mifepristone in combination with misoprostol at 75-85%(28-30) and 90-95%(27) respectively, with effectiveness defined as complete abortion without the need for surgical intervention. Additionally, a growing body of evidence on out-of-clinic models of abortion care have found that self-managed abortion with medication is highly effective.(99-101) In a recent scoping review of the literature on self-managed abortion, the authors identified 23 studies that described the experiences of people using mifepristone and misoprostol to self-manage their abortion, and 35 studies that described the experience of people using misoprostol alone; wide heterogeneity in the definition of effectiveness, how the study population was sampled, and the medication protocol used made it impossible to generate an overall meta-analytic estimate of effectiveness.(66)

However, studies that seek to measure the effectiveness of self-managed abortion in legally restricted contexts may be prone to misclassification and selection bias. For example, misclassification bias may arise if self-reported abortion outcomes are incorrectly ascertained. While clinical studies typically ascertain abortion outcomes via ultrasound or a negative pregnancy test confirmed by a provider, this type of ascertainment is impossible in studies where abortion largely occurs outside of the formal healthcare system, and where care-seeking may place people at legal risk. Additionally, selection bias can arise if those who enroll in the study are not representative of the source population, and/or if retention in the study is differential based on the outcome. Counselors at safe abortion hotlines and accompaniment organizations may never meet the people they support face-to-face. Given legal risks and social stigma, those who are seeking abortion support do not often disclose their real names and use temporary phones, making follow-up to ascertain effectiveness and safety outcomes challenging. As a result, high loss to follow-up, selective recruitment of participants from health facilities during post-abortion care seeking, and lack of clinically confirmed completion have been documented as a limitation that precludes the inclusion of existing research on self-managed abortion as “high-quality” evidence in the WHO task-shifting guidelines.

Given these methodological limitations, methods to quantify and adjust for these sources of bias are essential. Proportions (e.g. prevalence) are fundamental measures in epidemiology, and are subject to bias from misclassification and selection factors.(102) A naïve or basic sensitivity analysis could quantify the potential influence of these sources of bias by assuming specific Sensitivity and Specificity values for the diagnostic test assessing the outcome, and assumptions about the proportion of those lost to follow-up or who refused to participate with the outcome. However, probabilistic and semi- and full-Bayesian methods allow us to incorporate uncertainty around these bias parameters.(103-105) While use of probabilistic and Bayesian methods of bias analysis are becoming more

common analytic approaches, the use of these methods in formally assessing the effect of bias on proportions (versus measures of association) has been limited.(106, 107)

In this paper, we present a Monte Carlo Sensitivity Analysis (MCSA) technique for estimation of self-managed abortion effectiveness for two different medication abortion regimens that adjusts for misclassification and selection bias. We then compare these bias adjusted estimates to the observed measure of effectiveness in a study on callers to safe abortion hotlines and accompaniment groups.

3.3 Methods

Data and study subjects

The Studying Accompaniment Feasibility and Effectiveness (SAFE) study is a prospective cohort study to measure the effectiveness of self-managed medication abortion with accompaniment group or safe abortion hotline support. Participants were recruited from two accompaniment groups: one in Argentina, and one in Nigeria. These groups provide information on how to safely procure and take medications for abortion based on WHO-recommended protocol, and provide counseling and support on how to monitor symptoms, assess completion, and how and when to seek additional medical care if wanted or needed. People were eligible to participate in the study if they had contacted the accompaniment group for information on starting a new medication abortion process, had no contraindications to using medication abortion, and were at least 13 years of age. While the accompaniment groups provide counseling to anyone calling for support with self-managed medication abortion, people were ineligible for recruitment into the study if they were unwilling to be followed up after their counseling call, or were experiencing ongoing bleeding and cramping that may indicate an ongoing abortion or miscarriage (though this is not a contraindication to medication abortion, it would make it difficult to ascertain whether their abortion outcome was due to the recorded abortion attempt during study participation or due to an ongoing miscarriage or previous abortion attempt). Participants were recruited after they received counseling from the accompaniment group. Those who provided verbal consent to participate were enrolled in the study and completed an interviewer-administered baseline questionnaire and contacted by phone one week and three weeks after they reported taking medication for abortion. Full details on the study protocol, results from a 60-day pilot study, and study questionnaires have been previously published.(100, 108)

The primary outcome of interest is self-managed medication abortion effectiveness for each medication regimen, defined as the proportion of participants who report complete abortion without surgical intervention. Participants were categorized as “complete abortion without surgical intervention” if they responded “Yes” to the question, “Do you feel that your abortion process is complete?” and reported “No” when they were directly asked if they had received a manual vacuum aspiration (MVA) or dilation and curettage or evacuation procedure (D&C / D&E). Participants who responded “Yes” and reported an MVA or D&C/D&E were categorized as “complete abortion with surgical intervention;”

participants who responded "No" or "Unsure" were categorized as "Not complete/unsure." Data for this analysis are drawn from last recorded follow-up.

Between July 2019 and April 2020, a total of 1051 participants were enrolled in the study; 961 took medications and completed at least one follow-up. While both accompaniment groups provide information on both medication abortion regimens, given local availability of abortion medications in each context during the study period, participants in Argentina largely used mifepristone in combination with misoprostol (mifepristone + misoprostol), while participants in Nigeria largely used misoprostol alone. As a result, we do not compare effectiveness between regimens, as we are unable to distinguish whether these differences are due to the medication or due to differences such as health system structure and accessibility or differences in counseling. For the purposes of this bias analysis, we excluded participants who used misoprostol alone in Argentina (n = 2) and participants who used the mifepristone + misoprostol in Nigeria (n = 4).

Ethical approval for this study was obtained by the Allendale Investigational Review Board (Old Lyme, CT) and the Comité de Bioética de Fundación Huésped (Buenos Aires, Argentina). The study was registered with the ISRCTN registry, number ISRCTN95769543.

Statistical analysis

We present our assumptions of the bias framework for this study and details about adjustment of the effectiveness estimate based on proposed bias parameters below.

Bias framework

We conceptualized three main sources of potential systematic error (bias) in our estimates of self-managed abortion effectiveness (Figure 1).

First, selection bias arises at two time points: study enrollment and follow-up. If eligible participants who did not enroll in the study have different self-managed medication abortion experiences than those who chose to participate in the study, our measure of effectiveness may be biased. Similarly, among study participants, if those lost to follow up after completing baseline were more or less likely to have a complete abortion without surgical intervention that those who remain in the study, our measure of effectiveness may be biased.

Second, misclassification of our primary outcome (a three level variable: complete without surgical intervention, complete with surgical intervention, not complete/unsure) may arise because measurement of this outcome relies on self-report. Participants may unintentionally misclassify themselves as "complete," when their abortion process is not actually complete, or, may misclassify themselves as "not complete" when their abortion process is truly complete. We do not believe that participants intentionally misrepresented their abortion outcome. The extent of this misclassification depends on the Sensitivity (*Se*) and Specificity (*Sp*) of self-report as a method of abortion completion assessment. For participants that reported a complete abortion without surgical intervention, or that their

abortion was not complete, we do not know the *Se* or *Sp* for self-reporting this outcome. However, we assume that any abortion process that ended in surgical intervention was reported with 100% accuracy as all participants were directly asked if they had any surgical procedure, and intentional misclassification is unlikely. Consequently, we assume that the outcome "complete with surgical intervention" is measured with 100% Sensitivity and 100% Specificity.

Finally, there is the possibility that non-pregnant (and therefore, ineligible) participants enrolled in the study. Not all participants had clinical confirmation of pregnancy at enrollment – thus, there may be some proportion of study participants who were not pregnant at the time they took the medications for abortion. We assume that these participants would self-report that their abortion was complete without surgical intervention, and thus should be excluded from the denominator and numerator of any measure of effectiveness.

Overview of methods for bias-adjustment

To calculate bias-adjusted measures of effectives ($\pi_1^*, \pi_1', \pi_1^\dagger$, defined below) from our measure of observed effectiveness ($\tilde{\pi}_1$) based on the above bias framework, we removed ineligible participants, and then adjusted for each source of bias in the reverse of the order that it occurred.

In the below equations, $\mathbf{Y} = (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Y}_3)$ is a multinomially distributed random variable indexed by category i, where $i = 1,2,3$ and $n = \sum_{i=3}^{1} y_i$, where y_i denotes the realized outcome in category i. Category 1 (y_1) represents individuals who had a complete abortion without surgical intervention, y_2 represents those did not have a complete abortion or were unsure, and y_3 represents those who had a complete abortion with surgical intervention. The probability of each outcome is given by π_i, with $\sum_{i=3}^{1} \pi_i = 1$.

The objective was to estimate medication abortion effectiveness, given by the parameter π_1, which represents the probability of a complete abortion without surgical intervention, and is estimated by $\hat{\pi}_1 = y_1/n$. Instead of the true counts $\{y_i\}_{i=1}^3$ we observed the reported counts $\{\tilde{y}_i\}_{i=1}^3$, which we adjusted for ineligibility, misclassification, and selection bias below.

Adjustment for eligibility

To adjust for the potential participation of ineligible (not pregnant) participants in the study, we calculated the expected number of participants with a complete abortion without surgical intervention who were actually pregnant (and thus eligible for the study):

$$y_1^* = \tilde{y}_1(1 - \theta_{N,1}) \qquad \text{(Eq. 1)}$$

where $\theta_{N,1} = \Pr[\text{not pregnant} \mid i = 1]$ is the probability of a negative pregnancy test, which indicates ineligibility. We assumed that all participants who reported their abortion was

not complete ($i = 2$) or complete with surgical intervention ($i = 3$) were eligible ($y_2^* = \tilde{y}_2$ and $y_3^* = \tilde{y}_3$). Thus, the adjusted total number of participants in the study is $n^* = y_1^* + y_2^* + y_3^*$.

Adjustment for misclassification

Next, for those who do not report surgical intervention ($i = 1,2$), we adjusted for misclassification of self-reported abortion completion without surgical intervention. Given assumed Sensitivity (*Se*) and Specificity (*Sp*) of self-reported outcomes among those who did not have a surgical intervention, we calculated outcomes y_i' that are adjusted for eligibility and misclassification:

$$y_1' = \frac{y_1^* - (y_1^* + y_2^*) \times (1 - Sp)}{Se - (1 - Sp)} \quad \text{(Eq. 2)}$$

We made no adjustment to the number who report a complete abortion with surgical intervention ($y_3' = y_3^*$), and adjusted the number who report their abortion is not complete or not sure ($y_2' = n^* - y_1' - y_3'$). This serves to preserve the total number of participants so $n' = n^*$.

Adjustment for selection bias

To adjust for selection bias resulting from differential loss to follow-up and differential enrollment, we calculated probabilities of inclusion in the final analytic sample (denoted *S*) given outcome $i = k$, as the joint probability of completing follow-up and enrolling in the study:

$$\Pr[S = 1|i = k] = \Pr[F = 1|E = 1, i = k] \times \Pr[E = 1|i = k] \quad \text{(Eq. 3)}$$

where *F* and *E* are indicators of completing follow-up and enrollment in the study, respectively.

Using Bayes Rule, and assuming counts $\{y_i'\}_{i=1}^3$ adjusted for eligibility and misclassification, we calculated the probability of follow-up given enrollment and outcome type $i = k$:

$$\Pr[F = 1 \mid E = 1, i = k] = \frac{y_k'}{y_k' + (n - n')\Pr[i = k|F = 0, E = 1]} \quad \text{(Eq. 4)}$$

where *n* is the total number of participants enrolled in the study, and $(n - n')\Pr[y_k'|F = 0, E = 1]$ is the expected number of participants lost to follow-up in outcome group *k*. The relevant bias parameters in this step are:

$$\theta_{k,F,E} = \Pr[i = k|F = 0, E = 1] \quad \text{(Eq. 5)}$$

$$\theta_{E,k} = \Pr[E = 1|i = k] \quad \text{(Eq. 6)}$$

for the probability of outcome k among those lost to follow-up (Eq. 5) and the probability of enrollment (Eq. 6). Substituting these parameters into Equations 3 and 4, the probability of inclusion $\Pr[S = 1|i = k]$, denoted by $\theta_{S,k}$, can be expressed as:

$$\theta_{S,k} = \frac{y'_k}{y'_k + (n - n')\theta_{k,F,E}} \times \theta_{E,k} \quad \text{(Eq. 7)}$$

Estimation of bias-adjusted measures of effectiveness

Based on the above adjustments, we calculated bias-adjusted measures of effectiveness using the equations below.

Observed effectiveness is:

$$\tilde{\pi}_1 = \frac{\tilde{y}_1}{n} \quad \text{(Eq. 8)}$$

The bias-adjusted effectiveness after removal of ineligible participants is given by:

$$\pi_1^* = \frac{y_1^*}{n^*} \quad \text{(Eq. 9)}$$

Effectiveness adjusted for eligibility and misclassification is the proportion:

$$\pi'_1 = \frac{y'_1}{n'} \quad \text{(Eq. 10)}$$

Finally, our estimate of effectiveness, adjusted for eligibility, misclassification, and selection bias by inversely weighting participants based on probability of inclusion ($\omega_k = 1/\theta_{S,k}$), is the proportion:

$$\pi_1^\dagger = \frac{y'_1 \times \omega_1}{(y'_1 \times \omega_1) + (y'_2 \times \omega_2) + (y'_3 \times \omega_3)} \quad \text{(Eq. 11)}$$

Selection of bias parameters

To provide estimates of effectiveness that are adjusted for possible sources of bias, estimates of the following bias parameters are needed: Sensitivity of self-report of abortion completion (*Se*), Specificity of self-report of abortion completion (*Sp*), proportion of lost to follow-up who had a complete abortion without surgical intervention ($\theta_{1,F,E} = \Pr[y'_1|F = 0, E = 1]$, probability of enrolling in the study ($\theta_{E,k} = \Pr[E = 1|i = k]$), and probability of being ineligible for the study among those reporting a complete abortion ($\theta_{N,1} = \Pr[\text{not pregnant} \mid i = 1]$). As the true value of the above parameters are unknown, we assume that the above parameters are drawn from beta distributions, defined by shape

parameters alpha and beta (α, β), and described in-depth below. Shape parameters for each distribution are listed in Table 1.

Se and Sp of self-report: We were not able to identify any external validation data on *Se* or *Sp* for self-report, and given the context in which the study was conducted, internal validation of self-reported outcomes via ultrasound or pregnancy test was not possible. We drew values of *Se*(Beta[9,1]) and *Sp*(Beta[4,2]) from truncated beta distributions using *rtrunc* from the *truncdist* package in R.(109) We truncated the beta distribution to avoid negative cell counts using the observed effectiveness $\tilde{\pi}_1$and 1 – $\tilde{\pi}_1$ as the minimum value for the *Se* and *Sp* distributions, respectively.(110, 111)

Probability of inclusion: Among all eligible callers, 94% were enrolled into the study. Probabilities of enrollment based on eventual abortion outcome were drawn from a beta distribution; we assumed that those who had a complete abortion had a slightly higher probability of enrolling in the study (95% vs. 85% for those not complete or complete with surgical intervention). After the study period, we attempted to contact the 90 individuals who did not complete a follow-up survey; among these, we were able to reach 47. Among these 47, 3 did not proceed with medication abortion, 2 had a complete abortion with surgical intervention, 2 had a complete abortion (surgical intervention unknown), and 39 had a complete abortion without surgical intervention, corresponding to an estimated effectiveness among those lost to follow-up of 88% (39/44). Based on these data, we drew values of the effectiveness among those lost to follow-up from a beta distribution (Beta [39, 5]).

Probability of ineligibility: A random subset of participants (n = 102) were invited to take a pregnancy test at the time of enrollment into the study to confirm study (and abortion eligibility); of these, 2 had a negative pregnancy test (2.0%), and all decided to proceed with their abortion attempt as they had had a prior positive pregnancy test. Based on this, we drew values for probability of ineligibility from a beta distribution with shape parameters (Beta [2,100]) corresponding to a median probability of 2.0%.

Monte Carlo sampling

We calculated adjusted measures of effectiveness (π_1^*, π_1', and $\pi_1^\dagger$) based on equations 9, 10, and 11, described above, across 50,000 simulations with bias parameters drawn from the above specified beta distributions. The distribution of these bias-adjusted point estimates only characterize uncertainty given the distributions of the bias parameters; thus, to additionally account for random error inherent in the original (uncorrected) estimate, we resampled the adjusted measure of effectiveness from the sampling distribution of the estimator. Specifically, as the asymptotic distribution of the maximum likelihood estimator of the log-odds of a proportion is normal (Gaussian), we resampled the log-odds of effectiveness from a normal distribution with mean equal to the point estimate of the bias-adjusted log odds, and standard deviation equal to its standard error.(103, 112) We present 95% simulation-based confidence intervals based on 2.5% and 97.5% quantiles; point estimates represent the median estimate across simulations. All analyses were

completed in R 4.0.2;(113) code is available in the Supplementary Material (Supplementary Material 1).

3.4 Results

A total of 401 participants in Argentina were enrolled in the study; 7 reported not proceeding with medication abortion with mifepristone and misoprostol, resulting in data from 393 eligible and enrolled participants for analysis. A total of 350 (89.1%) completed at least one follow-up survey; among these, 328 (93.7%) reported a complete abortion without surgical intervention, 14 (4.0%) reported a complete abortion with surgical intervention, and 8 (2.3%) reported that their abortion was not complete or they were unsure. A total of 650 participants in Nigeria were enrolled in the study; 12 reported not proceeding with medication abortion with misoprostol alone, resulting in data from 638 eligible and enrolled participants for this analysis. A total of 589 completed at least one follow-up survey (92.3%): 585 (99.3%) reported a complete abortion without surgical intervention, 2 (0.3%) reported complete abortion with surgical intervention, and 2 (0.3%) reported that their abortion was not complete or they were unsure.

Table 1 displays the distribution of the bias parameters across 50,000 simulations. The median *Se* used for adjusting effectiveness in the misoprostol group was higher than the mifepristone and misoprostol group given the higher observed effectiveness, which placed a higher minimum bound on the distribution. Distributions of all bias parameters across the 50,000 simulations reflected the specified distributions. The simulated median overall probability of inclusion among those who do not have a complete abortion or had a complete abortion with surgical intervention was much lower among the misoprostol alone group as compared to the mifepristone + misoprostol group.

Bias-adjusted MCSA estimate of self-managed medication abortion effectiveness

After removing potentially ineligible participants from the mifepristone + misoprostol group, effectiveness dropped from 93.71% to 93.60% (Table 2). After additional adjustment for misclassification, the estimate of effectiveness was 93.98%; final adjustment for selection bias resulted in an adjusted effectiveness estimate of 92.68% (95% simulation interval: 87.80%, 95.74%).

In the misoprostol alone group, removal of ineligible participants and adjustment for misclassification had minimal effect. After removing ineligible participants and accounting for misclassification and selection bias, the bias-adjusted MCSA estimate of self-managed medication abortion effectiveness was 98.47% (95% simulation interval: 96.79%, 99.39%) for misoprostol alone.

3.5 Discussion

Given our assumed distributions on the bias parameters, after adjusting for potential misclassification, selection bias, and enrollment of ineligible participants, estimates of the effectiveness of self-managed medication abortion remained highly effective. Bias-adjusted

estimates of effectiveness for both regimens are similar or higher than effectiveness reported in clinical studies: in the most comparable clinical trial on the effectiveness of misoprostol alone, effectiveness was 84.2%(28). Pooled effectiveness across three comparable clinical trials for mifepristone and misoprostol (114-116) had an effectiveness of 94.0%. However, these are also naïve estimates that did not account for possible sources of bias.

In our simulations, we allowed for low specificity of self-report, reflecting concerns in the literature around self-assessment of abortion outcomes: namely, that individuals are more likely to inaccurately self-assess their abortion outcome as complete when it is incomplete than to inaccurately self-assess their abortion outcome as not complete when it is complete. However, it is possible that the true distribution of specificity is higher, which would influence our bias adjusted estimates towards higher effectiveness values. In order to assuage these concerns, we elected to use a lower distributional assumption. To our knowledge, there have been no formal validation studies on the accuracy of self-report of abortion completion compared to clinician-confirmed completion. However, studies have found that the use of a pregnancy checklist to confirm abortion completion had high Sensitivity and Specificity validated against an at-home pregnancy test.(117) As this was an observational study, participants confirmed their completion based on the information they had from counseling and other sources, knowledge about their own bodies, and additional resources available to them. Though outside the scope of this paper, the range of plausible specificity values would need to be much lower in order to meaningfully shift effectiveness in this study.

Our findings highlight the potential importance of selection bias in studies of effectiveness or prevalence estimation. Among those lost to follow-up who we were able to directly contact or extract outcomes from hotline or accompaniment group records, 89% had a complete abortion. Despite this modest difference in effectiveness among this group as compared to those who remained in the study, there was a relatively large difference in the median simulated probability of inclusion among those who had a complete abortion and those who did not (89% vs 33% for the misoprostol alone regimen). However, these differences did not correspond to having a large impact on estimates, likely because overall loss to follow-up and refusal to participate among eligible participants in this study was low. This highlights the importance of robust follow-up and rigorous data collection from eligible participants, as well as those who are censored, in order to meaningfully assess and adjust for these sources of bias.

Prior studies of self-managed abortion have found high levels of abortion completion, higher than effectiveness reported in clinical studies of medication abortion.(99) This has raised questions as to whether some participants in self-managed abortion research incorrectly identify themselves to be pregnant when in fact they are not,(99) and thus inflate estimates of effectiveness. For this reason, we included an assessment of ineligibility in our bias analysis. However, we have a high degree of confidence in a person's ability to self-assess whether or not they are pregnant; confidence informed by research that has demonstrated the accuracy of self-assessment of pregnancy duration as compared to ultrasound. In our study, almost all participants reported that they confirmed their

pregnancy via an at-home pregnancy test, facility-based urine or blood test, or ultrasound. Additionally, among the two participants who had a negative pregnancy test at enrollment, both had reported a previous positive pregnancy test. Given the known Se and Sp of home pregnancy tests, we believe it is more likely that the participants' second pregnancy tests were false negatives, rather than their initial pregnancy tests being a false positive. Furthermore, WHO technical guidance and revised clinical guidance in the wake of the COVID-19 pandemic do not require an ultrasound to assess eligibility for abortion.(73) We would not recommend including removal of potentially ineligible participants in future bias analyses of self-managed medication abortion effectiveness as it may actually introduce bias by removing participants who are in fact eligible.

While the bias-adjusted estimate of mifepristone in combination with misoprostol is similar to effectiveness documented in clinical studies, the observed and bias-adjusted estimate of the effectiveness of the misoprostol alone regimen is much higher than the clinical literature has previously established. This effectiveness, however, is aligned with what other studies in self-managed contexts have established.(99, 101, 118) Effectiveness may be artificially low in clinical studies given the shorter time frame to assess completion (one to two weeks after taking the medications, versus three to four weeks in self-managed studies), and the lower threshold to intervene surgically in clinical contexts; those in legally restrictive settings and/or those who are choosing to end their pregnancies outside the formal healthcare setting may be more likely to wait longer for the abortion process to complete on its own, or to take additional doses of misoprostol, and have much more restricted access to surgical intervention.(119)

This analysis has several limitations. Our selection of bias parameters was based on assumptions that those with a complete abortion have a lower probability of inclusion in the study and distributional assumptions around the Sensitivity and specificity of self-reported abortion completion. However, it is possible that probability of inclusion is not different across outcomes, or that those without a complete abortion have a higher probability of inclusion as they may be more likely to desire additional support from the hotline. Given critiques of studies on the experiences of those who self-managed their abortions, we chose parameters that reflected the most conservative assumptions about effectiveness. If these assumptions are incorrect, however, then the true self-managed abortion effectiveness would be higher than observed in our study. Furthermore, we were unable to formally compare the effectiveness between the two regimens as study site (Argentina versus Nigeria) almost completely determined the medication regimen available to participants. Due to differences in health system structure and accessibility, as well as differences between accompaniment group counseling regarding health care seeking, we are unable to identify which surgical interventions occurred based on clinical need/the effectiveness of the medication regimen, versus those that occurred solely based on recommendations from the accompaniment group about where to seek care. However, we take care to note that the aim of this study was not to compare effectiveness between the two regimens nor to formally test its difference from clinical studies, but rather to demonstrate the possible influence of various sources of bias on observed effectiveness in the SAFE study.

Formal assessment of the influence of bias on estimates of effectiveness and other prevalence measures can be easily implemented. In our study, we found that after adjustment for multiple potential sources of bias, self-managed medication abortion with mifepristone and misoprostol or misoprostol alone remain highly effective. Due to high levels of stigma and privacy concerns with regard to abortion in most contexts, accuracy of abortion measurement has many challenges. Sexual and reproductive health researchers should incorporate formal bias analyses into reporting on abortion and other sensitive sexual and reproductive health research. We offer this analysis as an example for all researchers reporting on prevalence measures from survey research, regardless of discipline.

3.6 Tables and Figures

Table 1. Distribution of bias parameters in Monte Carlo Sensitivity Analysis of data from the SAFE study on self-managed medication abortion effectiveness, Nigeria and Argentina (N = 1,031).

	Parameter distributions	Mife + Miso	Misoprostol alone
Bias Parameters*	**Beta(α, β) [min, max]**	**Median % (IQR)**	**Median % (IQR)**
Se: Sensitivity	Beta(9,1) [OE, 1]	98.86 (98.26, 99.45)	99.83 (99.75, 99.92)
Sp: Specificity	Beta(4,2)[1-OE,1]	68.32 (54.55, 80.59)	68.55 (54.66, 80.46)
$\theta_{1,F,E}$: Probability of complete abortion (Follow-up = 0)	Beta(39,5)[0,1]	89.23 (85.81, 92.14)	89.23 (85.78, 92.08)
$\theta_{E,1}$: Probability of enrolling (i = 1)	Beta(95,5)[0,1]	95.30 (93.73, 96.59)	95.30 (93.75, 96.58)
$\theta_{E,2}$: Probability of enrolling (i = 2)	Beta(85,15)[0,1]	85.20 (82.69, 87.52)	85.23 (82.76, 87.53)
$\theta_{E,3}$: Probability of enrolling (i = 3)	Beta(85,15)[0,1]	85.27 (82.76, 87.52)	85.23 (82.75, 87.51)
$\theta_{S,1}$: Overall probability of inclusion (i = 1)	See Equation 7	85.18 (83.46, 86.70)	89.33 (87.70, 90.76)
$\theta_{S,2}$: Overall probability of inclusion (i = 2)	See Equation 7	62.13 (48.41, 71.74)	33.44 (20.41, 45.82)
$\theta_{S,3}$: Overall probability of inclusion (i = 3)	See Equation 7	73.17 (67.95, 78.09)	39.49 (32.72, 47.25)
$\theta_{N,1}$: Probability of ineligibility (Negative pregnancy test)	Beta(2,100)[0,1]	1.66 (0.95, 2.65)	1.65 (0.95, 2.66)

*Bias parameters are indexed by $i = 1,2,3$ where $i = 1$ represents individuals who had a complete abortion without surgical intervention, $i = 2$ represents those did not have a complete abortion or were unsure, and $i = 3$ represents those who had a complete abortion with surgical intervention.

.

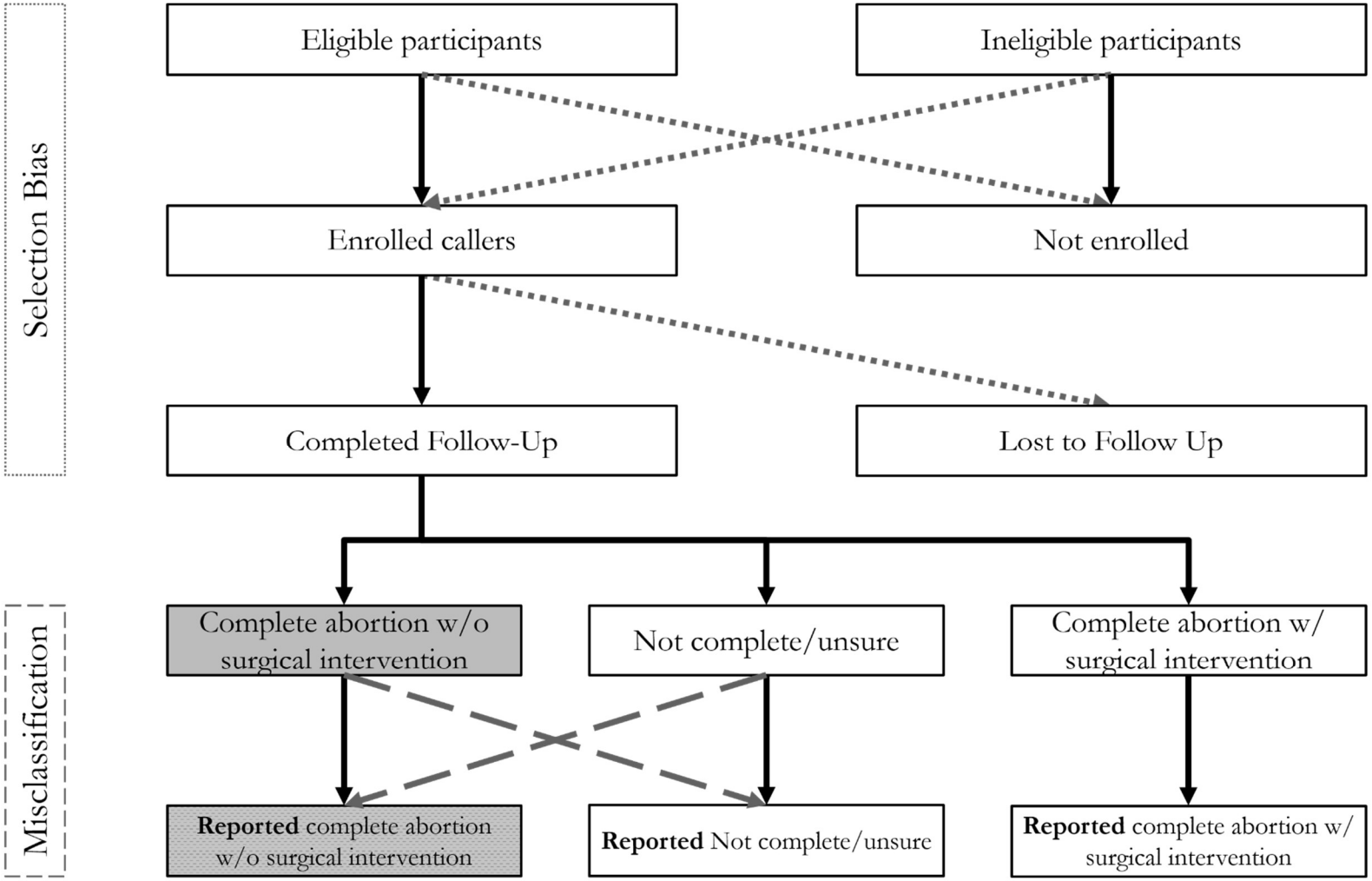
Table 2. Monte Carlo Sensitivity Analysis results of self-managed medication abortion effectiveness from the SAFE study, Nigeria and Argentina (N = 1,031).

	Mife + Miso (N = 393)		Misoprostol Alone (N = 638)	
	%	95% Intervals*	%	95% Intervals*
$\tilde{\pi}_1$: Observed effectiveness†	93.71	90.67, 95.81	99.32	98.27, 99.73
π_1^*: Adjustment for eligibility	93.60	90.51, 95.74	99.31	98.17, 99.74
π_1': Adjustment for eligibility + misclassification	93.98	89.09, 96.76	99.37	98.10, 99.84
$\pi_1^\dagger$: Adjustment for eligibility + misclassification + selection	92.68	87.80, 95.74	98.47	96.79, 99.39

* 95% simulation intervals for adjusted estimates based on 2.5% and 97.5% quantiles across 50,000 simulations

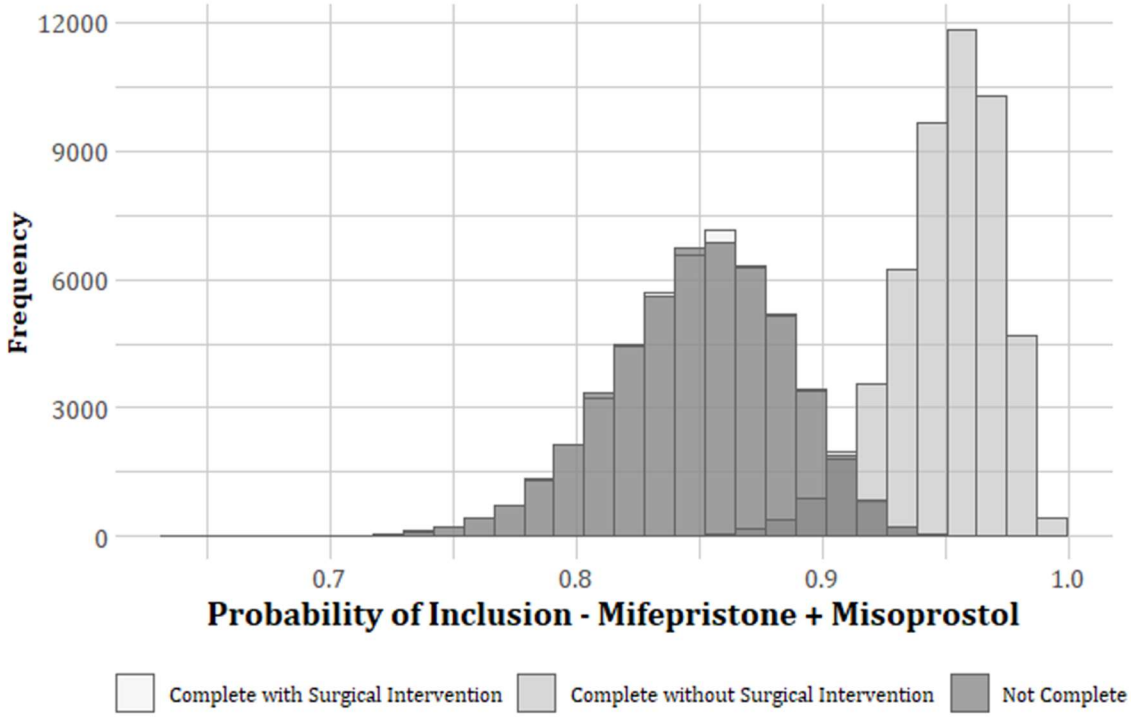
† Missing data on 43 participants in the mife + miso group and 49 participants in the misoprostol alone group are excluded from the denominator of observed effectiveness.

Figure 1. Bias framework for SAFE study on effectiveness of self-managed medication abortion, Nigeria and Argentina (N = 1,031).*



*Dashed lines indicate biasing paths due to selection or misclassification factors.

Figure 2. Simulated inclusion probabilities at enrollment in the SAFE study across 50,000 simulations.



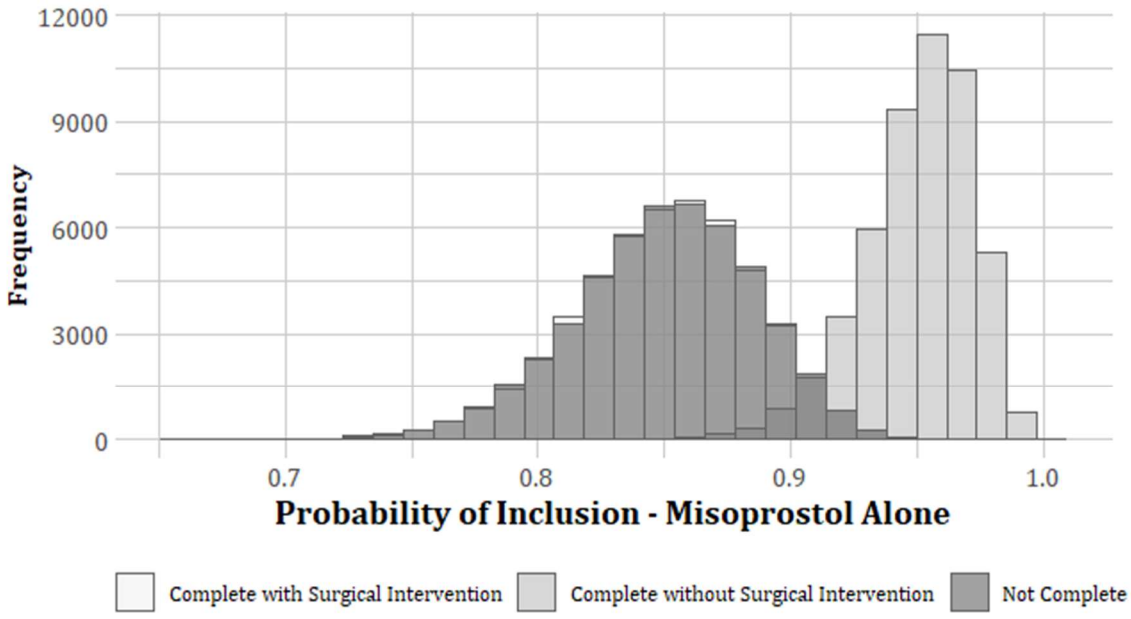
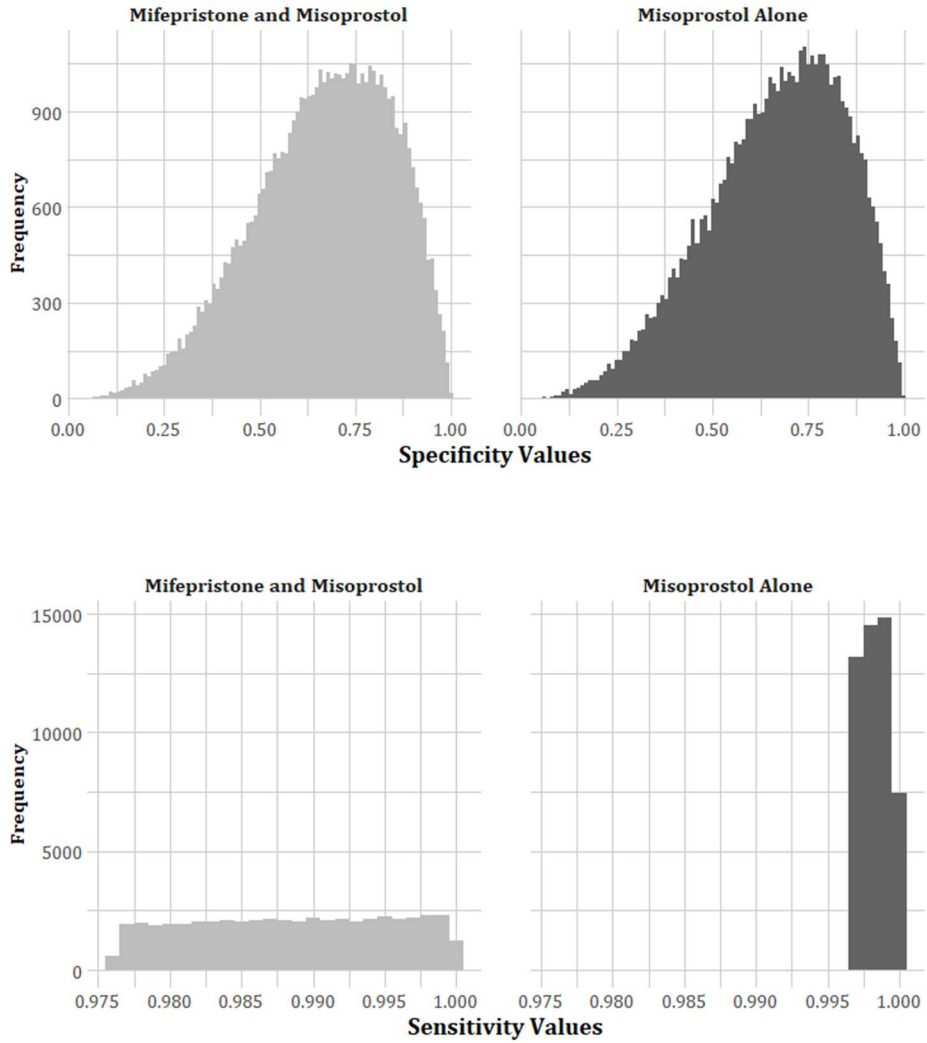


Figure 3. Simulated sensitivity and specificity values of self-reported abortion completion in the SAFE study of self-managed medication abortion effectiveness across 50,000 simulations.



3.7 Supplementary Material

Supplementary Material 1. R code for Monte Carlo Sensitivity Analysis of SAFE study data

```
# Notes:

#Using assumed normal dist of log odds to estimate random error for final MCSA estimate
# Mife and miso separately estimated
# 3 level misclassification model (complete abortion w/o surg, complete abortion w/ surg, not complete/unsure)
# Ineligibles removed prior to correcting for misclassification or selection factors

library(haven)
library(tidyr)
library(truncdist)

#functions
expit <- function(x) exp(x)/(1+exp(x))

#import dataset
# contact corresponding author for access to minimally reproducible dataset

#1. Se, Sp, LTFU, and Inclusion probabilities drawn from distributions & boostrapped CI's
#2. Se, Sp, LTFU, and Inclusion probabilities drawn from distributions, logit(p) drawn from normal dist for CI's
#3. Regression modeling for selection and misclassification & bootstrapped CI's
#4. Regression modeling for selection and misclassification, logit(p) drawn from normal dist for CI's

#####
#Parameter Assumptions
#####
# Sensitivity is at least = to proportion with outcome in the observed
# Specificity is at least 1 - minimum Se, mean 90%
# LTFU with outcome is mean 83%, minimum 43% (based on follow-up data)
# No difference in inclusion probabilities -- though maybe those with advanced GA less likely to participate

######
# 1. Mifepristone + Misoprostol -> 3 level outcome
######
set.seed(126)
N.samp <- 50000 #number of iterations
safe$Y <- safe$complete_nosurg_cat_last
mife <- safe[!is.na(safe$regimen) & safe$regimen==1 & safe$site==1 & !is.na(safe$Y),]
#table(mife$Y, useNA = "always")

p.bc.1 <-p.bc.2 <- p.bc.3 <-p_1 <- p_2 <- p_3 <- p.fu.Y1 <- p.fu.Y2 <- p.fu.Y0 <- rep(NA, N.samp) #store bias-corrected estimates
```

```
#N's for each outcome level
 n_y1 <- table(mife$Y)["1"] #number with outcome in data
 n_y0 <- table(mife$Y)["0"] #number without outcome in data
 n_y2 <- table(mife$Y)["2"] #number complete with surgical intervention (no misclassification)
 n_enrolled <- nrow(mife) + nrow(safe[safe$anyfu==0 & safe$site==1,]) #total number enrolled at study site 1
 n_observed <- (n_y1 + n_y0 + n_y2) #total number with any follow-up data

#observed proportion with complete abortion in data
 p.observed <- n_y1 / (n_y1 + n_y0 + n_y2)


#bias parameters - update these based on validation data/assumptions
 Se = Se.mife = rtrunc(N.samp, spec="beta", shape1 = 9, shape2 = 1, a = n_y1/(n_y1 + n_y0), b = 1)
 Sp = Sp.mife = rtrunc(N.samp, spec="beta", shape1 = 4, shape2 = 2, a = 1-(n_y1/(n_y1 + n_y0)), b = 1)
 p_ltfu = p_ltfu.mife = rtrunc(N.samp, spec="beta", shape1 = 39 , shape2 = 5, a= 0, b=1) #proportion of LTFU with outcome
 p.inclusion.Y1 = p.inclusion.Y1.mife = rtrunc(N.samp, spec="beta", shape1 = 95 , shape2 = 5 , a= 0, b=1) #probability of inclusion among those with outcome
 p.inclusion.Y0 = p.inclusion.Y0.mife = rtrunc(N.samp, spec="beta", shape1 = 85 , shape2 = 15 , a= 0, b=1) #probability of inclusion among those without outcome
 p.inclusion.Y2 = p.inclusion.Y2.mife = rtrunc(N.samp, spec="beta", shape1 = 85 , shape2 = 15 , a= 0, b=1) #probability of inclusion among those with surgical abortion
 p.negtest = p.negtest.mife = rtrunc(N.samp, spec="beta", shape1 = 2, shape2 = 100, a = 0, b = 1) #probability of having a negative pregnancy test prior to enrollment (not eligible)

 for (i in 1:N.samp){

  #0. Remove ineligible participants, assume only among those who ultimately report complete without surgical intervention
  Y1.star <- n_y1*(1-p.negtest[i]) #observed with outcome (complete without surgical intervention)
  Y0.star <- n_y0 #observed not complete/not sure
  Y2.star <- n_y2 #observed complete with surgical
  Y.total <- Y1.star + Y0.star + Y2.star #total after removing ineligible participants

  p_1[i] <- Y1.star / Y.total #observed proportion after removing ineligible participants

  log.odds.p_1 <- log(p_1[i]/(1-p_1[i])) #convert to log odds
  var.log.odds.p_1 <- 1/((Y.total)*p_1[i]*(1-p_1[i])) #calculate variance of log odds [1/np(1-p)]
  se.log.odds.p_1 <- sqrt(var.log.odds.p_1) #calculate standard error

  p.bc.1[i] <- expit((rnorm(1, log.odds.p_1, se.log.odds.p_1))) #resample from normal dist using estimated parameters to add sampling variability


  #1. Correct for misclassification

  Y1.c <- (Y1.star - (Y1.star + Y0.star)*(1-Sp[i]))/(Se[i] - (1-Sp[i])) #from Lash p 95, TRUE # with complete abortion w/o surgical
  Y2.c <- Y2.star #TRUE # with complete abortion with surgical intervention = observed
  Y0.c <- Y.total - Y1.c - Y2.c #TRUE number of not complete/not sure after correction for misclassification
```

```
p_2[i] <- Y1.c/(Y1.c+Y0.c+Y2.c) #proportion with outcome, corrected for misclassification

log.odds.p_2 <- log(p_2[i]/(1-p_2[i])) #convert to log odds
var.log.odds.p_2 <- 1/((Y1.c+Y0.c+Y2.c)*p_2[i]*(1-p_2[i])) #calculate variance of log odds [1/np(1-p)]
se.log.odds.p_2 <- sqrt(var.log.odds.p_2) #calculate standard error of log odds

p.bc.2[i] <- expit((rnorm(1, log.odds.p_2, se.log.odds.p_2))) #resample from normal dist using estimated parameters to add sampling variability


#3. Calculated selection probabilities of follow-up
p.fu.Y1[i] = (Y1.c)/(Y1.c + (n_enrolled - n_observed)*p_ltfu[i]) #probability of follow-up = number followed / (number followed + those LTFU with same outcome)
p.fu.Y0[i] = (Y0.c)/(Y0.c + (n_enrolled - n_observed)*((1-p_ltfu[i])/2))
p.fu.Y2[i] = (Y2.c)/(Y2.c + (n_enrolled - n_observed)*((1-p_ltfu[i])/2))

#4. Correct for selection bias at enrollment and differential LTFU
Y1 <- Y1.c/(p.fu.Y1[i]*p.inclusion.Y1[i])
Y0 <- Y0.c/(p.fu.Y0[i]*p.inclusion.Y0[i])
Y2 <- Y2.c/(p.fu.Y2[i]*p.inclusion.Y2[i])

p_3[i] <- Y1/(Y1+Y0+Y2) #proportion corrected for misclassification AND selection bias

log.odds.p_3 <- log(p_3[i]/(1-p_3[i])) #convert to log odds
var.log.odds.p_3 <- 1/((Y1+Y0+Y2)*p_3[i]*(1-p_3[i])) #calculate variance of log odds [1/np(1-p)]
se.log.odds.p_3 <- sqrt(var.log.odds.p_3) #standard error

p.bc.3[i] <- expit((rnorm(1, log.odds.p_3, se.log.odds.p_3))) #resample from normal dist using estimated parameters to add sampling variability



}

mife.results <- data.frame(p_1, p_2, p_3, p.bc.1, p.bc.2, p.bc.3, p.fu.Y1, p.fu.Y2, p.fu.Y0 )

bias.results.1.mife <- quantile(mife.results$p.bc.1, c(.5, .025, .975)) #MCSA
bias.results.1.mife
bias.results.2.mife <- quantile(mife.results$p.bc.2, c(.5, .025, .975)) #MCSA
bias.results.2.mife
bias.results.3.mife <- quantile(mife.results$p.bc.3, c(.5, .025, .975)) #MCSA
bias.results.3.mife



#####
```

```
# 2. Misoprostol Only
#####
set.seed(276)
N.samp <- 50000
safe$Y <- safe$complete_nosurg_cat_last
miso <- safe[!is.na(safe$regimen) & safe$regimen==2 & !is.na(safe$Y) & safe$site==2,]
#table(miso$Y, useNA = "always")

p.bc.1 <-p.bc.2 <- p.bc.3 <-p_1 <- p_2 <- p_3 <- p.fu.Y1 <- p.fu.Y2 <- p.fu.Y0 <- rep(NA, N.samp) #store bias-corrected estimates

#N's for each outcome level
n_y1 <- table(miso$Y)["1"] #number with outcome in data
n_y0 <- table(miso$Y)["0"] #number without outcome in data
n_y2 <- table(miso$Y)["2"] #number complete with surgical intervention (no misclassification)
n_enrolled <- nrow(miso) + nrow(safe[safe$anyfu==0 & safe$site==1,]) #total number enrolled at study site 1
n_observed <- (n_y1 + n_y0 + n_y2) #total number with any follow-up data

#observed proportion with complete abortion in data
p.observed <- n_y1 / (n_y1 + n_y0 + n_y2)


#bias parameters - update these based on validation data/assumptions
Se = Se.miso = rtrunc(N.samp, spec="beta", shape1 = 9, shape2 = 1, a = n_y1/(n_y1 + n_y0), b = 1)
Sp = Sp.miso = rtrunc(N.samp, spec="beta", shape1 = 4, shape2 = 2, a = 1-(n_y1/(n_y1 + n_y0)), b = 1)
p_ltfu = p_ltfu.miso = rtrunc(N.samp, spec="beta", shape1 = 39 , shape2 = 5, a= 0, b=1) #proportion of LTFU with outcome
p.inclusion.Y1 = p.inclusion.Y1.miso = rtrunc(N.samp, spec="beta", shape1 = 95 , shape2 = 5 , a= 0, b=1) #probability of inclusion among those with outcome
p.inclusion.Y0 = p.inclusion.Y0.miso = rtrunc(N.samp, spec="beta", shape1 = 85 , shape2 = 15 , a= 0, b=1) #probability of inclusion among those without outcome
p.inclusion.Y2 = p.inclusion.Y2.miso = rtrunc(N.samp, spec="beta", shape1 = 85 , shape2 = 15 , a= 0, b=1) #probability of inclusion among those with surgical abortion
p.negtest = p.negtest.miso = rtrunc(N.samp, spec="beta", shape1 = 2, shape2 = 100, a = 0, b = 1) #probability of having a negative pregnancy test prior to enrollment (not eligible)

for (i in 1:N.samp){

 #0. Remove ineligible participants, assume only among those who ultimately report complete without surgical intervention
 Y1.star <- n_y1*(1-p.negtest[i]) #observed with outcome (complete without surgical intervention)
 Y0.star <- n_y0 #observed not complete/not sure
 Y2.star <- n_y2 #observed complete with surgical
 Y.total <- Y1.star + Y0.star + Y2.star #total after removing ineligible participants

 p_1[i] <- Y1.star / Y.total #observed proportion after removing ineligible participants

 log.odds.p_1 <- log(p_1[i]/(1-p_1[i])) #convert to log odds
 var.log.odds.p_1 <- 1/((Y.total)*p_1[i]*(1-p_1[i])) #calculate variance of log odds [1/np(1-p)]
 se.log.odds.p_1 <- sqrt(var.log.odds.p_1) #calculate standard error
```

```
p.bc.1[i] <- expit((rnorm(1, log.odds.p_1, se.log.odds.p_1))) #resample from normal dist using estimated parameters to add sampling variability


#1. Correct for misclassification

Y1.c <- (Y1.star - (Y1.star + Y0.star)*(1-Sp[i]))/(Se[i] - (1-Sp[i])) #from Lash p 95, TRUE # with complete abortion w/o surgical
Y2.c <- Y2.star #TRUE # with complete abortion with surgical intervention = observed
Y0.c <- Y.total - Y1.c - Y2.c #TRUE number of not complete/not sure after correction for misclassification

p_2[i] <- Y1.c/(Y1.c+Y0.c+Y2.c) #proportion with outcome, corrected for misclassification

log.odds.p_2 <- log(p_2[i]/(1-p_2[i])) #convert to log odds
var.log.odds.p_2 <- 1/((Y1.c+Y0.c+Y2.c)*p_2[i]*(1-p_2[i])) #calculate variance of log odds [1/np(1-p)]
se.log.odds.p_2 <- sqrt(var.log.odds.p_2) #calculate standard error of log odds

p.bc.2[i] <- expit((rnorm(1, log.odds.p_2, se.log.odds.p_2))) #resample from normal dist using estimated parameters to add sampling variability


#3. Calculated selection probabilities of follow-up
p.fu.Y1[i] = (Y1.c)/(Y1.c + (n_enrolled - n_observed)*p_ltfu[i]) #probability of follow-up = number followed / (number followed + those LTFU with same outcome)
p.fu.Y0[i] = (Y0.c)/(Y0.c + (n_enrolled - n_observed)*((1-p_ltfu[i])/2))
p.fu.Y2[i] = (Y2.c)/(Y2.c + (n_enrolled - n_observed)*((1-p_ltfu[i])/2))

#4. Correct for selection bias at enrollment and differential LTFU
Y1 <- Y1.c/(p.fu.Y1[i]*p.inclusion.Y1[i])
Y0 <- Y0.c/(p.fu.Y0[i]*p.inclusion.Y0[i])
Y2 <- Y2.c/(p.fu.Y2[i]*p.inclusion.Y2[i])

p_3[i] <- Y1/(Y1+Y0+Y2) #proportion corrected for misclassification AND selection bias

log.odds.p_3 <- log(p_3[i]/(1-p_3[i])) #convert to log odds
var.log.odds.p_3 <- 1/((Y1+Y0+Y2)*p_3[i]*(1-p_3[i])) #calculate variance of log odds [1/np(1-p)]
se.log.odds.p_3 <- sqrt(var.log.odds.p_3) #standard error

p.bc.3[i] <- expit((rnorm(1, log.odds.p_3, se.log.odds.p_3))) #resample from normal dist using estimated parameters to add sampling variability



}

miso.results <- data.frame(p_1, p_2, p_3, p.bc.1, p.bc.2, p.bc.3, p.fu.Y1, p.fu.Y2, p.fu.Y0 )
```

```
bias.results.1.miso <- quantile(miso.results$p.bc.1, c(.5, .025, .975)) #MCSA Scenario 1
bias.results.1.miso
bias.results.2.miso <- quantile(miso.results$p.bc.2, c(.5, .025, .975)) #MCSA Scenario 2
bias.results.2.miso
bias.results.3.miso <- quantile(miso.results$p.bc.3, c(.5, .025, .975)) #MCSA Scenario 3
bias.results.3.miso



####
# SAVE RESULTS
####

bias.mife.results <- apply(mife.results, 2, quantile , probs = c(.5, .025, .975) , na.rm = TRUE)
write.csv(bias.mife.results, "Bias Results - Mife 3 level.csv")

bias.miso.results <- apply(miso.results, 2, quantile , probs = c(.5, .025, .975) , na.rm = TRUE)
write.csv(bias.miso.results, "Bias Results - Miso 3 level.csv")


#### NAIVE ESTIMATES
prop.test(328,350,correct=FALSE) #mife = 0.9371429 ( 0.9066670 0.9581271)
prop.test(585, 589, correct=FALSE) #miso = 0.9932088 (0.982670 0.997356)
```

IV. CONCLUSION

This study responds to an urgent and critically neglected public health problem. The underlying causes of unsafe abortion have been described as "apathy and disdain for women—they suffer and die because they are not valued."(5) The Sustainable Development Goals(120), a set of 17 social and economic development goals set by the United Nations General Assembly that determines global health agendas, do not include any specific indicators for abortion access, safety, or quality of care.(121) As a result, abortion access is frequently excluded from global health funding priorities focused on improving maternal health, despite its substantial contribution to maternal morbidity and mortality(122) and its role as a determining factor of achieving reproductive autonomy.

Morbidity and mortality from unsafe abortion and forced child-bearing persist globally, despite the existence of safe and effective methods of abortion. Medication abortion, in particular, is highly effective in clinical settings, relatively easily stored and transported, and has straightforward dosing regimens. Additional research that demonstrates the safety and effectiveness of medication abortion outside of clinical settings is sorely needed in order to expand WHO recommendations for abortion task-shifting and provide evidence for the scale-up of novel interventions such as telemedicine models, pharmacy access, accompaniment models, and community distribution models. However, legal restrictions on abortion, widespread abortion stigma, lack of research funding, and limited or no infrastructure or incentive for basic surveillance make research on these models particularly challenging.

The first chapter presented formative qualitative research to develop a framework for measuring abortion complications from medication abortion based on self-report. This chapter highlighted the range of experiences in bleeding, cramping, pain, and reasons for seeking medical care reported by those who have had an abortion. Complications from medication abortion (hemorrhage, infection) have clearly defined clinical signs as demonstrated by the abortion complications literature. However, we found that classification schema often include other undesired outcomes—such as ongoing pregnancy or incomplete abortion—as complications. Existing frameworks use receipt of medical treatment, such as surgical intervention for any reason, as an indication of a complication, which conflates the experience of care-seeking with true medical complications. Reframing medical treatment as a positive outcome if the person receives the treatment they wanted in a timely manner, rather than considering it as an indicator of a "complication," may help us shift to a more patient-centered consideration of abortion outcomes and care. This chapter proposed a new framework that reserves the use of the term "complication" for moderate or severe unanticipated problems that arise following a treatment, procedure, or condition, and shifts our focus on measuring outcomes from medication abortion to measuring other dimensions of quality that center the needs and desired experiences of individuals.

The second chapter analyzed whether several core assumptions of respondent-driven sampling were met in a study estimating the cumulative lifetime incidence of abortion

among women of reproductive age in Soweto, South Africa. In this study, we found that several core assumptions were met: respondents are being able to identify others as members of the target population, seeds independent of final composition, and sampling approximated sampling with replacement. However, the following assumptions were not met: participants are able to accurately report their degree and participants randomly recruit within their social network. The latter two assumptions are core to the validity of RDS weights in generating a probability-based sample. The implications of not meeting these assumptions on the estimate are difficult to verify directly, given the lack of population data on abortion incidence (hence the need for using this tool). One possible suggestion is to utilize peer-to-peer methods as a recruitment strategy, rather than a sampling method, and potentially use other characteristics about the population to weight the population to account for differences in socio-demographics.

The third chapter examined the potential impact of selection and misclassification bias on the measure of self-managed abortion effectiveness in a prospective observational study among callers to safe abortion hotlines and accompaniment groups in Nigeria and Argentina where effectiveness was determined via self-report of the outcome. In this study, the observed effectiveness of self-managed abortion was high. Under our specification for the distribution of the bias parameters, after adjusting for selection bias, misclassification bias, and removal of potentially non-pregnant (ineligible) participants, we found that the bias-adjusted estimate of effectiveness of self-managed abortion was lower than observed effectiveness, though still highly effective. The observed and bias-adjusted effectiveness of misoprostol alone is higher than what has been documented in clinical trials. However, effectiveness in this chapter for misoprostol alone was similar to what other studies on self-managed abortion have found – some possibilities for the higher effectiveness in studies on self-managed abortion are longer time period to allow for or assess completion before surgical intervention (as compared to clinical trials), protocols that allow for additional doses of misoprostol if needed, and contextual differences in the availability of clinic-based care.

Proposed future work

Overall, this dissertation seeks to address critical gaps in our ability to measure the safety, incidence, and effectiveness or self-managed abortion. Findings from this dissertation can be used to design research studies that more accurately capture outcomes from self-managed abortion. Building on this body of work, I, together with colleagues, developed a draft self-report questionnaire based on the proposed framework in chapter one that centers the individual's perspective and needs, and shifts away from a strict biomedical frame focused on complication "severity" to one that centers abortion "quality" and outcomes. Together with colleagues, I will conduct a pilot validation study of this self-report questionnaire in a health facility that serves a people seeking post-abortion care and medication abortion to test how the questionnaire distinguishes between abortion complications and abortion symptoms, as well as captures reasons for care seeking among clients. This pilot validation study will also measure the symptoms, severity, and treatment of complications from a provider perspective (using a chart abstraction questionnaire), and assess for alignment between patient preferences, patient needs, and the care patients

ultimately receive. This pilot study will contribute to the development of a full-scale validation study for the self-report tool, which we hope will eventually lead to changes in the perceptions of care seeking and provide data that more accurately reflects and documents the safety, symptoms, and complications associated with medication abortion self-use.

Given the limitations of RDS as a sampling strategy for estimating abortion incidence, we are exploring the use of RDS as a recruitment tool to construct a sample of people who have ever had an abortion experience in refugee settlements in Uganda and Kenya. There is almost no research on the abortion experiences of people living in humanitarian contexts. While recruiting a target population of individuals who have ever had an abortion experience will not allow us to estimate abortion incidence using RDS weighting methods, it will allow us to generate a sample of individuals with any abortion experiences. Though findings from the second chapter suggest that this sample, even after adjusting for degree, may not approximate a probability sample, by collecting more in-depth data on recruitment experiences, conducting follow-up with all participants (regardless of success in recruiting), and collecting more in-depth data about the composition of participant's social networks will allow us to rigorously test RDS assumptions. Furthermore, we will be piloting two different approaches to estimate abortion incidence in these contexts: population-size estimation using data on social network structures from the RDS sample, and a modified Abortion Incidence Complications Methodology (AICM) using data on care-seeking behaviors from the RDS sample.

Finally, we will apply the bias-adjustment approach detailed in chapter three to recently collected data on the observed effectiveness of self-managed abortion collected from safe abortion hotline callers in Indonesia. Additionally, given the need for additional research on the effectiveness of self-managed abortion using misoprostol alone, we are exploring the possibility of pooling data from various observational studies on self-managed abortion using misoprostol alone and calculating bias-adjusted estimates. We hope that these findings can be used to update guidance around who is qualified to be an abortion provider, as well as provide additional evidence on the effectiveness of misoprostol alone regimens. Furthermore, we hope that our proposed analytic approach can be used by researchers conducting other observational studies that rely on self-report.

Conclusion

Our capacity to conduct valid analytic epidemiological studies, such as studies that examine factors associated with abortion complications or assess the safety of alternative innovative models of abortion service delivery, are hindered by our current inability to accurately measure the outcome of interest.(3, 64, 65) Research is needed to document gaps in access to safe abortion and the development and evaluation of effective programs, policies, and interventions to increase access to safe abortion. In settings where abortion is legal, accurately characterizing the burden of unsafe abortion can inform local advocacy efforts to hold governments accountable for providing equitable access to legal and safe services.(123) In legally restricted settings, a lack of trained providers is one of the most

critical barriers that limit access to safe abortion;(31) evidence is sorely needed in order to expand the cadre of recommended providers for medication abortion in WHO task-shifting guidelines.(31, 64, 65, 124)

Results from this dissertation can contribute to the wider field's understanding of self-managed abortion, and ideally be used to inform the development of interventions and policies, and, ultimately, improve access to safe, effective, and supported self-managed abortion.

V. REFERENCES

1. Ross L, Solinger R. Reproductive justice: An introduction: Univ of California Press; 2017.
2. World Health Organization. Unsafe abortion: global and regional estimates of incidence of unsafe abortion and associated mortality in 2008. Geneva: WHO; 2011.
3. Ganatra B, Gerdts C, Rossier C, Johnson BR, Jr., Tuncalp O, Assifi A, et al. Global, regional, and subregional classification of abortions by safety, 2010-14: estimates from a Bayesian hierarchical model. Lancet. 2017 Nov 25;390(10110):2372-81.
4. Ganatra B, Tuncalp O, Johnston HB, Johnson BR, Jr., Gulmezoglu AM, Temmerman M. From concept to measurement: operationalizing WHO's definition of unsafe abortion. Bull World Health Organ. [Editorial]. 2014 Mar 1;92(3):155.
5. Grimes DA, Benson J, Singh S, Romero M, Ganatra B, Okonofua FE, et al. Unsafe abortion: the preventable pandemic. The Lancet. 2006;368(9550):1908-19.
6. Center for Reproductive Rights. The world's abortion laws. Center for Reproductive Rights New York; 2018 [10/1/2018]; Available from: http://worldabortionlaws.com/questions.html.
7. Heino A, Gissler M, Apter D, Fiala C. Conscientious objection and induced abortion in Europe. Eur J Contracept Reprod Health Care. 2013 Aug;18(4):231-3.
8. Jones RK, Jerman J. Abortion Incidence and Service Availability In the United States, 2014. Perspect Sex Reprod Health. 2017 Mar;49(1):17-27.
9. Fuentes L, Lebenkoff S, White K, Gerdts C, Hopkins K, Potter JE, et al. Women's experiences seeking abortion care shortly after the closure of clinics due to a restrictive law in Texas. Contraception. 2016 Apr;93(4):292-7.
10. Baum SE, White K, Hopkins K, Potter JE, Grossman D. Women's Experience Obtaining Abortion Care in Texas after Implementation of Restrictive Abortion Laws: A Qualitative Study. PLoS One. 2016;11(10):e0165048.
11. Jewkes RK, Gumede T, Westaway MS, Dickson K, Brown H, Rees H. Why are women still aborting outside designated facilities in metropolitan South Africa? BJOG. 2005 Sep;112(9):1236-42.
12. Gerdts C, Fuentes L, Grossman D, White K, Keefe-Oates B, Baum SE, et al. Impact of clinic closures on women obtaining abortion services after implementation of a restrictive law in Texas. Journal Information. 2016;106(5).
13. Culwell KR, Hurwitz M. Addressing barriers to safe abortion. Int JGynaecol Obstet. 2013 May;121 Suppl 1:S16-9.
14. Kumar A, Hessini L, Mitchell EM. Conceptualising abortion stigma. Culture, health & sexuality. 2009 Aug;11(6):625-39.
15. Jayaweera RT, Ngui FM, Hall KS, Gerdts C. Women's experiences with unplanned pregnancy and abortion in Kenya: A qualitative study. PLoS One. 2018;13(1):e0191412.
16. Coelho HL, Teixeira AC, Santos AP, Forte EB, Morais SM, La Vecchia C, et al. Misoprostol and illegal abortion in Fortaleza, Brazil. Lancet. 1993 May 15;341(8855):1261-3.
17. Costa SH, Vessey MP. Misoprostol and illegal abortion in Rio de Janeiro, Brazil. Lancet. 1993 May 15;341(8855):1258-61.
18. Barbosa RM, Arilha M. The Brazilian experience with Cytotec. Stud Fam Plann. 1993 Jul-Aug;24(4):236-40.

19. Hyman A, Blanchard K, Coeytaux F, Grossman D, Teixeira A. Misoprostol in women's hands: a harm reduction strategy for unsafe abortion. Contraception. 2013;87(2):128-30.
20. World Health Organization. Safe abortion: technical and policy guidance for health systems. Geneva, Switzerland: WHO; 2012.
21. Gynuity Health Projects. Instructions for use: Mifepristone plus misoprostol or misoprostol-alone for abortion induction in pregnancies 12 - 24 weeks' LMP. 2014.
22. Briozzo L, Vidiella G, Rodríguez F, Gorgoroso M, Faúndes A, Pons J. A risk reduction strategy to prevent maternal deaths associated with unsafe abortion. Int J Gynecol Obstet. 2006;95(2):221-6.
23. Fiol V, Briozzo L, Labandera A, Recchi V, Piñeyro M. Improving care of women at risk of unsafe abortion: implementing a risk-reduction model at the Uruguayan-Brazilian border. International Journal of Gynecology & Obstetrics. 2012;118:S21-S7.
24. Stifani BM, Couto M, Lopez Gomez A. From harm reduction to legalization: The Uruguayan model for safe abortion. Int J Gynecol Obstet. 2018;143:45-51.
25. Erdman JN. Harm reduction, human rights, and access to information on safer abortion. Int J Gynecol Obstet. 2012;118(1):83-6.
26. Erdman JN. Access to information on safe abortion: a harm reduction and human rights approach. Harvard Journal of Law & Gender. 2011;34:413-62.
27. Raymond EG, Shannon C, Weaver MA, Winikoff B. First-trimester medical abortion with mifepristone 200 mg and misoprostol: a systematic review. Contraception. 2013;87(1):26-37.
28. Von Hertzen H, Piaggio G, Huong NTM, Arustamyan K, Cabezas E, Gomez M, et al. Efficacy of two intervals and two routes of administration of misoprostol for termination of early pregnancy: a randomised controlled equivalence trial. The Lancet. 2007;369(9577):1938-46.
29. Ngoc NTN, Shochet T, Raghavan S, Blum J, Nga NTB, Minh NTH, et al. Mifepristone and misoprostol compared with misoprostol alone for second-trimester abortion: a randomized controlled trial. Obstet Gynecol. 2011;118(3):601-8.
30. Blum J, Raghavan S, Dabash R, Ngoc NtN, Chelli H, Hajri S, et al. Comparison of misoprostol-only and combined mifepristone–misoprostol regimens for home-based early medical abortion in Tunisia and Vietnam. Int J Gynecol Obstet. 2012;118(2):166-71.
31. World Health Organization. Health worker roles in providing safe abortion care and post-abortion contraception. 2015.
32. Winikoff B, Sheldon W. Use of medicines changing the face of abortion. Int Perspect Sex Reprod Health. 2012:164-6.
33. Miller S, Lehman T, Campbell M, Hemmerling A, Brito Anderson S, Rodriguez H, et al. Misoprostol and declining abortion-related morbidity in Santo Domingo, Dominican Republic: a temporal association. BJOG. 2005;112(9):1291-6.
34. Anderson BA, Katus K, Puur A, Silver BD. The validity of survey responses on abortion: evidence from Estonia. Demography. 1994 Feb;31(1):115-32.
35. Jones EF, Forrest JD. Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988. Demography. 1992;29(1):113-26.
36. Jagannathan R. Relying on surveys to understand abortion behavior: some cautionary evidence. Am J Public Health. 2001;91(11):1825-31.

37. Jones EF, Forrest JD. Use of a supplementary survey of abortion patients to correct contraceptive failure rates for underreporting of abortion. New York, NY: Population Division, Department of Internaional Economic and Social Affairs, United Nations1991.
38. London K, Williams L. A comparison of abortion underreporting in an in-person interview and self-administered questionnaire. 1990.
39. Singh S, Prada E, Juarez F. The Abortion Incidence Complications Method: A Quantitative Technique. In: Singh S, Remez L, Tartaglione A, editors. Methodologies for Estimating Abortion Incidence and Abortion-Related Morbidity: A Review. New York, NY: Guttmacher Institute; International Union for the Scientific Study of Population; 2010. p. 71-85.
40. Rossier C. The anonymous third party reporting method. In: Singh S, Remez L, Tartaglione A, editors. Methodologies for estimating abortion incidence and abortion-related morbidity: a review. New York and Paris: Guttmacher Institute and IUSSP; 2010. p. 99-106.
41. Singh S, Maddow-Zimet I. Facility-based treatment for medical complications resulting from unsafe pregnancy termination in the developing world, 2012: a review of evidence from 26 countries. BJOG : an international journal of obstetrics and gynaecology. [Review]. 2016 Aug;123(9):1489-98.
42. Gerdts C, Vohra D, Ahern J. Measuring Unsafe Abortion-Related Mortality: A Systematic Review of the Existing Methods. PLOS ONE. 2013;8(1):e53346.
43. Say L. Importance of accurate information on causes of maternal deaths for informing health care programmes. Indian J Med Res. 2010 Aug;132:123-4.
44. Hill K, Stanton C. Measuring maternal mortality through the census: rapier or bludgeon? Journal of Population Research. 2011 March;28(1):31-47.
45. Abouzahr C. New estimates of maternal mortality and how to interpret them: choice or confusion? Reprod Health Matters. 2011 May;19(37):117-28.
46. Lara D, Strickler J, Olvarrieta CD, Ellertson C. Measuring induced abortion in Mexico: A comparison of four methodologies. Sociological Methods & Research. 2004;32(4):529-58.
47. Wang J, Carlson RG, Falck RS, Siegal HA, Rahman A, Li L. Respondent-driven sampling to recruit MDMA users: a methodological assessment. Drug Alcohol Depend. 2005;78(2):147-57.
48. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Social problems. 1997:174-99.
49. Heckathorn DD. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. Social problems. 2002;49(1):11-34.
50. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. Sociological methodology. 2004;34(1):193-240.
51. Decker MR, Marshall BD, Emerson M, Kalamar A, Covarrubias L, Astone N, et al. Respondent-Driven Sampling for an Adolescent Health Study in Vulnerable Urban Settings: A Multi-Country Study. J Adolesc Health. 2014;55(6):S6-S12.
52. Reisner SL, Mimiaga MJ, Johnson CV, Bland S, Case P, Safren SA, et al. What makes a respondent-driven sampling “seed” productive? Example of finding at-risk Massachusetts men who have sex with men. J Urban Health. 2010;87(3):467-79.
53. Ludlam AH, Saxton PJ, Dickson NP, Adams J. Respondent-driven sampling among gay and bisexual men: experiences from a New Zealand pilot study. BMC research notes. 2015;8(1):549.

54. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Soc Probl. 1997;44(2):174-99.
55. Goel S, Salganik MJ. Assessing respondent-driven sampling. Proceedings of the National Academy of Sciences. 2010;107(15):6743-7.
56. Gile KJ, Johnston LG, Salganik MJ. Diagnostics for Respondent-driven Sampling. J R Stat Soc Ser A Stat Soc. 2015 Jan;178(1):241-69.
57. Gerdts C, Tuncalp O, Johnston H, Ganatra B. Measuring abortion-related mortality: challenges and opportunities. Reprod Health. 2015 Sep 16;12:87.
58. Sedgh G, Bearak J, Singh S, Bankole A, Popinchalk A, Ganatra B, et al. Abortion incidence between 1990 and 2014: global, regional, and subregional levels and trends. Lancet. [Research Support, Non-U.S. Gov't]. 2016 Jul 16;388(10041):258-67.
59. Singh S, Maddow-Zimet I. Facility-based treatment for medical complications resulting from unsafe pregnancy termination in the developing world, 2012: a review of evidence from 26 countries. BJOG : an international journal of obstetrics and gynaecology. 2016 Aug;123(9):1489-98.
60. Calvert C, Owolabi OO, Yeung F, Pittrof R, Ganatra B, Tuncalp O, et al. The magnitude and severity of abortion-related morbidity in settings with limited access to abortion services: a systematic review and meta-regression. BMJ global health. 2018;3(3):e000692.
61. Singh S, Prada E, Juarez F. The abortion incidence complications method: a quantitative technique. Methodologies for estimating abortion incidence and abortion-related morbidity: a review, New York: Guttmacher Institute. 2010:71-98.
62. Harries J, Gerdts C, Momberg M, Greene Foster D. An exploratory study of what happens to women who are denied abortions in Cape Town, South Africa. Reprod Health. 2015 Mar 21;12:21.
63. Gerdts C, Tuncalp O, Johnston H, Ganatra B. Measuring abortion-related mortality: challenges and opportunities. Reproductive health. 2015 Sep 16;12:87.
64. Gerdts C, Jayaweera RT, Baum SE, Hudaya I. Second-trimester medication abortion outside the clinic setting: an analysis of electronic client records from a safe abortion hotline in Indonesia. BMJ sexual & reproductive health. 2018 Jul 18.
65. Kapp N, Blanchard K, Coast E, Ganatra B, Harries J, Footman K, et al. Developing a forward-looking agenda and methodologies for research of self-use of medical abortion. Contraception. 2018 Feb;97(2):184-8.
66. Moseson H, Herold S, Filippa S, Barr-Walker J, Baum SE, Gerdts C. Self-managed abortion: a systematic scoping review. Best Practice & Research Clinical Obstetrics & Gynaecology. 2019:In Press.
67. Medical Management of Abortion. Geneva: World Health Organization; 2018.
68. Ngo TD, Park MH, Shakur H, Free C. Comparative effectiveness, safety and acceptability of medical abortion at home and in a clinic: a systematic review. Bull World Health Organ. 2011 May 1;89(5):360-70.
69. Raymond EG, Harrison MS, Weaver MA. Efficacy of Misoprostol Alone for First-Trimester Medical Abortion: A Systematic Review. Obstet Gynecol. 2019 Jan;133(1):137-47.
70. Dzuba IG, Winikoff B, Peña M. Medical abortion: a path to safe, high-quality abortion care in Latin America and the Caribbean. The European Journal of Contraception & Reproductive Health Care. 2013;18(6):441-50.

71. Briozzo L, Gómez Ponce de León R, Tomasso G, Faúndes A. Overall and abortion-related maternal mortality rates in Uruguay over the past 25years and their association with policies and actions aimed at protecting women's rights. Int JGynaecol Obstet. 2016;134(1):004.
72. Erdman JN, Jelinska K, Yanow S. Understandings of self-managed abortion as health inequity, harm reduction and social change. Reprod Health Matters. 2018 Nov;26(54):13-9.
73. Raymond EG, Grossman D, Mark A, Upadhyay UD, Dean G, Creinin MD, et al. Commentary: No-test medication abortion: A sample protocol for increasing access during a pandemic and beyond. Contraception. 2020;101(6):361-6.
74. Grossman D, Grindlay K, Buchacker T, Lane K, Blanchard K. Effectiveness and acceptability of medical abortion provided through telemedicine. Obstet Gynecol. 2011;118(2 Pt 1):296-303.
75. World Health Organization. Health worker roles in providing safe abortion care and post-abortion contraception. Geneva, Switzerland: WHO; 2015.
76. Upadhyay UD, Desai S, Zlidar V, Weitz TA, Grossman D, Anderson P, et al. Incidence of emergency department visits and complications after abortion. Obstet Gynecol. 2015 Jan;125(1):175-83.
77. Bankole A, Kayembe P, Chae S, Owolabi O, Philbin J, Mabika C. The severity and management of complications among postabortion patients treated in Kinshasa health facilities. International perspectives on sexual and reproductive health. 2018;44(1):1.
78. Taylor D, Upadhyay UD, Fjerstad M, Battistelli MF, Weitz TA, Paul ME. Standardizing the classification of abortion incidents: the Procedural Abortion Incident Reporting and Surveillance (PAIRS) Framework. Contraception. 2017;96(1):1-13.
79. Owolabi OO, Cresswell JA, Vwalika B, Osrin D, Filippi V. Incidence of abortion-related near-miss complications in Zambia: cross-sectional study in Central, Copperbelt and Lusaka Provinces. Contraception. 2017 Feb;95(2):167-74.
80. Moustakas C. Phenomenological research methods: Sage; 1994.
81. Sanders P. Phenomenology: A new way of viewing organizational research. Acad Manage Rev. 1982;7(3):353-60.
82. Software V. MAX QDA. 2020 ed. Berlin: VERBI Software; 2019.
83. Creinin MD, Chen MJ. Medical abortion reporting of efficacy: the MARE guidelines. Contraception. 2016 Aug;94(2):97-103.
84. Gerdts C, Raifman S, Daskilewicz K, Momberg M, Roberts S, Harries J. Women's experiences seeking informal sector abortion services in Cape Town, South Africa: a descriptive study. BMC Womens Health. 2017 Oct 2;17(1):95.
85. Gile KJ, Johnston LG, Salganik MJ. Diagnostics for Respondent-driven Sampling. J R Stat Soc Ser A Stat Soc. 2015 Jan;178(1):241-69.
86. Johnston LG, Hakim AJ, Dittrich S, Burnett J, Kim E, White RG. A Systematic Review of Published Respondent-Driven Sampling Surveys Collecting Behavioral and Biologic Data. AIDS Behav. 2016;20(8):1754-76.
87. Lindberg L, Kost K, Maddow-Zimet I, Desai S, Zolna M. Abortion Reporting in the United States: An Assessment of Three National Fertility Surveys. Demography. 2020;57(3):899-925.
88. Ramirez A, Jayaweera R, Mtimkulu M, Bessenaar T, Chollampat N, Gerdts C. Understanding abortion seeking and care outside of legal public sector services in South Africa: a qualitative study. Contraception. 2017;96(4):277-8.

89. Salganik MJ. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. J Urban Health. 2006;83(1):98.
90. McCreesh N, Frost S, Seeley J, Katongole J, Tarsh MN, Ndunguse R, et al. Evaluation of respondent-driven sampling. Epidemiology (Cambridge, Mass). 2012;23(1):138.
91. Volz E, Heckathorn DD. Probability based estimation theory for respondent driven sampling. Journal of official statistics. 2008;24(1):79.
92. McLaughlin KR, Handcock MS, Johnston LG, Japuki X, Gexha-Bunjaku D, Deva E. Inference for the visibility distribution for respondent-driven sampling. JSM Proceedings Alexandria, VA: American Statistical Association. 2015:2259-67.
93. Handcock MS, Fellows IE, Gile KJ. RDS Analyst: Software for the Analysis of Respondent-Driven Sampling Data. 0.65 ed. Los Angeles2016.
94. LP SC. Stata Statistical Software Release 15: Stata Press Publication; 2017.
95. Statistics South Africa. Quarterly Labour Force Survey Quarter 3: 2018. Pretoria: Statistics South Africa2018.
96. Wejnert C. An Empirical Test of Respondent-Driven Sampling: Point Estimates, Variance, Degree Measures, and out-of-Equilibrium Data. Sociol Methodol. 2009;39(1):73-116.
97. Handcock MS, Gile KJ, Mar CM. Estimating hidden population size using respondent-driven sampling data. Electronic journal of statistics. 2014;8(1):1491.
98. Organization WH. WHO recommendations on self-care interventions: self-management of medical abortion: World Health Organization2020. Report No.: WHO/SRH/20.11.
99. Foster AM, Arnott G, Hobstetter M. Community-based distribution of misoprostol for early abortion: evaluation of a program along the Thailand–Burma border. Contraception. 2017;96(4):242-7.
100. Moseson H, Jayaweera RT, Raifman S, Keefe-Oates B, Filippa S, Motana R, et al. Self-managed medication abortion outcomes: results from a prospective pilot study. Under Review. 2020.
101. Stillman M, Owolabi O, Fatusi AO, Akinyemi AI, Berry AL, Erinfolami TP, et al. Women's self-reported experiences using misoprostol obtained from drug sellers: a prospective cohort study in Lagos State, Nigeria. BMJ Open. 2020;10(5):e034670.
102. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol. 2014;43(6):1969-85.
103. Greenland S. Multiple-bias modelling for analysis of observational data. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2005;168(2):267-306.
104. Gustafson P, McCandless LC. Probabilistic approaches to better quantifying the results of epidemiologic studies. Int J Environ Res Public Health. 2010 Apr;7(4):1520-39.
105. MacLehose RF, Gustafson P. Is probabilistic bias analysis approximately Bayesian? Epidemiology. 2012 Jan;23(1):151-8.
106. Petersen JM, Ranker LR, Barnard-Mayers R, MacLehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiological research. Int J Epidemiol. 2021.
107. Flor M, Weiß M, Selhorst T, Müller-Graf C, Greiner M. Comparison of Bayesian and frequentist methods for prevalence estimation under misclassification. BMC Public Health. 2020 2020/07/20;20(1):1135.

108. Moseson H, Jayaweera R, Raifman S, Keefe-Oates B, Filippa S, Motana R, et al. Self-managed medication abortion outcomes: results from a prospective pilot study. Reprod Health. 2020;17(1):1-12.
109. Novomestky F, Nadarajah S. truncdist: Truncated Random Variables. 1.0-2 ed2016. p. R package.
110. Hilden J. A further comment on "Estimating prevalence from the results of a screening test". Am J Epidemiol. 1979;109(6):721-2.
111. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data: Springer Science & Business Media; 2011.
112. Greenland S. Interval estimation by simulation as an alternative to and extension of confidence intervals. Int J Epidemiol. 2004;33(6):1389-97.
113. Team RC. R: A Language and Environment for Statistical Computing. 4.0.2 ed. Vienna, Austria: R Foundation for Statistical Computing; 2020.
114. Tang OS, Chan CC, Kan AS, Ho PC. A prospective randomized comparison of sublingual and oral misoprostol when combined with mifepristone for medical abortion at 12-20 weeks gestation. Hum Reprod. 2005 Nov;20(11):3062-6.
115. Tang OS, Xu J, Cheng L, Lee SW, Ho PC. Pilot study on the use of sublingual misoprostol with mifepristone in termination of first trimester pregnancy up to 9 weeks gestation. Hum Reprod. 2002;17(7):1738-40.
116. Von Hertzen H, Huong N, Piaggio G, Bayalag M, Cabezas E, Fang A, et al. Misoprostol dose and route after mifepristone for early medical abortion: a randomised controlled noninferiority trial. BJOG: An International Journal of Obstetrics & Gynaecology. 2010;117(10):1186-96.
117. Schmidt-Hansen M, Cameron S, Lohr PA, Hasler E. Follow-up strategies to confirm the success of medical abortion of pregnancies up to 10 weeks' gestation: a systematic review with meta-analyses. Am J Obstet Gynecol. 2020;222(6):551-63.
118. Moseson H, Jayaweera R, Raifman S, Keefe-Oates B, Filippa S, Motana R, et al. Self-managed medication abortion outcomes: results from a prospective pilot study. Reprod Health. 2020 Oct 27;17(1):164.
119. Jayaweera RT, Moseson H, Gerdts C. Misoprostol in the era of COVID-19: a love letter to the original medical abortion pill. Sexual and reproductive health matters. 2020 Dec;28(1):1829406.
120. Indicators S. Revised list of global Sustainable Development Goal indicators. New York: United Nations. 2017.
121. Gemzell-Danielsson K, Cleeve A. Estimating abortion safety: advancements and challenges. The Lancet. 2017;390(10110):2333-4.
122. Crane BB, Dusenberry J. Power and politics in international funding for reproductive health: the US Global Gag Rule. Reprod Health Matters. 2004 Nov;12(24):128-37.
123. Singh S, Remez L, Tartaglione A. Methodologies for estimating abortion incidence and abortion-related morbidity: a review: Guttmacher Institute; 2010.
124. Gerdts C, Hudaya I. Quality of Care in a Safe-Abortion Hotline in Indonesia: Beyond Harm Reduction. Am J Public Health. 2016 Nov;106(11):2071-5.