DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 43, ARTICLE 56, PAGES 1651–1684
PUBLISHED 16 DECEMBER 2020
https://www.demographic-research.org/Volumes/Vol43/56/
DOI: 10.4054/DemRes.2020.43.56

Research Article

Estimating abortion incidence using the network scale-up method

Elizabeth Sully

Margaret Giorgio

Selena Anjur-Dietrich

Contents

Demographic Research: Volume 43, Article 56
Research Article

Estimating abortion incidence using the network scale-up method

Elizabeth Sully[1]

Margaret Giorgio[2]

Selena Anjur-Dietrich[2]

Abstract

BACKGROUND
A major challenge in abortion research is accurately measuring the incidence of induced abortion, particularly in restrictive settings. This study tests the network scale-up method (NSUM) to measure abortion incidence, which uses respondent social network data to estimates the size of hidden populations.

METHODS
Using NSUM modules added to the Ethiopia and Uganda 2018 Performance Monitoring for Action (PMA) community-based surveys, we compute NSUM abortion incidence ratios, and adjust these ratios to account for transmission bias. We conduct internal validity checks to assess the NSUM performance.

RESULTS
The unadjusted NSUM abortion ratios were likely underestimates (Uganda: 15.3 per 100 births, Ethiopia: 3.6 per 100 births). However, the transmission bias-adjusted NSUM abortion ratios grossly overestimated abortion (Uganda: 151.4 per 100 births, Ethiopia: 73.9 per 100 births), which was likely due to selection bias, question wording, and the use of lifetime abortions to measure transmission bias. Internal validity checks revealed problems with the NSUM application in Ethiopia. Unadjusted NSUM estimates of intrauterine device/implant use performed well compared to established external estimates, but adjusting for transmission bias again resulted in overestimation.

CONCLUSION
The NSUM resulted in overestimates of abortion incidence in Ethiopia and Uganda. We discuss several modifications that may improve future applications of the NSUM for measuring abortion.

[1] Guttmacher Institute, USA. Email: esully@guttmacher.org.
[2] Guttmacher Institute, USA.

CONTRIBUTION

This is the first test of the NSUM to estimate national abortion incidence. Our findings highlight the critical need to assess the validity of abortion estimates, a key feature of the NSUM that is lacking in most other indirect abortion measurement methods.

1. Introduction

A major challenge in the field of abortion research is accurately measuring the incidence of induced abortion, particularly in clandestine or restrictive settings. Direct estimation methods, such as using administrative health records or surveying individuals directly, tend to result in an underestimation of abortion incidence (Rossier 2003; Jones and Kost 2007), especially in settings where abortion is highly stigmatized and/or hard to access. As a result, the most widely used method to estimate abortion incidence in restrictive contexts is an indirect technique known as the Abortion Incidence Complications Method (AICM). The AICM uses estimates of post-abortion care caseloads in health facilities, in combination with a multiplier that accounts for the proportion of abortions that will not necessitate facility-based medical care, to create an induced abortion incidence estimate (Singh, Prada, and Juarez 2010). This method has been successfully implemented in over 25 countries with varying degrees of restrictive abortion climates, including 11 in sub-Saharan Africa (Bankole et al. 2015; Basinga et al. 2012; Chae et al. 2017; Henshaw et al. 1998; Keogh et al. 2015; Levandowski et al. 2013; Mohamed et al. 2015; Moore et al. 2016; Polis et al. 2017; Prada et al. 2016; Sedgh et al. 2011; Singh et al. 2005; Singh et al. 2010; Sully et al. 2018).

The recent rise in the use of medication abortion, which can potentially be accessed from drug stores or the informal sector and has lower rates of complications that require care, means that abortions often do not require any interaction with the medical system. With fewer complications, the AICM estimate would need to rely more heavily on experts' judgments of abortion complication and treatment rates, making traditional AICMs increasingly less desirable to estimate abortion incidence (Sedgh and Keogh 2019). In response to these limitations, new abortion estimation methods are being developed and tested through community-based surveys, including list experiments, the confidante method, and modified approaches to the AICM (Sedgh and Keogh 2019). While these methods have several strengths, many lack a clear process for assessing the accuracy and robustness of the estimates they produce.

One promising indirect method for estimating abortion incidence is the network scale-up method (NSUM). The NSUM uses information about respondents' social networks to estimate the size of hidden populations. The NSUM was initially developed to estimate the number of deaths during the Mexico City earthquake of 1985 (Bernard et al. 1991). Since then, this method has been used to improve the measurement of hidden populations that are defined by engagement in stigmatized behaviors, such as female sex workers, men who have sex with men, and injecting drug users (Ezoe et al. 2012; RBC/IHDPC 2012; Salganik et al. 2011; Wang et al. 2015). The NSUM includes a number of internal validation checks, an assuring and rigorous feature that provides an advantage over other newly proposed indirect methods for measuring abortion incidence. Only the list experiment method includes a similar internal validation check for a design effect. However, this has not been tested by most studies employing this method, and when it has, the design effect test has shown mixed results (Bell and Bishai 2019; Elewonibi et al. 2020; Moseson et al. 2017).

Another potential advantage of the NSUM is that asking aggregate information about people in respondents' social networks may reduce the likelihood of underreporting sensitive behaviors such as abortion, as respondents may be less concerned about disclosing anonymous information about other people. Despite these potential strengths, little is known about the NSUM's ability to accurately measure abortion incidence; only one previous study has used the NSUM to measure abortion incidence (Rastegari et al. 2014). However, that study, which was conducted in Iran, was not nationally representative and did not employ the NSUM's internal validation checks.

This paper capitalizes on newly available data collected in Ethiopia and Uganda using the NSUM. The Guttmacher Institute added an NSUM module to the 2018 rounds of the Performance Monitoring for Action (PMA) female questionnaires in both countries, which was administered in interviews to a nationally representative sample of reproductive-aged women. While nonbinary individuals and transgender men also have abortions, recruitment into this sample involved respondents identified as cisgender women by the primary respondent of the household, and as such, we refer to our sample as composed of 'women.' This paper uses the NSUM to estimate each country's abortion incidence, compares the NSUM estimates to other sources of abortion data for both countries, and assesses the application of the NSUM in both settings using validation checks. These analyses provide a needed assessment of the NSUM for estimating national abortion incidence.

2. Methods

2.1 Network scale-up method (NSUM)

The foundational assumption of the NSUM is that the underlying social networks of individuals are, on average, representative of the general population. Given this, the proportion of a hidden population among the social networks of a representative sample will approximate the true proportion of that hidden population among the general population. Therefore, in order to estimate the size of a hidden population, one must first determine the size of individuals' social networks. This can be difficult, as most people find it challenging to accurately report the number of people in their social network.

In this study, we use the 'known population' approach to estimate social network sizes (Killworth et al. 1998; McCarty et al. 2001). The known population approach involves asking each respondent to report the number of people they know who have a certain characteristic. In selecting characteristics, two criteria must be met: the number of individuals in a population that have the characteristic must be known, and the characteristic must be rare enough that an individual could reasonably count all of the people in their social network with that characteristic. For example, respondents in Ethiopia were asked, "How many women do you know who live in a household that owns a camel?" If a respondent reports that she knows one woman who fits this description, and we know from the most recent Ethiopian Demographic and Health Survey (DHS) that approximately 539,000 women ages 15–49 live in a household that owns a camel, we estimate that the respondent knows 1 out of 539,000 women of reproductive age in Ethiopia. We can then estimate that the size of her social network is 50 by multiplying 1/539,000 by the total number of women age 15–49 living in Ethiopia (26,725,476). The more known population questions asked of each respondent, the more accurate the network size estimate becomes. We use an established formula for calculating personal network sizes using the maximum likelihood method (Killworth et al. 1998):

$$\widehat{c_i} = \frac{\sum_j m_{ij}}{\sum_j e_j} * t.$$

Here, $\widehat{c_i}$ is the estimated personal network size of respondent i, m_{ij} is the number of people with a particular characteristic j that respondent i knows, e_j is the size of the subpopulation with characteristic j, and t is the size of the general population (Bernard and McCarty 2009; McCarty et al. 2001).

In order to prevent outlier responses from unduly biasing social network size estimates, all known population responses are top-coded at 30, as has been done in previous NSUM studies (McCormick, Salganik, and Zheng 2010; RBC/IHDPC 2012; Salganik et al. 2011; Zheng, Salganik, and Gelman 2006).

Once an estimate of someone's personal network size is reached, the next step in the method is to estimate the size of the key population of interest (in this study, women who have had an induced abortion in the past year). Each respondent is asked how many women they know who have ever done anything to induce an abortion ("Of the women you have had contact with in the past 12 months, how many have ever done something to intentionally end a pregnancy?"). We can use this information, in combination with the personal network size estimates, to estimate the number of women who have had an abortion in each country using the following formula:

$$\widehat{e} = \frac{\sum_i (m_{ij} * \pi_i)}{\sum_i (\hat{c}_i * \pi_i)} * t.$$

In this case, $\widehat{e}$ is the estimated number of women who had an induced abortion in each country, m_{ij} is the number of women that respondent *i* knows with characteristic *j* (induced abortion), π_i is the inverse probability of selection for respondent *i*, $\widehat{c_i}$ is the estimated personal network size of each respondent *i*, and *t* is the size of the general population (Bernard and McCarty 2009; McCarty et al. 2001). In this study the sample and population of interest is only women of reproductive age; as such, *t* is defined as the number of women aged 15–49 in each country.

2.2 Adjusting for transmission effects

One assumption of the NSUM is that all respondents have perfect knowledge about all people in their social network (i.e., if someone in your social network has cancer, then you know they have cancer). Violations of this assumption are called ‘transmission effects.’ However, someone’s abortion(s) are not likely to be known by everyone in their social network. When knowledge about the hidden population is not complete in a social network, it is necessary to estimate the ‘visibility,’ or τ, which can be used to adjust the NSUM estimate for transmission bias. For example, if women who have abortions only told 20% of their social network, then we would consider the visibility of abortion to be 0.2. We then adjust the NSUM estimator by 1/0.2 to account for the transmission rate. Without this adjustment, the NSUM estimator would underestimate the number of women who had abortions by a factor of 5. Transmission bias adjustments are calculated for all women and are not subgroup specific. Subgroup specific transmission bias adjustments would require detailed characteristics for all reported people in the known population, which we do not have. The updated NSUM estimator adjusting for transmission bias is as follows (Salganik et al. 2011):

$$\widehat{e} = \frac{\sum_i (m_{ij} * \pi_i)}{\sum_i (\hat{c}_i * \pi_i)} * t * \frac{1}{\tau}.$$

One drawback of the NSUM is the difficulty in determining the value of τ. Several previous studies testing different methods to estimate transmission bias have proven unsuccessful for a variety of reasons (Guo et al. 2013; Killworth et al. 2006; Shelley et al. 2006). The most rigorous and valid estimates of the social visibility of hidden groups are derived from the Game of Contacts method, developed by Salganik et al. (2011). This method involves recruiting a separate sample of members of the hidden population of interest and estimating visibility through a game-like activity. While results from this method have been promising (Maghsoudi et al. 2014; Salganik et al. 2011; Sirbiladze et al. 2015), the main downside is the additional resources required to nest this smaller study within a larger community-based survey. While the Game of Contacts has been tested for estimating socially connected hidden populations such as intravenous drug users or men who have sex with men, its applicability to measuring transmission bias for abortion is unknown. Obtaining an abortion does not necessarily involve social interaction with other women obtaining abortions, which could make women who have had abortions less socially connected than other hidden populations that the Game of Contacts has been used to estimate. In the current study, we test a novel method for estimating transmission bias that does not require a separate data collection effort. Instead, we ask women who directly report their abortions in the PMA surveys how many people in their social network they have told about their abortion(s). We then use the inverse of this proportion as our estimate of τ.

2.3 Data sources and sample

Data for this analysis come from the 2018 PMA surveys in Uganda and Ethiopia (PMA 2018). These surveys include a nationally representative sample of enumeration areas (EAs) in each country. EAs are sampled using a two-stage cluster design for rural/urban residential areas and geographic regions. Households in the selected EAs are then randomly selected, and all women aged 15–49 residing in selected households are invited to participate in the survey. In 2018, there were 4,288 women interviewed in Uganda and 7,546 in Ethiopia, half of which were randomized to answer the NSUM module (Uganda: n = 2,161; Ethiopia: n = 3,755). The other half of respondents completed a separate module testing the confidante method, a different social network-based method for estimating abortion incidence.

Respondents were excluded from the final analytic sample if they did not provide valid responses to all 12 known population questions included in the NSUM module (Uganda: n = 77, Ethiopia: n = 191). This resulted in a final analytic sample of 2,084 respondents in Uganda (95% of those randomized to the NSUM) and 3,564 respondents in Ethiopia (93% of randomized respondents). In order to determine whether the exclusion of these women may have introduced bias into our final estimates, we examined differences in key demographic characteristics based on inclusion status. In Uganda, the only significant difference was based on region. However, the magnitude of this difference was small (Supplemental Table A). In Ethiopia, we found that larger proportions of excluded women had no or low levels of education compared to women included in the sample (Supplemental Table B). We consider the implications of these differences later in the Discussion.

We used 2016 DHS data from Uganda and Ethiopia to identify appropriate characteristics for the known population questions. Known population sizes were calculated using the 2016 DHS and the 2019 revision of the World Population Prospects' estimates for the population of women age 15–49 in 2018.

2.4 Data collection

Before fielding the NSUM, pilot tests in the community and focus group discussions with resident enumerators (REs) were conducted in each country. The pilots included face validity questions to ensure that language used to describe social networks and abortions would be interpreted correctly by participants in each context. Face validity questions included (1) "Please describe how you interpreted which people in your life we were referring to in the last question [NSUM introduction/prompt]" and (2) "Please describe how you interpreted the phrase 'intentionally ended a pregnancy.'" The latter question was intended to ensure differentiation between successful and unsuccessful abortions and miscarriages, and provided an opportunity to ensure that REs knew how to properly clarify the question or probe the respondent if needed. The focus group discussions held among REs provided insight into the appropriate criteria for defining a member of a respondent's social network as well as an opportunity to ensure that selected known populations were appropriate in each country.

The main data collection effort occurred in 2018, April–May in Uganda and June–July in Ethiopia. Surveys were administered by female REs in several major languages in each country (with translators assisting REs in some EAs in Ethiopia). The PMA core survey included several modules on various health-related topics, including contraceptive use and fertility history. The NSUM module occurred after the PMA core questions but prior to the section on direct reports of abortion.

2.5 Measures

Previous NSUM studies have typically used two different methods for defining what it means to 'know' someone, which has implications for how someone's social network is identified (RBC/IHDPC 2012). The more conservative definition aims to only include stronger network ties. When using this definition, a respondent is asked to think of individuals who they know by sight and name, who also know the respondent by sight and name, who live in a specified geographic area, and who the respondent shared a meal or drink with in the past 12 months. This study originally planned to use this 'meal' definition. However, during the pilot it was determined that women in Uganda and Ethiopia generally do not socialize with their friends and extended family members in this way; thus, using this definition would systematically exclude appropriate social ties. Instead, the more basic definition of to 'know' was used, which removes the meal/drink requirement and instead stipulates that contact has occurred (in person, by phone, over computer) in the past 12 months (see Appendix A).

Several NSUM questions were added to the 2018 female PMA surveys in each country. First, 12 known population measures were included, each using the more basic definition of a social network tie (see Table 1 and Table 2). Criteria for selecting known populations in each country included that the population be rare enough that respondents would be able to accurately count the number of women within their social network who were members of that population, or populations that were notable enough to easily recall (e.g., women who had a live birth in the last 12 months). As a result, the size of the known populations, on average, was 3.3% of women of reproductive age in Ethiopia and 8.3% in Uganda. Additional criteria for selection were that the characteristic that defined the population must be knowable or visible to the respondents, and that there was balance in the likelihood of knowing women in different populations (e.g., if some known populations were correlated with urban residence, other populations were selected that were correlated with rural residence). Appropriate known populations were determined using 2016 DHS data; five questions were asked in both countries, and seven questions were specific to Ugandan and Ethiopian contexts.

The questions used to measure induced abortion in both Uganda and Ethiopia were "Of the women you have had contact with in the past 12 months, how many have ever done something to intentionally end a pregnancy?" and "Thinking of these X women who you have had contact with in the past 12 months and who have ever ended a pregnancy, how many have ended a pregnancy in the past 12 months?" Finally, questions were included to measure transmission bias. Women who self-reported ever having an induced abortion were asked how many women in their social networks know that they had ever intentionally ended a pregnancy.

2.6 Analytic plan

We calculated abortion incidence estimates for each country using the NSUM procedures described above. To produce a certainty interval for the NSUM estimates, we used a rescaled bootstrap variance estimation procedure to generate 5,000 replicate samples with which to produce replicate estimates. We drew our 95% certainty intervals from this set of estimates; it is important to note that previous NSUM studies have found that this method produces extremely narrow confidence intervals representing actual coverage of closer to 10% (Feehan and Salganik 2016). The rescaled bootstrapping technique is appropriate in this study over a standard bootstrap procedure, which assumes a random sample. PMA data are collected using a complex sample design (primary sampling unit = EA, stratified by region and urban/rural residence), which can mostly be accounted for in the rescaled bootstrapping method, with the exception of post-stratification adjustment for household-level response rate within EAs (Feehan and Salganik 2016).

We compare the NSUM abortion ratios (with and without the transmission bias adjustment) with direct report abortion incidence estimates from the PMA survey, and with the most recent AICM estimates for both countries (Moore et al. 2016; Prada et al. 2016). Direct report abortion incidence was calculated only among self-reported abortions in the last 12 months to be comparable to the NSUM abortion estimate. We use abortion ratios rather than rates to take into account changes in fertility between each AICM (2013 in Uganda and 2014 in Ethiopia) and our NSUM and direct report estimates in 2018. While it is possible that abortion trends have changed over time in both countries, limiting the comparability with older AICM estimates, regional data for Eastern Africa show a relatively stable abortion rate in 2010–2014 and 2015–2019 (Bearak et al. 2020). We do not use the abortion ratios reported in the respective AICM studies; instead, we compare our results to an updated abortion ratio calculated using the number of abortions reported by each study together with more recent estimates of total births in the corresponding years, from the 2019 revision of the World Population Prospects. This update resulted in slightly higher abortion ratios than those reported by Prada et al. (2016) with an abortion ratio in Uganda of 20.6 per 100 births, updated from 19, and Moore et al. (2016) with an abortion ratio in Ethiopia of 18.5 per 100 births, updated from 17.6. The AICM is considered the current best available estimate of abortion in both Uganda and Ethiopia; there is not a gold standard estimate of abortion with which to compare the NSUM estimates.

NSUM analyses were originally conducted using Stata version 16.0 (StataCorp LP, College Station, TX), and then NSUM estimates were calculated again with the rescaled bootstrapping in R 4.0.2 (R Core Team 2020) using the networkreporting and surveybootstrap packages (Feehan 2016; Feehan and Salganik 2014).

2.7 Internal validity checks

In order to test how the NSUM performed, we step-wise removed one known population, treated that known population as if it is a hidden population, and produced an NSUM estimate of that known population (Feehan et al. 2016). Take women who are current smokers as an example. We pretend that we do not in fact know the number of current smokers. We then then use the remainder of the known populations to estimate the respondents' network sizes and apply the scale-up method to estimate the number of current smokers. To test the accuracy of this back-estimate, the newly estimated population size is compared to the known number of women who smoke from the DHS. This can be repeated for every known population measured in the survey. The extent to which the back-estimates mirror the known population sizes provides confidence in the relative accuracy in the estimates for the size of the hidden population, which in this study, was induced abortion. This back-estimation of known populations has been successfully used in a number of previous NSUM studies, with back-estimates of 50–200% of the known population size considered suitable for the method and 100% as an indicator of ideal performance of the NSUM (Guo et al. 2013; Habecker, Dombrowski, and Khan 2015; Kadushin et al. 2006).

We also use the NSUM to estimate another reproductive health behavior, the use of IUDs or implants, for which we have a known estimate. Assuming that sharing information on contraceptive use may be similar to how women share information on abortion, we asked respondents about IUD and implant use, treating this as a temporarily unknown population. We similarly measure the transmission bias for IUD and implant use. The accuracy of the NSUM estimate of women using IUDs and implants is used to assess the validity of the NSUM estimates; while this does not directly validate the NSUM abortion estimate, it indicates overall how well the NSUM performed and whether it was able to accurately estimate other reproductive health behaviors.

2.8 Sensitivity tests

In recent years, a number of alternative strategies to generating the NSUM estimator have been proposed (Guo et al. 2013; Habecker, Dombrowski, and Khan 2015). In this paper, we test the sensitivity of our estimates to changes in the inclusion criteria for respondents as well as for the accuracy in estimating the size of known populations. First, we restrict our analytic sample to respondents who provided a minimum of two non-zero responses to the NSUM questions. Second, we use a recursive process to identify and remove the known populations whose size is predicted with less precision from the generation of respondents' social network sizes. These tests are described in further detail in Appendix B, and their results summarized in Supplemental Table C.

3. Results

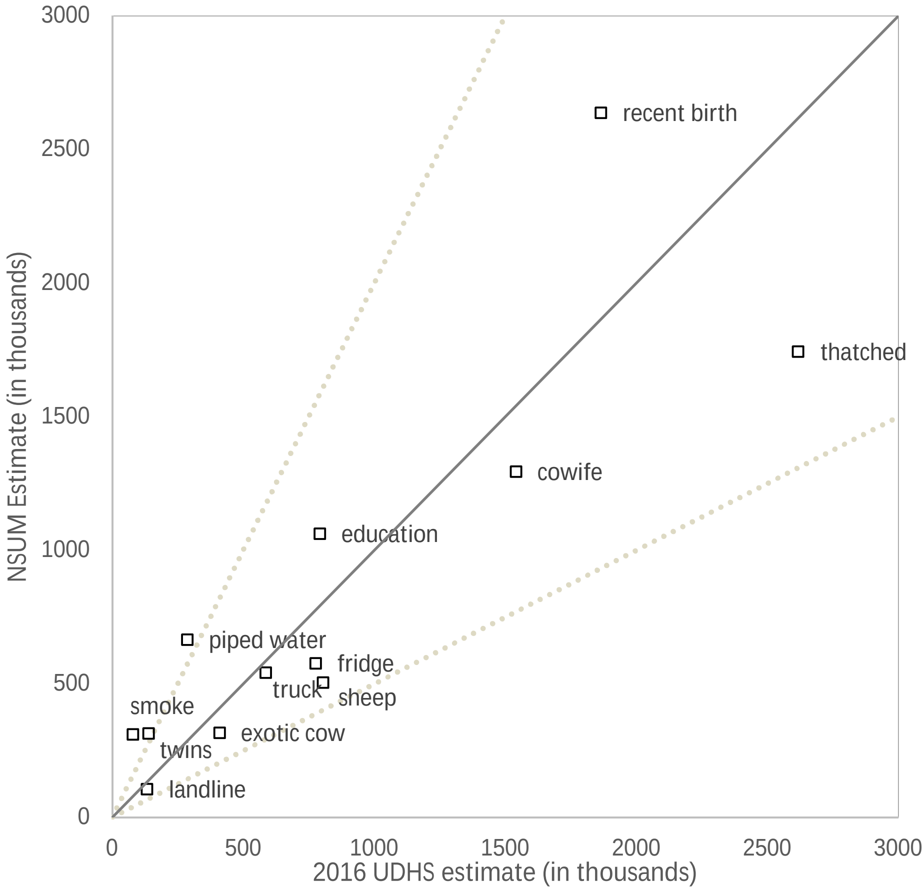
Table 1 shows the size of each known population, the mean number of social network connections reported for each known population in Uganda, and the mean number of connections after top-coding at 30. Figure 1 shows the relationship between the known size of each population (based on the 2016 Uganda DHS) and our estimated population sizes, generated using the back-estimation process as part of the internal validity check (also displayed in Table 1).

Table 1: Known populations used in Uganda, women ages 15–49

Category of population	Size	Mean number of connections	Mean number of connections, top-coding at 30	Initial NSUM estimate (as % of DHS estimate)
Gave birth in last 12 months	1,866,029	6.1	4.2	141%
Most recent birth was a multiple birth	139,930	0.6	0.6	224%
Has at least one co-wife	1,542,103	2.5	2.5	84%
Attended any education past senior six	793,764	2.2	2.0	134%
Smokes a pipe or cigarettes	80,308	0.6	0.6	386%
Lives in a household:				
...with a thatched roof	2,619,087	5.9	4.2	67%
...that owns a car or truck	587,313	1.2	1.1	92%
...that has a refrigerator	777,497	1.3	1.2	74%
...that owns an exotic cow	410,925	0.6	0.6	77%
...that owns at least one sheep	806,281	1.2	1.2	62%
...that has a landline	134,188	0.2	0.2	78%
...that has piped water inside the home	288,430	1.3	1.2	231%

Source: 2016 Uganda DHS, female (FQ) and household (HHQ) questionnaires.

Figure 1: Comparison of known population sizes from 2016 Uganda DHS to network scale-up estimates



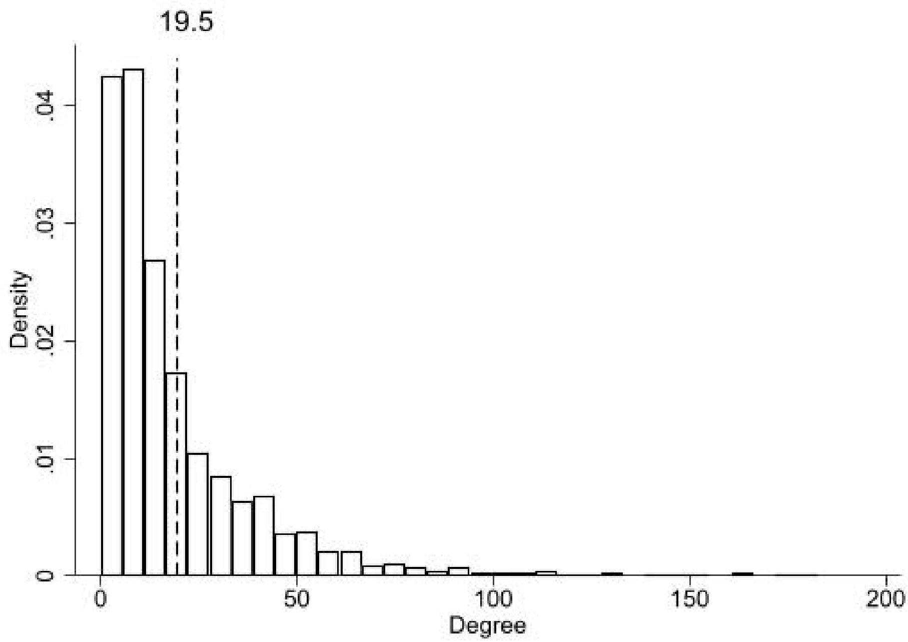
Note: Dashed lines indicate NSUM estimates that are <50% or >200% of the known population size.

This validity check shows that the NSUM estimates of known populations are relatively close to the true population size without a consistent pattern of over- or underestimation. Out of all 12 populations, the largest outliers are women who smoke a pipe or cigarettes (386% of DHS estimate), women who live in a household with piped water (231% of DHS estimate), and women whose most recent birth was a multiple birth

(224% of the DHS estimate). These results suggest that the NSUM is performing relatively well at measuring non-hidden population sizes in Uganda.

The average social network size for women in Uganda is 19.5, with a range of 0–182 (Figure 2). This means that, on average, each respondent knows 19.5 women by sight and name who are between the ages 15 and 49, who live in Uganda, and who the respondent was in contact with (in person, by phone, over computer) in the past 12 months.

Figure 2: Degree distribution from NSUM in Uganda



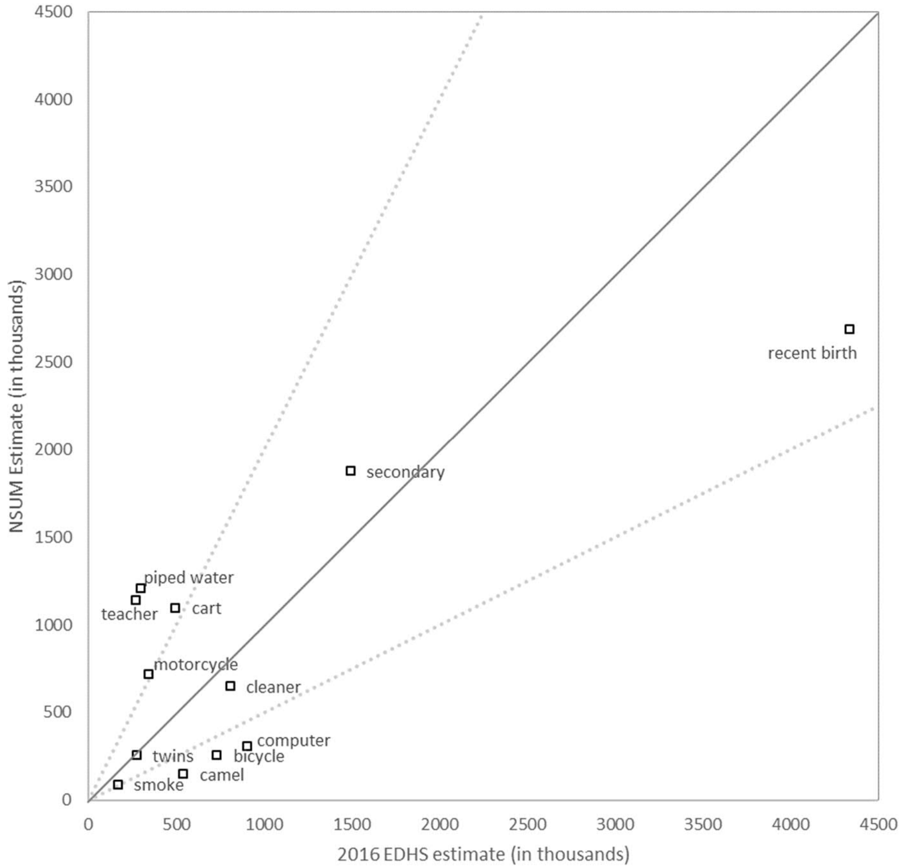
We applied the same initial analytic process and internal validity checks in Ethiopia. From the beginning of the analysis, there was evidence that the NSUM did not perform as well in Ethiopia as it did in Uganda. The mean number of connections is equal to or less than one in eight out of the 12 known population questions, indicating that a large number of respondents reported knowing no one in the specific known population (Table 2). Additionally, the estimated/known population size ratios are outside of the acceptable range for seven questions in this initial analysis and close to the cut-off for one additional question. Figure 3 graphically shows this lack of precision found through the internal validity check process. Given the weaker results of the internal validity checks in Ethiopia, we tested for but did not find evidence of enumerator effects on data quality in terms of the proportions of valid and zero responses to NSUM questions (Supplemental Table D).

Table 2: Known populations used in Ethiopia, women ages 15–49

Category of population	Size	Mean number of connections	Mean number of connections, top-coding at 30	Initial NSUM estimate (as % of DHS estimate)
Gave birth in last 12 months	4,335,398	3.2	2.7	62%
Most recent birth was a multiple birth	274,902	0.2	0.2	93%
Works as a cleaner	806,794	1.0	0.9	80%
Works as a teacher	269,614	1.9	1.2	424%
Smokes a pipe or cigarettes	165,682	0.1	0.1	55%
Attended any school past secondary	1,494,784	2.5	2.3	126%
Lives in a household:				
...that has piped water inside the home	295,552	1.8	1.6	409%
...that owns a computer	901,877	0.6	0.6	34%
...that owns an animal-drawn cart	495,617	0.7	0.7	221%
...that owns a scooter or motorcycle	341,059	0.5	0.5	211%
...that owns a bicycle	732,308	0.3	0.3	35%
...that owns at least one camel	538,923	0.2	0.2	27%

Source: 2016 Ethiopia DHS, female (FQ) and household (HHQ) questionnaires.

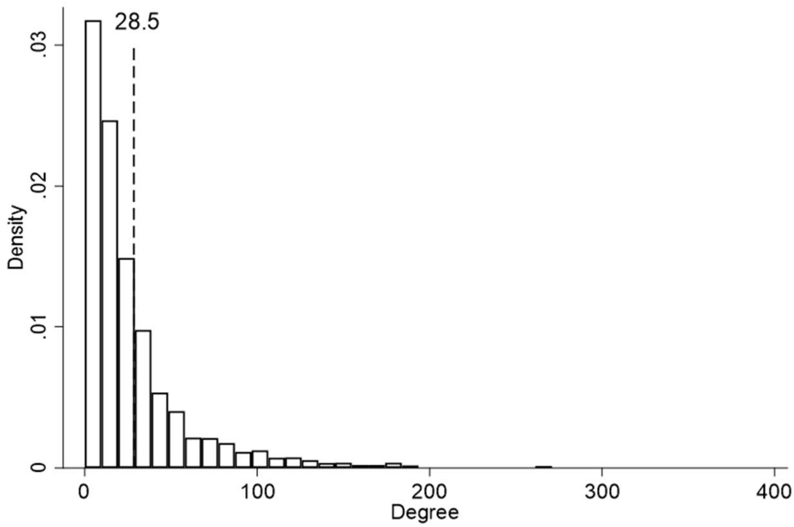
Figure 3: Comparison of known population sizes from 2016 Ethiopia DHS to network scale-up estimates



Note: Dashed lines indicate NSUM estimates that are <50% or >200% of the known population size.

The average degree for women in Ethiopia is 28.5 when estimated based on all 12 known populations, with a range of 0–339 (Figure 4).

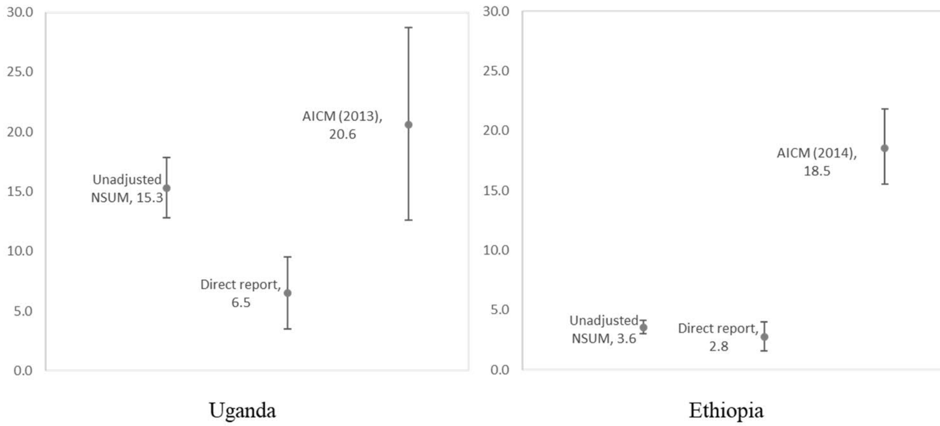
Figure 4: Degree distribution from NSUM in Ethiopia



3.1 NSUM estimation of abortion incidence

Figure 5 compares abortion ratios for Uganda and Ethiopia from the unadjusted NSUM, the most recent AICM estimates in each country (Moore et al. 2016; Prada et al. 2016), and from the direct report abortion questions in the PMA 2018 surveys. In both countries, the direct report estimates are the lowest (Uganda: 6.5 per 100 births, 95% CI 3.5–9.5; Ethiopia: 2.8 per 100 births, 95% CI 1.6–4.0). Before adjusting for the visibility of this hidden population, the NSUM estimates are higher than the direct reports but lower than the most recent AICM abortion ratios (Uganda: 15.3 per 100 births, 95% CI 12.8–17.9; Ethiopia: 3.6 per 100 births, 95% CI 3.0–4.2). While the unadjusted NSUM estimate in Ethiopia was still quite low (19% of the AICM estimate), the Ugandan unadjusted estimate was much closer to the recent AICM (74% of the estimate).

Figure 5: Comparison of abortion ratio estimates: Unadjusted NSUM, most recent AICM, and direct report (abortions per 100 births)



Note: AICM abortion ratios calculated using total number of abortions reported by Prada et al. (2016) for Uganda and Moore et al. (2016) for Ethiopia, and estimates for total births in the corresponding years from the World Population Prospects (2019).

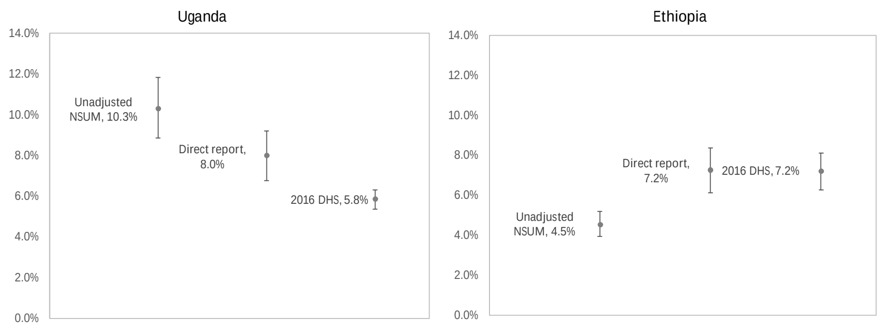
Next, we tested our novel approach to assessing transmission bias by calculating the visibility factor among women who self-report an abortion (n = 121 in Uganda, n = 114 in Ethiopia); women in Uganda told approximately 10%, and women in Ethiopia told approximately 5% of their social networks about their abortions. Supplemental Figure A displays the NSUM abortion incidence estimates after adjusting for transmission bias, which inflates the unadjusted ratio to 156.3 abortions per 100 births in Uganda and 74.2 abortions per 100 births in Ethiopia.

There are important limitations to these adjusted estimates. Transmission bias varies across key sociodemographic characteristics, which we cannot account for in this study. In both Uganda and Ethiopia women living in rural areas, who had never attended school and who were currently married, reported that more women in their social networks knew about their abortions (Supplemental Table E). Further, in Ethiopia there are also important differences in the social network size of the women who self-reported abortions compared to all women. The average social network size in Ethiopia overall was 28.5, but the subset of women who self-reported an abortion, from which transmission bias was estimated, had a social network size of 44.6 (data not shown). Conversely, in Uganda there was little difference in the social network size of women self-reporting abortions (19.5 overall versus 20.2 among women self-reporting an abortion).

3.2 NSUM estimation of IUD and implant use

We calculated the NSUM 12-month prevalence of IUD and implant use to assess how well the method performed in estimating a different and less socially stigmatized reproductive health behavior. Figure 6 compares the 12-month IUD/implant prevalence estimates for Uganda and Ethiopia calculated using the NSUM (unadjusted for transmission bias), the 2016 DHS in each country, and from the direct report questions about contraception use in the PMA 2018 surveys. In Uganda, the results show that the NSUM estimate is comparable to the PMA direct report, but it is slightly higher than the older DHS estimate from 2016 (NSUM: 10.3%, 95% CI 8.9–11.8; direct report: 8.0%, 95% CI 6.8–9.2; DHS: 5.8%, 95% CI 5.4–6.3). In Ethiopia, the direct report and DHS estimates are similar, while the NSUM estimate is somewhat lower (direct report: 7.2%, 95% CI 6.1–8.4; DHS: 7.2%, 95% CI 6.3–8.1; NSUM: 4.5%, 95% CI 3.9–5.2).

Figure 6: Comparison of IUD/implant prevalence estimates: Unadjusted NSUM, 2016 DHS, and direct report



We again calculated the visibility factor among women who self-reported current implant and IUD use (n = 170 in Uganda, n = 258 in Ethiopia) and found that in Uganda, these women estimated that approximately 14% of their social network knew about their use of an IUD or implant, and women in Ethiopia reported that, on average, 10% of their social network knew about their contraceptive method. Supplemental Table C displays the NSUM IUD/implant prevalence estimates after adjusting for transmission bias, which inflates the baseline estimate to 71.9% in Uganda and 45.6% in Ethiopia.

3.3 NSUM estimates from sensitivity tests

The three different estimation techniques produced similar abortion incidence estimates, and similar estimates of IUD/implant prevalence (see Supplemental Table C). The lack of change in results suggests that our findings are not sensitive to the inclusion criteria implemented in our analysis nor to the choice of known populations among the original 12 for which we collected data.

4. Discussion

The results of this study highlight several strengths and weaknesses of the NSUM's utility in measuring abortion incidence across different contexts. Overall, we find that the NSUM's application in measuring abortion incidence was not successful but that there is scope for modifications that could reduce potential sources of bias. While it is likely that the unadjusted estimate is an improvement over direct reports, the adjusted estimate is clearly an overestimate of the true incidence of abortion. The bias in the adjusted estimate could be driven by the method we used to measure abortion visibility, underlying biases in the unadjusted estimate, or a combination of the two. In this discussion, we provide several recommendations for removing potential biases in both the unadjusted estimate and the measure of transmission bias.

Our novel attempt to adjust for transmission bias resulted in an unreasonably high estimate of abortion, suggesting that the true value of abortion incidence likely falls somewhere between the NSUM estimate without transmission bias and the gross overestimate produced with the transmission bias adjustment. Transmission bias for the NSUM has previously been assessed through a companion respondent-driven sampling study of individuals with the hidden characteristics of interest. We did not have the ability to field such a study in 2018; we instead tested a novel approach to measuring transmission bias by asking questions directly to respondents who self-reported an abortion. The failure of this first attempt at implementing our novel approach to the transmission bias adjustment could be due to a number of factors. First, there is likely selection bias in the sample of respondents who self-reported their abortion. We know that individuals who self-report an abortion are likely not representative of all people having abortions. If those who self-report abortions are also different in how they talk about abortion in their social networks, then the transmission bias adjusted estimate may be biased. In addition, the failure of the transmission bias adjustment may be due to the wording of the question that measured transmission bias; we asked respondents who had abortions how many people in their social network know about their abortion. However, social network members may have knowledge of that abortion either from being told directly or indirectly, and respondents may not have full awareness of third parties disclosing their abortions. If many people are finding out about abortions through other, indirect channels, such as family connections or gossip, then the transmission bias is overcorrecting for the visibility of abortion and leading to an overestimation. Little is known about the motivations, circumstances, and timing of abortion disclosures between social network members. Future studies should investigate these factors to identify the best way of measuring transmission bias.

We used data on abortion visibility from respondents who reported ever having had an abortion, not just those who did so in the last 12 months, due to the small size of the sample of respondents self-reporting abortions in our survey in each country. It is likely that more individuals know about lifetime abortions compared to one that happened very recently. As such, it is possible that the average portion of these individuals' social networks that know about their abortion is larger than for the (more recent) abortions used to generate the NSUM estimate. This would result in a smaller correction estimate for transmission effects than might have been calculated with a larger sample of more recent abortions. However, it still appears that our adjusted estimates are vastly over-inflated.

This study's definition of a social tie may also have contributed to a lack of visibility of recent abortions in respondents' social networks, as it only required that the respondent had contact with the other individual in the last 12 months. Abortions may have occurred in the respondent's social network after the most recent contact. Further work should investigate how differences in the time component of social tie definitions influences incidence estimates of rare outcomes such as abortion. In addition, prior research on mixed-gender networks found that a narrower tie definition (sharing a meal) outperformed the social tie definition we used in our analysis (RBC/IHDPC 2012). While the meal tie definition was deemed inappropriate for women's social networks in Ethiopia and Uganda, future research should experimentally test whether narrower tie definitions specific to women's social networks further reduces the error in the internal validity checks.

To date, the only known way to successfully generate transmission bias estimates for other hidden populations is through the use of respondent-driven sampling (RDS) and the Game of Contacts methodology (Salganik et al. 2011). However, while several hidden populations have been successfully recruited through RDS (e.g., injecting drug users, female sex workers) (Maghsoudi et al. 2014; Salganik et al. 2011), it is unknown whether this network-based methodology is appropriate for sampling people who have abortions. In addition, employing the Game of Contacts adds additional cost to a study that must also field a community-based survey, as is the case for the NSUM. Conversely, the results of this study highlight potential areas of improvement for the measurement of transmission bias through community-based surveys. This could be done through (1) modifying question wording to clearly differentiate whether respondents know indirectly or were told directly about an abortion, (2) expanding the sample size to try to capture enough self-reported abortions to generate a transmission bias estimate from only the last 12 months, and (3) exploring the use of a narrower social tie definition. Future research is needed to both test the Game of Contacts methodology in its application for people who have abortions as well as test our proposed improvements for the measurement of transmission bias through community-based surveys.

A second reason why our application of the NSUM failed to produce valid abortion incidence estimates is that the underlying unadjusted estimate may have also been biased; such a bias would only be further magnified through the transmission adjustment. For example, the interviewer training manual included explicit instructions to only report successful abortion attempts. However, the exact language in the English version of the questionnaire did not specify that only successful abortions should be included. The translated questionnaires, which were almost exclusively used during data collection, used local terminology for describing abortions, and such terminology may have left open the possibility of respondents reporting unsuccessful abortion attempts or miscarriages. Further, while we tested this question prior to data collection to ensure that respondents were able to accurately report successful induced abortions only, our pilot was only conduced in urban areas. If there was geographic variation in a respondent's interpretation of the study definition of induced abortion, our results may have been unduly biased. Finally, the sensitive and stigmatized nature of abortion may have resulted in continued underreporting of the practice, even with third-party reporting. If this was the case in either study setting, our unadjusted estimate may have been biased downwards.

We did find, however, that the unadjusted NSUM performed much better in the validation exercise of estimating IUD and implant use than for estimating abortion incidence in both countries. This suggests that for less stigmatized reproductive health behaviors, the NSUM is able to produce reliable estimates. However, we expected some adjustment for transmission effects to be necessary; while IUD and implant use is less stigmatized than abortion, people are unlikely to disclose their contraceptive use to everyone they know. Yet, the transmission bias adjustment for IUD and implant use resulted in an overestimate of these behaviors as well. Visibility of IUD and implant use was calculated from a larger sample with less selection bias than abortion visibility, suggesting that the resulting transmission adjustment would also be less biased. Given this, we would have expected the transmission bias adjustments to be smaller and more successful for IUD and implant use. The fact that the adjusted NSUM estimate was a gross overestimate suggests that our transmission bias adjustment is not working correctly. It is not surprising, then, that this approach also did not work well for abortion, where transmission effects are likely to be much larger.

While it is assuring that the unadjusted NSUM is able to approximate the prevalence of IUD and implant use, the value of the method comes from its utility in estimating hidden behaviors that are underreported on traditional surveys. More work is needed to improve on the measurement of transmission bias in order to fully assess the NSUM's application for estimating abortion incidence.

The poor performance of the NSUM in Ethiopia, indicated by the limited success in back-estimating known population sizes, suggests that NSUM estimates of abortion incidence should be interpreted cautiously. A higher proportion of respondents stated that they did not initially understand the NSUM instructions in Ethiopia (4.6%) as opposed to Uganda (1.1%), suggesting that our sample in Ethiopia may have had more difficulty in correctly interpreting the module and providing accurate responses. It is likely that not all respondents who did not understand the instructions to the NSUM module would be willing to state this to the RE, in which case more respondents may have struggled to accurately complete the module. In addition, the respondents in the excluded sample were more likely to have had no education. It is possible that in settings with lower education levels the NSUM is too cognitively taxing and difficult for people with lower numeracy skills to answer accurately.

There were also challenges in the fielding of the methodology to consider, including an urban-only pilot, a large number of interviewers being trained at once, new REs (a group that yielded higher rates of invalid responses) receiving a separate training, and some use of translators in the field who were not present at training and therefore unfamiliar with the NSUM. The NSUM is a complex methodology to field, making it particularly sensitive to the quality and type of training enumerators receive. Any of these factors may have limited interviewer or translator comprehension of the method, which in turn may have resulted in lower quality data being collected.

Future applications of this method should also carefully consider the selection of known populations; we found poorer back-estimation of the known populations that were either the smallest or largest population groups, such as women who smoke cigarettes or a pipe in both Uganda and Ethiopia. We selected known populations in Ethiopia that were on average around 3.3% of the population of women of reproductive age, but this ranged between 0.6% and 16.2%. In Uganda, the known population average was 8.3% of the population of women of reproductive age but ranged from 0.8% to 26.0%. The tendency to overestimate smaller population groups and underestimate larger population groups has been documented in prior NSUM studies (Maltiel et al. 2015). Known populations should also be selected to carefully avoid populations that may be more or less likely to be socially connected. For example, the estimate of women who worked as teachers was nearly 400% of the DHS estimate, but teachers, by nature of their profession, are often more socially connected and may have very different social network structures. Known populations' social network structure and overall network size should be considered carefully in selecting populations for NSUM surveys.

The foundational assumption of the NSUM is that people's social networks are, on average, representative. The only way of testing this assumption is through the back-estimation internal validity checks, which we found performed well in Uganda but not in Ethiopia. However, we are unable to determine whether the back-estimation process failed due to the challenges of fielding the survey, potentially some incorrectly selected known populations, or because the foundational assumption may have been violated.

Finally, it may be that social network approaches to investigate sensitive behaviors are not appropriate in some cultural contexts. The poor performance of the NSUM in Ethiopia may not only be due to poorly estimated social network sizes (as evidenced through the back-estimation process), but they may also reflect a lack of sharing information about sensitive reproductive health behaviors within social networks. Future research is needed to better understand whether the application of social network-based methods for estimating abortion incidence is appropriate in diverse cultural settings.

While the poor performance of the NSUM in Ethiopia is concerning, our ability to assess the performance of the method is a crucial feature of the NSUM. Most other abortion estimation methods in restrictive settings lack a similar feature. There are rarely internal validation checks available to determine how well a method performed, or how accurate our estimates are. This is one critical advantage to the NSUM. However, these internal validity checks can only assess the validity of the denominator; they do not address biases in the numerator. The numerator relies on respondents being able to accurately report the abortions they do know about, and transmission bias adjustments to adequately adjust for abortion visibility. Even if the internal validity checks had performed well in both countries, this would not guarantee that the NSUM estimates were unbiased.

Overall, our application of the NSUM to measure abortion incidence in Ethiopia and Uganda shows that social network-based methods for measuring abortion are viable, as respondents knew and were willing to report on abortions in their social networks. With modifications to the transmission bias adjustment, the NSUM may be a promising method for measuring abortion incidence. However, until such adjustments have been tested and validated, the NSUM does not present an improvement over the AICM, the current most commonly used method. Other indirect methods being tested, such as the list experiment method or confidante method, are in a similar position to the NSUM with regard to what we understand about their usefulness in measuring abortion incidence. The list experiment has a design effect test, but it has either not been employed in prior studies or shown mixed results when tested (Bell and Bishai 2019; Elewonibi et al. 2020; Moseson et al. 2017). The confidante method lacks appropriate validations strategies to assess its performance, making it difficult to understand the many ways in which biases are likely influencing the resulting estimates. Future research into indirect methods of improving abortion incidence must prioritize internal and external validation checks, which are essential in both assessing how the method performed and being able to evaluate it in comparison to other currently available methods of abortion measurement.

5. Conclusions

This study is an important first step in testing the applicability of the NSUM to estimate abortion incidence in restrictive settings. However, we do not believe that the estimates provided in this paper accurately reflect the true incidence of abortion in Ethiopia or Uganda. In particular, more work is needed to better understand the magnitude of transmission bias for abortion in each country. Further, a better understanding of the reasons for the poor performance of the NSUM, particularly in Ethiopia, is needed. This additional information will help refine the application of the NSUM for estimation of abortion incidence in order to put forward recommendations for its potential expanded use in the field of abortion measurement.

6. Acknowledgments

We greatly appreciate the work of our partners at PMA Ethiopia, Solomon Shiferaw and Assefa Seme; PMA Uganda, Frederick Makumbi, Simon Peter Kibira; and Johns Hopkins University School of Public Health, Suzanne Bell, for their collaboration in fielding this module. We would also like to thank our colleagues for their feedback and support of this work, including Ann Biddlecom, Laura Lindberg, Onikepe Owolabi, and Isaac Maddow-Zimet.

References

Bankole, A., Adewole, I.F., Hussain, R., Awolude, O., Singh, S., and Akinyemi, J.O. (2015). The incidence of abortion in Nigeria. *International Perspectives on Sexual and Reproductive Health* 41(4): 170–181. doi:10.1363/intsexrephea.41.4.0170.

Basinga, P., Moore, A.M., Singh, S.D., Carlin, E.E., Birungi, F., and Ngabo, F. (2012). Abortion incidence and postabortion care in Rwanda. *Studies in Family Planning* 43(1): 11–20. doi:10.1111/j.1728-4465.2012.00298.x.

Bearak, J.M., Popinchalk, A., Ganatra, B., Moller, A.B., Tunçalp, Ö., Beavin, C., Kwok, L., and Alkema, L. (2020). Unintended pregnancy and abortion by income, region, and the legal status of abortion: Estimates from a comprehensive model for 1990–2019. *Lancet Global Health* online. doi:10.1016/S2214-109X(20)30315-6.

Bell, S.O. and Bishai, D. (2019). Can a list experiment improve validity of abortion measurement? *Studies in Family Planning* 50(1): 43–61. doi:10.1111/sifp.12082.

Bernard, H.R. and McCarty, C. (2009). The network scale-up method: Background and theory. http://nersp.osg.ufl.edu/~ufruss/scale-up/scale-up%20method%20theory%20and%20history%20with%20notes.pdf.

Bernard, H.R., Johnsen, E.C., Killworth, P.D., and Robinson, S. (1991). Estimating the size of an average personal network and of an event subpopulation: Some empirical results. *Social Science Research* 20(2): 109–121. doi:10.1016/0049-089X(91)90012-R.

Chae, S., Kayembe, P.K., Philbin, J., Mabika, C., and Bankole, A. (2017). The incidence of induced abortion in Kinshasa, Democratic Republic of Congo, 2016. *PLOS ONE* 12(10): e0184389. doi:10.1371/journal.pone.0184389.

Elewonibi, B., Amour, C., Gleason, S., Msuya, S., Canning, D., and Shah, I. (2020). Estimating the lifetime incidence of induced abortion and understanding abortion practices in a Northeastern Tanzania community through a household survey. *Contraception.* doi.org/10.1016/j.contraception.2020.10.013.

Ezoe, S., Morooka, T., Noda, T., Sabin, M.L. and Koike, S. (2012). Population size estimation of men who have sex with men through the network scale-up method in Japan. *PLOS ONE* 7(1): e31184. doi:10.1371/journal.pone.0031184.

Feehan, D.M. (2016). *Surveybootstrap: Tools for the bootstrap with survey data. R package version 0.0.1. https://CRAN.R-project.org/package=surveybootstrap.*

Feehan, D.M. and Salganik, M.J. (2014). *The networkreporting package*. *http://cran.r-project.org/package=networkreporting*.

Feehan, D.M. and Salganik, M.J. (2016). Generalizing the network scale-up method: A new estimator for the size of hidden populations. *Sociological Methodology* 46(1): 153–186. doi:10.1177/0081175016665425.

Feehan, D.M., Umubyeyi, A., Mahy, M., Hladik, W., and Salganik, M.J. (2016). Quantity versus quality: A survey experiment to improve the network scale-up method. *American Journal of Epidemiology* 183(8): 747–757. doi:10.1093/aje/kwv287.

Guo, W., Bao, S., Lin, W., Wu, G., Zhang, W, Hladik, W., Abdul-Quader, A., Bulterys, M., Fuller, S., and Wang, L. (2013). Estimating the size of HIV key affected populations in Chongqing, China, using the network scale-up method. *PLOS ONE* 8(8): e71796. doi:10.1371/journal.pone.0071796.

Habecker, P., Dombrowski, K., and Khan, B. (2015). Improving the network scale-up estimator: Incorporating means of sums, recursive back estimation, and sampling weights. *PLOS ONE* 10(12): e0143406. doi:10.1371/journal.pone.0143406.

Henshaw, S.K., Singh, S., Oye-Adeniran, B.A., Adewole, I.F., Iwere, N., and Cuca, Y.P. (1998). The incidence of induced abortion in Nigeria. *International Family Planning Perspectives* 24(4): 156. doi:10.2307/2991973.

Jones, R. and Kost, K. (2007). Underreporting of induced and spontaneous abortion in the United States: An analysis of the 2002 National Survey of Family Growth. *Studies in Family Planning* 38(3): 187–197. doi:10.1111/j.1728-4465.2007.00130.x.

Kadushin, C., Killworth, P.D., Bernard, H.R., and Beveridge, A.A. (2006). Scale-up methods as applied to estimates of heroin use. *Journal of Drug Issues* 36(2): 417–440. doi:10.1177/002204260603600209.

Keogh, S.C., Kimaro, G., Muganyizi, P., Philbin, J., Kahwa, A., Ngadaya, E., and Bankole, A. (2015). Incidence of induced abortion and post-abortion care in Tanzania. *PLOS ONE* 10(9): e0133933. doi:10.1371/journal.pone.0133933.

Killworth, P.D., Johnsen, E.C., McCarty, C., Shelley, G.A., and Bernard, H.R. (1998). A social network approach to estimating seroprevalence in the United States. *Social Networks* 20(1): 23–50. doi:10.1016/S0378-8733(96)00305-X.

Killworth, P.D., McCarty, C., Bernard, H.R., Shelley, G.A., and Johnsen, E.C. (1998). Estimation of seroprevalence, rape, and homelessness in the United States using a social network approach. *Evaluation Review* 22(2): 289–308. doi:10.1177/0193841X9802200205.

Killworth, P.D., McCarty, C., Johnsen, E.C., Bernard, H., and Shelley, G.A. (2006). Investigating the variation of personal network size under unknown error conditions. *Sociological Methods and Research* 35(1): 84–112. doi:10.1177/0049124106289160.

Levandowski, B.A., Mhango, C., Kuchingale, E., Lunguzi, J., Katengeza, H., Gebreselassie, H., and Singh, S. (2013). The incidence of induced abortion in Malawi. *International Perspectives on Sexual and Reproductive Health* 39(2): 88–96. doi:10.1363/3908813.

Maghsoudi, A., Baneshi, M.R., Neydavoodi, M., and Haghdoost, A.A. (2014). Network scale-up correction factors for population size estimation of people who inject drugs and female sex workers in Iran. *PLOS ONE* 9(11). doi:10.1371/journal.pone.0110917.

Maltiel, R., Raftery, A.E., McCormick, T.H., and Baraff, A.J. (2015). Estimating population size using the network scale-up method. *The Annals of Applied Statistics* 9(3): 1247–1277. doi:10.1214/15-AOAS827.

McCarty, C., Killworth, P.D., Bernard, H.R., Johnsen, E.C., and Shelley, G.A. (2001). Comparing two methods for estimating network size. *Human Organization* 60(1): 28–39. doi:10.17730/humo.60.1.efx5t9gjtgmga73y.

McCormick, T.H., Salganik, M.J., and Zheng, T. (2010). How many people do you know? Efficiently estimating personal network size. *Journal of the American Statistical Association* 105(489): 59–70. doi:10.1198/jasa.2009.ap08518.

Mohamed, S.F., Izugbara, C., Moore, A.M., Mutua, M., Kimani-Murage, E.W., Ziraba, A.K., Bankole, A., Singh, S.D., and Egesa, C. (2015). The estimated incidence of induced abortion in Kenya: A cross-sectional study. *BMC Pregnancy and Childbirth* 15: 185. doi:10.1186/s12884-015-0621-1.

Moore, A.M., Gebrehiwot, Y., Fetters, T., Wado, Y.D., Bankole, A., Singh, S., Gebreselassie, H., and Getachew, Y. (2016). The estimated incidence of induced abortion in Ethiopia, 2014: Changes in the provision of services since 2008. *International Perspectives on Sexual and Reproductive Health* 42(3): 111–120. doi:10.1363/42e1816.

Moseson, H., Gerdts, C., Dehlendorf, C., Hiatt, R.A., and Vittinghoff, E. (2017). Multivariable regression analysis of list experiment data on abortion: Results from a large, randomly-selected population based study in Liberia. *Population Health Metrics* 15(1): 40. doi:10.1186/s12963-017-0157-x.

PMA2020 (2018). *Performance, monitoring, and accountability 2020 (PMA2020)*. Baltimore, MD: Johns Hopkins Bloomberg School of Public Health.

Polis, C.B., Mhango, C., Philbin, J., Chimwaza, W., Chipeta, E., and Msusa, A. (2017). Incidence of Induced Abortion in Malawi, 2015. *PLOS ONE* 12(4): e0173639. doi:10.1371/journal.pone.0173639.

Prada, E., Atuyambe, L.M., Blades, N.M., Bukenya, J.N., Orach, C.G., and Bankole, A. (2016). Incidence of induced abortion in Uganda, 2013: New estimates since 2003. *PLOS ONE* 11(11): e0165812. doi:10.1371/journal.pone.0165812.

R Core Team (2020). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Rastegari, A., Baneshi, M.R., Haji-maghsoudi, S., Nakhaee, N., Eslami, M., Malekafzali, H., and Haghdoost, A.A. (2014). Estimating the annual incidence of abortions in Iran applying a network scale-up approach. *Iranian Red Crescent Medical Journal* 16(10). doi:10.5812/ircmj.15765.

RBC/IHDPC (2012). *Estimating the size of populations through a household survey*. Calverton: Rwanda Biomedical Center/Institute of HIV/AIDS, Disease Prevention and Control Department (RBC/IHDPC), School of Public Health (SPH), UNAIDS, and ICF International.

Rossier, C. (2003). Estimating induced abortion rates: A review. *Studies in Family Planning* 34(2): 87–102. doi:10.1111/j.1728-4465.2003.00087.x.

Salganik, M.J., Fazito, D., Bertoni, N., Abdo, A.H., Mello, M.B., and Bastos, F.I. (2011). Assessing network scale-up estimates for groups most at risk of HIV/AIDS: Evidence from a multiple-method study of heavy drug users in Curitiba, Brazil. *American Journal of Epidemiology* 174(10): 1190–1196. doi:10.1093/aje/kwr246.

Salganik, M.J., Mello, M.B., Abdo, A.H., Bertoni, N., Fazito, D., and Bastos, F.I. (2011). The game of contacts: Estimating the social visibility of groups. *Social Networks* 33(1): 70–78. doi:10.1016/j.socnet.2010.10.006.

Sedgh, G. and Keogh, S.C. (2019). Novel approaches to estimating abortion incidence. *Reproductive Health* 16(1): 44. doi:10.1186/s12978-019-0702-0.

Sedgh, G., Rossier, C., Kaboré, I., Bankole, A., and Mikulich, M. (2011). Estimating abortion incidence in Burkina Faso using two methodologies. *Studies in Family Planning* 42(3): 147–154. doi:10.1111/j.1728-4465.2011.00275.x.

Shelley, G.A., Killworth, P.D., Bernard, H.R., McCarty, C., Johnsen, E.C., and Rice, R.E. (2006). Who knows your HIV Status II? Information propagation within social networks of seropositive people. *Human Organization* 65(4): 430–444. doi:10.17730/humo.65.4.08mwg9d3nfy8w9th.

Singh, S, Fetters, T., Gebreselassie, H., Abdella, A., Gebrehiwot, Y., Kumbi, S., and Audam, S. (2010). The estimated incidence of induced abortion in Ethiopia, 2008. *International Perspectives on Sexual and Reproductive Health* 36(1): 16–25. doi:10.1363/3601610.

Singh, S., Prada, E., and Juarez, F. (2010). The abortion incidence complications method: A quantitative technique. In: Singh, S., Remez, L., and Tartaglione, A. (eds.). *Methodologies for estimating abortion incidence and abortion-related morbidity: A review*. New York: Guttmacher Institute and International Union for the Scientific Study of Population: 71–98.

Singh, S., Prada, E., Mirembe, F., and Kiggundu, C. (2005). The incidence of induced abortion in Uganda. *International Family Planning Perspectives* 31(04): 183–191. doi:10.1363/3118305.

Sirbiladze, T., Tavzarashvili, L., Chikovani, I., Shengelia, N., and Sulaberidze, L. (2015). *Population size estimation of people who inject drugs in Georgia 2014*. Curatio International Foundation and Bemoni Public Union.

Sully, E.A., Madziyire, M.G., Riley, T., Moore, A.M., Crowell, M., Nyandoro, M.T., Madzima, B., and Chipato, T. (2018). Abortion in Zimbabwe: A national study of the incidence of induced abortion, unintended pregnancy and post-abortion care in 2016. *PLOS ONE* 13(10): e0205239. doi:10.1371/journal.pone.0205239.

Wang, J., Yang, Y., Zhao, W., Su, H., Zhao, Y., Chen, Y., Zhang, T., and Zhang, T. (2015). Application of network scale up method in the estimation of population size for men who have sex with men in Shanghai, China. *PLOS ONE* 10(11). doi:10.1371/journal.pone.0143118.

Zheng, T., Salganik, M.J., and Gelman, A. (2006). How many people do you know in prison? Using overdispersion in count data to estimate social structure in networks. *Journal of the American Statistical Association* 101(474): 409–423. doi:10.1198/016214505000001168.

Appendix A: Known population questions for Ethiopia and Uganda PMA Round 6

Uganda	**Ethiopia**
Tie Definition	
a) Individuals you know by sight AND name, and who also know you by sight and name. In other words, you should not consider famous people who you know about but who do not know about you; (b) Individuals you have had some contact with – either in person, over the phone, or on the computer -- in the past 12 months. These could be family members, friends, co-workers, neighbors, or other people you have contact with; and (c) Individuals 15-49 of age who currently live in Uganda.	(a) Individuals you know by sight AND name, and who also know you by sight and name. In other words, you should not consider famous people who you know about but who do not know about you; (b) Individuals you have had some contact with – either in person, over the phone, or on the computer – in the past 12 months. These could be family members, friends, co-workers, neighbors, or other people you have contact with; and (c) Individuals 15-49 years of age who currently live in Ethiopia.
Known Populations	
Women who have given birth in the last 12 months	Women who have given birth in the last 12 months
Women whose most recent birth was a multiple (twins, etc.)	Women whose most recent birth was a multiple (twins, etc.)
Women with at least one co-wife	Women who work as a domestic helper/cleaner
Women with education past senior six	Women with any education at a level higher than secondary school
Women who smoke a pipe or cigarette	Women who smoke a pipe or cigarette
Women who live in a household with a thatched roof	Women who work as a teacher
Women who live in a household that owns a car or truck	Women who live in a household that owns an animal-drawn cart
Women who live in a household that has a refrigerator	Women who live in a household that owns a motorcycle or scooter
Women who live in a household that owns an exotic cow	Women who live in a household that owns a bicycle
Women who live in a household that owns at least one sheep	Women who live in a household that owns at least one camel
Women who live in a household that has a landline	Women who live in a household that has a computer
Women who live in a household that has piped water inside the home	Women who live in a household that has piped water inside the home
Unknown Populations	
Women who have ever done something to intentionally end a pregnancy	Women who have ever done something to intentionally end a pregnancy
Women who ended a pregnancy in the past 12 months	Women who ended a pregnancy in the past 12 months
Validation Population	
Women who are currently using an IUD or implant	Women who are currently using an IUD or implant

Appendix B: Sensitivity analyses

We calculated two abortion incidence estimates in addition to our main analysis. First, we were concerned that providing a response of 0 to all 12 NSUM (known population) questions may be a reflection of poor comprehension of the method. In addition, respondents who give a 0 response to all or most of the NSUM questions may not be providing enough information about their networks, and we hypothesized that this could potentially bias the overall NSUM estimate. Therefore, our first sensitivity test restricts the NSUM estimate to respondents who provided non-zero responses to a minimum of two of the 12 NSUM questions.

In the second sensitivity test, we use the original analytic sample as well as a back-estimation process to identify problematic known populations that may be biasing the personal network size estimates. Previous work has suggested identifying and ultimately removing these items by creating a ratio that compares the back-estimate to the known population size for each NSUM indicator (Guo 2013). The closer this ratio is to 1, the more accurate the estimate. In order to identify problematic known populations, the current study employs a recursive approach developed by Habecker and colleagues; after the initial back-estimate known population ratios are calculated, the worst performing NSUM indicator is removed, and personal network size is re-estimated using the remaining known populations (Habecker, Dombrowski, and Khan 2015). This process is repeated recursively until all back-estimate/known population ratios are no less than .5 and no greater than 2. Only those known populations in the ratio range are used to estimate $\widehat{c_i}$, or the social network size, in the NSUM estimation equation.

Both abortion incidence estimates are then adjusted for transmission bias. For the approach restricting to respondents with at least two non-zero responses, the transmission bias correction was calculated only including respondents who self-reported abortions and who also met those criteria. To correct the estimate produced using known populations selected in the recursive back-estimation process, we use the same adjustment factor as was calculated for the main NSUM analysis, as both versions share the same exclusion criteria and therefore final sample.

We additionally provide prevalence estimates using the same exclusion criteria for the proportion of women currently using IUDs or implants. Uncertainty intervals were generated using the same bootstrap variance estimation procedure as in the main analysis. For the second sensitivity test, the 5,000 replicate estimates were generated assuming the same final set of known populations in each country as were selected by our step-wise selection process to calculate respondents' degrees.

Uganda

When we applied the criteria for the first sensitivity test, 184 respondents (12%) were excluded for not providing a minimum of two non-zero responses. This resulted in a new analytic sample of 1,898 women for this first sensitivity test. After employing the recursive back-estimation process to identify and remove problematic known population questions, we are left with nine known populations in Uganda. The excluded known populations are women who smoke cigarettes or pipes, women who live in households with piped water, and women whose most recent birth was a multiple birth. Supplemental Figure B shows the relationship between the NSUM population estimates and the DHS estimates for Uganda after the final results of the recursive back-estimation process, along with the results of the internal validity check for the main analysis and the sensitivity test restricting to respondents with at least two non-zero responses.

Ethiopia

In Ethiopia, a larger portion of respondents were excluded from the first sensitivity analysis because they did not provide a non-zero response to at least two NSUM question (26%, n = 913), resulting in a final analytic sample for this sensitivity test of 2,651 women. We performed the same recursive back-estimation process to identify and remove problematic known population questions. After this process, we are left with only five known populations in Ethiopia. The dropped known populations include teachers, women who live in households with piped water, women who live in households that own a computer, women who live in households that own an animal-drawn cart, women who live in households that own a motorcycle or scooter, women who live in households that own a bicycle, and women who live in households that own a camel. Supplemental Figure C shows the relationship between the NSUM population estimates and the DHS estimates for all three versions of the NSUM (main analysis and two sensitivity tests).

Results

The two estimation techniques employed in these sensitivity tests produced similar incidence estimates, both to each other as well as to the main NSUM estimate provided in this paper. (Supplemental Table C displays the unadjusted and transmission bias adjusted baseline, minimum two non-zero, and back-estimation process NSUM abortion incidence estimates and IUD/implant prevalence estimates for both Uganda and Ethiopia.) In Uganda, abortion incidence prior to adjusting for transmission bias ranged from 15.3 per 100 births to 17.0 per 100 births across the different estimation techniques. In Ethiopia, the estimates ranged from 3.5 per 100 births to 3.8 per 100 births.

Similarly, the IUD/implant prevalence estimates did not vary much across the three NSUM techniques. In Uganda, the unadjusted prevalence of IUD/implant use ranged 10.0–10.7%, and in Ethiopia 4.1–4.8%. These findings suggest that our analysis was not sensitive to our choice of inclusion criteria and selection of known populations (among the 12 originally included in data collection).