MONITORING HIV/AIDS DYNAMICS: AN APPLICATION OF REAL TIME ASSESSMENT METHODS FOR ESTIMATING PREVALENCE AT DISTRICT LEVELS IN UGANDA

Joseph Douglas Oguti Ouma

Supervisor:
Prof. Jonathan Levin

Co-supervisors:
Prof. Joseph J. Valadez and Dr. Caroline Jeffery

A thesis submitted by published work to the school of Public Health, Faculty of Health Sciences, University of the Witwatersrand (WITS), Johannesburg, South Africa in fulfillment of the requirements for the degree of Doctor of Philosophy, Public Health

Johannesburg, South Africa

February 2022

Declaration

I, Joseph D.O. Ouma, declare that this thesis is my own, original unaided work. Where there is any contribution from anybody or work by other people, it has been duly acknowledged. This work is being submitted for the degree of Doctor of Philosophy in Public Health in the University of the Witwatersrand, Johannesburg, South Africa. It has not been submitted before for any degree or examination at this or any other University.

Name: Joseph Douglas Oguti Ouma

Signature:  Date: February 2022

Joseph Ouma,

Student number: 1608215

Institution: School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Supervisors:

1. Prof. Jonathan Levin, Division of Epidemiology and Biostatistics, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa
2. Prof. Joseph J. Valadez, Department of International Public Health, Liverpool School of Tropical Medicine, Liverpool, United Kingdom
3. Dr. Caroline Jeffery, Ph.D., Department of International Public Health, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

Dedication

This work is dedicated to the memory of my late parents.

Publications

1. Joseph Ouma, Caroline Jeffery, Joseph J. Valadez, Rhoda K. Wanyenze, Jim Todd & Jonathan Levin (2020) Difference in HIV prevalence by testing venue: results from population level survey in Uganda, AIDS Care, DOI: 10.1080/09540121.2020.1734179

2. Ouma, J., Jeffery, C., Valadez, J.J. et al. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health 20, 379 (2020). https://doi.org/10.1186/s12889-020-8436-z

3. Ouma J, Jeffery C, Awor CA, Muruta A, Musinguzi J, Wanyenze RK, Biraro S, Levin J, Valadez JJ. Model-based small area estimation methods and precise district-level HIV prevalence estimates in Uganda. PLoS One. 2021 Aug 6;16(8):e0253375. doi: 10.1371/journal.pone.0253375. PMID: 34358233; PMCID: PMC8345831.

4. Paper 4. Factors associated with decline in HIV prevalence among adults aged 15-49 years in Uganda: A decomposition analysis of 2011 and 2016 population survey data. Under review

5. Case Study: Assessing HIV testing data from healthcare facilities in Uganda, Friday, September 11, 2020 case studies published by the Academy of African Sciences (AAS). https://www.aasciences.africa/news/assessing-hiv-testing-data-healthcare-facilities-uganda

This PhD study used data collected from Uganda AIDS Indicator Survey (UAIS) 2011, Uganda Population and HIV/AIDS Impact Assessment 2016 (UPHIA), Community Lot Quality Assurance Sampling (LQAS) survey conducted in 2016, Uganda Population and Housing Census, 2014 and the District Health Management Information System (DHIS2). The PhD candidate was the statistician for the community LQAS surveys in Uganda from 2009 to 2016, where he was responsible for survey sample selection, supervision of data collection, entry, management, analysis and preparation of district specific reports and the consolidated country report. The Country level report was a consolidation of data from all districts surveyed in a given year. For this thesis, the PhD candidate was responsible for data extraction, cleaning, analysis and wrote the first drafts of all manuscripts. All authors have agreed to these papers and manuscripts to be included as part of this PhD thesis.

Declaration: Student's contribution to articles and agreement of Co-authors

I, Joseph Ouma, student number 1608215 declare that this is my own work and that I contributed adequately to research findings published in the articles included in this thesis.

14th February, 2022

Signature of student date

Prof Jonathan Levin

Name of primary supervisor

14th February, 2022

Signature of primary supervisor date

Conference presentations

1. Poster presentation: “Monitoring HIV/AIDS dynamics: Application of real time assessment methods for estimating prevalence”.1st DELTAS Africa, Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) research conference, at the Hilton Hotel, Windhoek, Namibia. 1st - 3rd November 2017. Abstract number 10.
2. Oral presentation: “Estimating district level HIV burden through multilevel logistic regression modelling and a statistical annealing method using population survey and Health facility testing data in Uganda”. 2nd DELTAS Africa, Sub-Saharan Africa Consortium for Advanced Biostatistics SSACAB research conference, at the InterContinental Hotel, Nairobi Kenya, 10-14 September, 2018.
3. Oral presentation: “Combining national survey data with facility-based HIV testing data to obtain more accurate estimates of HIV prevalence in districts in Uganda”. 1st Joint conference of the Sub-Saharan Network (SUSAN) of the international Biometric Society (IBS) and DELTAS Africa, Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB). Biostatistics Research unit of the South African Medical Research Council (SAMREC), Cape Town, South Africa. 8th -11th, September 2019. Abstract number: 084.

List of abbreviations

AAS	African Academy of Sciences
AIDS	Acquired Immune Deficiency Syndrome
AIS	AIDS Indicator Survey
ANC	Antenatal Care
DELTAS	Developing Excellence in Leadership, Training and Science in Africa
DHIS2	District Health Information System Version 2
DHS	Demographic Health Survey
HMIS	Health Management Information System
HIV	Human Immunodeficiency Virus
HPE	Hybrid Prevalence Estimate
HREC	Human Research Ethics Committee
LQAS	Lot Quality Assurance Sampling
MOH	Ministry of Health
MOLG	Ministry of Local Government
NPHC	National Population and Housing Census
PEPFAR	Presidential Emergency Plan for AIDS Relief
PHIA	Population HIV Impact Assessment
PLHIV	People Living with HIV
PMTCT	Prevention of Mother to Child Transmission of HIV
SAE	Small Area Estimation
SSA	Sub-Saharan Africa
SSACAB	Sub-Saharan Africa Consortium for Advanced Biostatistics
TASO	The AIDS Support Organization
UAIS	Uganda AIDS Indicator Survey
UBOS	Uganda Bureau of Statistics
UNAIDS	The Joint United Nations Program on HIV/AIDS
UNCST	Uganda National Council for Science and Technology
UNICEF	The United Nations Children’s Fund
UPHIA	Uganda Population and HIV Impact Assessment
USAID	United States Agency for International Development
WHO	World Health Organization

Abstract

Background

More than 70% of the HIV cases and >75% of the HIV-related deaths globally occur in Sub-Saharan Africa. In Uganda, 1.5 million people were living with HIV in 2019 with approximately 80% of them being adults aged 15-49 years. The general population prevalence of HIV in Uganda was at 6.2% (7.6% for females and 4.7% among males; regional prevalence varies from 3.1% to 8.0%) in 2016/17. Although the annual number of new infections has dropped significantly since 1990 due to the available effective interventions, the number of new infections remain high. In 2019, there were 53,000 new infections, 21,000 deaths although 87% of the PLHIV were on treatment.

To curb the further spread of the HIV virus, information on population subgroups who are more likely to contribute to new infections is required at more decentralised levels such as districts. Existing databases such as the population surveys and the Health Management Information System (HMIS) do not provide reliable data at district levels. Statistical tools can however be utilised to ensure availability of more timely, precise and reliable district level information to guide planning and decision-making.

Aim of the study

The overall aim of this study was to identify and apply advanced statistical methodologies to estimate HIV prevalence and its predictors among adults aged 15-49 years at district levels in Uganda using facility HIV testing, population survey, community Lot Quality Assurance Sampling survey and Census data.

Methods

The data sources used to address the specific study objectives were: the nationally representative population-based surveys conducted in 2011 and 2016 among adults aged 15-64 years, who consented to interview and blood draw for HIV testing; HIV testing data collected during service delivery at health facilities, which is reported to the National District Health Information System version 2; Lot Quality Assurance Sampling survey data conducted in 2016 and the National Population and Housing Census conducted in 2014.

The study assessed comparability of health facility and population survey testing data. The 2011 Uganda AIDS Indicator Survey (UAIS) data was stratified according to venue of the most recent HIV test, i.e. tested in a health facility or tested in a community setting. HIV prevalence ratio and the 95% confidence interval was computed using the KATZ methodology.

The Hybrid Prevalence Estimation (HPE) methodology was applied to obtain HIV prevalence estimates for districts in Uganda, through the following steps: (I) a multilevel logistic regression model was applied to the 2011 UAIS data to obtain the propensity of testing for HIV in a health facility. (II) Denominators for the health facility data were adjusted based on the population survey design. (III) District level average propensity to test for HIV in a health facility was then used to combine the adjusted health facility data and the population survey data of non-facility testers to obtain district HIV prevalence estimates for the 111 districts in Uganda. Percentage change in standard errors for the direct survey and HPEs was computed and assessed for improvement in efficiency of the HPEs. Consistency/agreement between HPEs and the direct survey estimates was assessed using the Bland Altman analysis.

Small Area Estimation techniques were applied and assessed for robustness in estimating district level HIV prevalence estimates. Four different approaches were applied namely (i) the direct survey method, where district HIV prevalence estimate was computed based on only sampled individuals from the district; (ii) the area level, Fay-Herriot (FH) model, applied to the 2016 population survey data with health facility antenatal data as the auxiliary variable. The methodology is a two-stage hierarchical model that links the outcome with possible predictors defined at area level; (iii) the Battese, Harter and Fuller (BHF) unit level model, was applied to the 2016 population survey data using the LQAS survey data as the source of auxiliary information. The BHF methodology uses auxiliary information at individual/unit level; and lastly the Spatial Fay-Herriot (SFH) model was applied, assuming a spatial correlation between districts.

The mean squared errors for BHF estimate was computed as a weighted combination of variance among the sample observations and variance of the estimates for the un-sampled observations. Similarity and consistency of the direct survey, FH, SFH and BHF model estimates were assessed using regression methods. Percentage improvement in the standard errors were also computed.

Finally, participants' distribution by the socio-demographic characteristics was analysed separately for the UAIS 2011 and UPHIA 2016 surveys. Descriptive analysis was also carried out to assess change in HIV prevalence between the two surveys. Pooled logistic regression was used to assess factors associated with HIV positivity while multivariable logistic decomposition model was used to determine the contribution of characteristics and coefficients of the characteristics to change in HIV prevalence between the two surveys. The Blinder-Oaxaca decomposition method was used to display the contribution of each characteristics or coefficient of the characteristics to change in HIV prevalence

between the two surveys. All analyses were weighted using the population survey weights and analysis carried out using R version 3.5 and Stata Release 15.1.

Results

Generally, testing in a health facility was associated with a higher HIV positivity compared to testing in a community setting (Prevalence Ratio=1.8, 95% CI: 1.4-2.2). There was a two-fold or higher increase in HIV positivity among health facility testers compared to non-facility testers if they were male; aged 15-19 years, 30-39 years and 40-49 years; had no formal education; never married or were previously married; reported no sexual partner in the 12 months preceding the survey; not in formal employment and resident in urban areas.

The propensity to test in a health facility was generally higher among females across all regions of the country. It was also higher in Mid-Northern, North East and Kampala regions of the country. Of the 111 districts included in the analysis, 105 (95.5%) had narrower confidence intervals compared to confidence intervals from direct survey-based estimates. Overall, the HPE standard errors decreased by 28.9% (95% CI: 23.4-34.4) compared to survey-based Standard Errors (SE).

There was no overall difference between survey and HP estimates. Average difference for males was -0.01 units (95% CI: -0.05, 0.03) while for females was 0.00 units (95% CI: -0.06, 0.06). The mean difference between the HPE and DHIS2 HIV prevalence estimates was 0.01 units (95% CI: -0.05, 0.06). Average difference for males was -0.01 units (95% CI: -0.07, 0.06) while for females was 0.02 units (95% CI: -0.05, 0.09).

Although FH, SFH and BHF models and the direct survey estimates were similar, direct survey estimate (0.065) had a larger average SE of 0.055 compared to the FH estimate

(0.054 (SE=0.020)), SFH estimate (0.054 (SE=0.021)), and BHF estimate (0.058 (SE=0.026)). Regression results showed that the BHF model estimates were strongly correlated with the direct survey estimates ($\beta_1 = 0.73$, r^2=0.891) compared to FH model against direct survey estimates ($\beta_1 = 0.27$, r^2=0.477).

Model and survey based estimates were similar for large survey sample sizes in the districts but significantly different for districts with small survey sample sizes. For example Nwoya district with survey sample of 44 individuals, had an HIV prevalence estimate of 0.13 (SE=0.68); 0.062 (0.014) and 0.083 (0.021) while Mbale district with a survey sample of 874 individuals had estimates of 0.053 (0.007); 0.053 (0.007); and 0.052 (0.004) for direct survey, FH and BHF models respectively. Overall average improvement in the precision of the estimates, were 37.5% and 33.1% for the BHF and FH model estimates respectively compared to the direct survey estimates.

Coefficient of variation of the direct survey estimates was generally larger compared to CV of FH and BHF model. If estimates with CV less than 20% were considered suitable for decision making, then only 21 (18%) of the districts would have reliable information for decision making based on direct survey estimates while 53 (47.3%) and 35 (50%) of the districts would have reliable information for decision making based on the FH and BHF models respectively.

HIV prevalence dropped by 1.3 percentage points between 2011 and 2016. A greater reduction in HIV prevalence was observed among Males (1.8); individuals aged 20-24 years (2.0); individuals with primary level of education (1.4); individuals with two or more sexual partners in the 12 months preceding the survey (2.7); and individuals who were divorced/separated (3.5). An increase in HIV prevalence between the 2011 and 2016

survey was observed among individuals with no formal education (0.3), individuals with no sexual partners in the 12 months preceding the survey (0.3) and among individuals aged 40-49 years (1.6). Change in the composition of survey participants contributed to 54% increase in HIV prevalence between 2011 and 2016 time points, while the coefficients contributed to 154% decline in HIV positivity between the two survey time points. While the coefficients contributed to a larger reduction in HIV prevalence, the contribution of most of the variables was not statistically significant except for the age group 40-49years.

Conclusion

Information for decision making for districts in Uganda was obtained. The findings show that it's feasible to use existing population survey and routine facility HIV testing data either combined through the HPE methodology or as auxiliary variables in the FH area level SAE model to obtain more precise indicator estimates for district level planning and decision making. Availability of individual level data such as those obtained using LQAS surveys leads to greater increase in the precision and reliability of the estimates.

Higher risk of infection among individuals accessing health facilities especially male, age groups 15-19 years, never married and those who reported no sexual partners in the 12 months preceding the survey, implies screening protocols should focus on these risk categories to minimise missed opportunities of identifying those who are likely to be HIV positive.

These findings also imply that to use the datasets in complementarity, there is need to take into account HIV prevalence differences by socio-demographic characteristics. Additionally, the compositions of the population subgroups with higher risks of infection

also contribute significantly to change in HIV prevalence. Intervention efforts should therefore focus on such subgroups to increase the impact on HIV prevention.

The HPE methodology presents dual benefits: (I) improves precision of the estimates and (II) enables computation of standard errors for routine health facility data whose denominator is difficult to ascertain directly.

Further research should consider incorporating more HIV risk factors in the models, using subgroup HIV prevalence differences as prior information in the models and applying mechanisms to deal with missing data while applying the methods.

Acknowledgements

I thank family and friends who supported me throughout the PhD; it was at the most difficult time. You individually played a role to help me move in the direction I did and moving forward was an inevitable direction by God's grace. We have all achieved and will always be thankful to God for his providence.

My promoters and PhD advisors were individually needed for the PhD journey to take shape and indeed to complete. Prof Jonathan Levin for the vast experience and kind guidance. We were able to walk through each paper and always reminded me on how each component/paper fits within the PhD context. You always "whipped" me back into the line". I vividly remember this question, "What did we plan to do?" Dr. Caroline Jeffrey who carefully read through the manuscripts from draft zero and kept reminding me of the possible questions the reviewers might ask, very candid indeed. Professor Jim Todd, your quick turn around and deep questions were always needed. Prof Joseph Valadez introduced me to this area fitting well with my passion of serving the most needy communities better and finally Professor, Rhoda Wanyenze, you dived right into my work almost half way the PhD journey but it was as if you were part of the conceptualization. Your experience with Uganda's HIV/AIDS landscape was indeed needed for me to contextualise whatever I was doing.

I would like to thank my sponsor, the Sub-Saharan African Consortium of Advanced Biostatistics (SSACAB) under the leadership of Professor Tobias Chirwa for offering me the opportunity to carry on with my dreams. The program administrators Dr Pascalia Munyewende and Gloria Bowes, for their unwavering support during the study period.

I equally thank the WITs school of health sciences for offering me the opportunity to study at the prestigious school. Not all who apply to your programs are admitted. I am privileged

to have attained the PhD at your prestigious school and of course the University of the Witwatersrand, Johannesburg, South Africa. PhD administrator Paul and coordinator Dr Jude Igumbor for the encouragement, I vividly remember our lunch time conversation during a PhD retreat at Chumba safari lodge which always re-ignited my energy to learn new things, the Idea that we all students was very inspiring indeed. My PhD colleagues, Christopher, Zvifadzo, Lawrence, Noel and Jonathan, thank you for the encouragements.

Ministry of Health, Uganda, ICF Macro international, Centres for Disease Control and prevention and ICAP, thank you for the permission to use data from your programs, that you dedicatedly collected and/or financed.

Table of Contents

Content **Page**

List of tables

List of figures

1.0 CHAPTER ONE: INTRODUCTION

1.1 Problem statement

Traditionally, population surveys are used to obtain reliable indicator estimates at national and regional levels or for large demographic subgroups [1–4]. For example, the Uganda Demographic and Health Survey (UDHS) is used to obtain reliable estimates of institutional deliveries or skilled assistance of deliveries at national and in the geographical regions of the country [5,6]. The Uganda AIDs Indicator Survey (UAIS) similarly, is designed to provide reliable HIV/AIDS indicator estimates including HIV prevalence at national and regional levels of the country [7–9].

Although these surveys have continually expanded to improve the availability of more decentralized data for planning and decision-making purposes, through increased sample sizes and stratification of the country into smaller geographical areas comprising more homogeneous population groups, they have not been powered well enough to provide reliable estimates for the third administrative level (districts). This implies that attempts to obtain estimates for districts based on these surveys will generate estimates with large standard errors. Under-sampling of high risk groups such as Commercial Sex Workers (CSW), non-response and other survey related sampling errors further impact representativeness of these surveys at district level. Additionally, the 2011 UAIS found up-to a 5-fold difference in HIV prevalence estimates between regions [7]. Estimates at regional levels are also generalised and fail to capture district level variation in HIV transmission patterns and service coverage [10]. The significant variation across regions and over aggregation of estimates necessitates more refined and accurate estimates to enable strategic resource allocation at district levels.

HIV testing data from antenatal attendance and other health care facilities are often used as an alternate source of data to monitor HIV prevalence. These however comprise data from only individuals who access health facilities, and from public health facilities, excluding private health facilities. Additionally health facilities tend to be located in urban and accessible areas [1,11,12]. This renders HMIS data biased and cannot therefore be used to describe the general population distribution of HIV.

Targeted district level surveys cover specific subpopulation groups and are often conducted when the supporting agency/organisation needs to address a specific research/health problem [13,14]. Information generated from these surveys is therefore limited to the specific population group and cannot be used for general population inference and for comparison across districts.

Lack of accurate and reliable information at the district level thus undermines HIV/AIDS control efforts despite the available effective HIV/AIDS interventions. Modelling approaches embedded in statistical software packages such as THEMBISA [15] and SPECTRUM [16] provide prevalence estimates taking into account the bias and limitations associated with using population survey and facility antenatal testing data. While the models are continuously updated based on new data and emerging knowledge of the epidemic [17,18], they have their own limitations and assumptions including the need for regular training, input requirements such as HIV prevalence rates, number of patients on antiretroviral drugs (ART), fertility and mortality rates, which information are often not often available or varying significantly based on the data source.

Geo-spatial statistical models for estimating HIV prevalence, apply simple Kriging, Kernel density [19] and Bayesian geo-statistical [20,21] approaches. The models assume a continuous prevalence surface; some models do not incorporate error terms while others

use data from population surveys only. Use of population survey data only, implies that the precision of the parameter estimates remain constrained by the available survey sample size in the district.

Simple and easy to use statistical approaches are therefore necessary to provide timely information for planning and decision-making at district levels. Small Area Estimation (SAE) methods using antenatal HIV prevalence as auxiliary/predictor variable has been applied to obtain HIV prevalence estimates for districts in South Africa [22]. The application found more precise HIV prevalence estimates for districts in South Africa compared to the direct prevalence estimates from the population survey.

SAE methods use explicit linking models with random-area specific effects, accounting for between area variation, to obtain parameter estimates based on data from surveyed and non-surveyed individuals/areas. Supplemental data is incorporated as predictors in the regression models and are selected through variable selection approaches or based on prior information on their relationship with the target outcome.

In this study, we test and apply the statistical annealing approach [23], the Hybrid Prevalence Estimation (HPE) and SAE methodologies, while addressing the limitations of using the data sources independently, challenges with small survey sample size at the district level and bias associated with using health facility data for surveillance.

To apply the statistical approaches, an understanding of the comparability of the datasets was explored. Many scholars have found HIV prevalence from health facility/antenatal testing data to underestimate general population prevalence [24,25] while in other studies, they were found to be similar [12,26]. The population based survey data contain information on HIV testing venues which can be analysed to understand the comparability

of prevalence estimates for individuals tested in a health facility and those tested in a community setting.

Furthermore, paucity of precise district level HIV estimates implies that comparison across districts will be flawed. We assessed factors associated with HIV positivity and the contribution of these factors to change in HIV prevalence over time using the Blinder-Oaxaca decomposition logistic regression model to inform possible specific interventions that could be rolled out in identified high districts. The decomposition model identifies the factors that have a higher impact on changes in HIV prevalence that could be targeted to improve the impact of interventions [27,28]..

1.2 Background to the Problem

The HIV/AIDS epidemic continues to cause significant morbidity and mortality worldwide. Globally, more than 38 million people were living with HIV in 2019 with approximately 19% of them not aware of their status [29]. There have been over 35 million deaths since the start of the epidemic. Sub-Saharan Africa remains the most affected region of the world. More than 70% of the cases living with HIV and more than 75% of the deaths occur in the region [29].

In Uganda, the number of people living with HIV increased from 850,000 in 1990 to 1.5 million in 2019 with approximately 80% of People Living with HIV (PLHIV) being the family breadwinners (active/working adults aged 15-49 years) [29]. The annual number of deaths and new infections have however dropped significantly since 1990 due to the available effective interventions. New infections dropped from 110,000 in 1990 to 53,000 in 2019 while total annual deaths dropped from 84,000 in 2000 to 21,000 in 2019 [29]. In 2019, 85% of the adults living with HIV were on Anti-retroviral Treatment (ART) due to the increased efforts scaling-up treatment [29].

HIV distribution in Uganda is mainly generalised, i.e. the majority of new infections occurring in the general population and infection is not restricted to specific high-risk population sub-groups. The population level prevalence of HIV in Uganda is at 6.2% and is disproportionally distributed across population subgroups and regions of the country. Prevalence is 7.6% for females and 4.7% among males; regional prevalence varies from 3.1% to 8.0%; while population subgroups at higher risk of new infections include commercial sex workers, young girls and adolescent women, men who have sex with men, people who inject drugs and the fishing communities [8].

UNAIDS set ambitious goals/targets to fast track an end to the epidemic. They anticipated an end to the HIV epidemic by 2030 if by 2020, 90 percent of the people living with HIV knew their status, 90 percent of those who knew their status are on treatment and 90 percent of those on treatment achieved viral suppression. By 2019, 84% of the PLHIV in Uganda knew their status, 87% of those who knew their status were on treatment while 88% of those on treatment, were virally suppressed. The country still lags behind in achieving the set target and is unlikely to achieve them by 2020. A key effort to get all HIV infected people on treatment is the Test and Treat approach that was recommended by the WHO and rolled out in the country in 2013 [30,31]. This policy has enabled a significant majority of HIV infected individuals to initiate ART upon testing HIV positive. Other interventions implemented over time to curb the spread of the epidemic include medical male circumcision and the ABC (Abstinence from sexual intercourse until marriage; Being faithful to one sexual partner; and Correct and consistent condom use) strategy.

To achieve this however, requires information to enable targeting/re-allocation of the resources according to the HIV/AIDS situation in a given area. Population subgroups who are more likely to contribute to new infections require differentiated packaging and intensity of interventions to accelerate movement towards the targets/goals to end the HIV epidemic. Targeting interventions has the following benefits:

(i) The smaller the area the greater the effectiveness of policy interventions;

(ii) Disease concentration varies across settings, making it inefficient to implement same set of interventions across settings;

(iii) PEPFAR, the largest donor for HIV/AIDS interventions implements a prioritized strategy whereby there is intensified support in districts classified as scale-up

saturated and scale-up aggressive. These are classified that way due to the gaps in HIV testing and ART uptake;

(iv) Uganda implements a decentralized system of governance whereby district level authorities lead planning, decision-making and intervention implementation.

District level planning and decision-making is however constrained as existing databases do not provide reliable data at district levels. Statistical tools can be utilised to ensure that districts have more timely, precise and reliable information to guide planning.

In this study we test and apply the HPE and SAE techniques to obtain more precise estimates of HIV prevalence for districts in Uganda. Annealing techniques such as the HPE methodology, combine data from routine service delivery and a small survey sample to obtain estimates that are more precise [23]. Health facility HIV testing data was combined with population survey data to obtain HIV prevalence estimates for districts.

Additionally, district level prevalence estimates were obtained using random effects SAE models. SAE models capture both group and individual level effects [32,33] and are therefore very useful in HIV modelling where the factors may relate to an individual or the structural setting. The models were applied based on availability of auxiliary information as described in the methods section of this thesis, i.e. unit level model where unit/individual level auxiliary information are available and area level model where area/district level information are available. The models were also compared for efficiency. SAE methods borrow information from other areas/locations through linking models to obtain more precise indicator estimates.

The study further assessed comparability of HIV prevalence estimates for individuals whose most recent HIV test was in a health facility and those tested in a community setting [34]. Antenatal HIV testing data complement estimates from population surveys to provide more accurate information for planning. A study conducted in five countries in Sub-Saharan Africa however found that the antenatal HIV prevalence estimate was higher than the population survey estimate in four out of the five countries assessed [12]. Differences in HIV prevalence estimates from antenatal versus population survey data have also been observed across age groups and places of residence [1].

1.3 Aims and Objectives of the Study

1.3.1 Aim

The overall aim of this study was to identify and apply advanced statistical methodologies to estimate HIV prevalence and its predictors among adults aged 15-49 years at district levels in Uganda using facility HIV testing, population survey, community Lot Quality Assurance Sampling (LQAS) and census data.

1.3.2 Objectives

1. To apply and assess robustness of the Hybrid Prevalence Estimation methodology in estimating district level adult HIV prevalence using 2011 population survey and facility HIV testing data in Uganda and assess any potential biases associated with application of this methodology.
2. To apply and compare Small Area Estimation techniques in estimating HIV prevalence for districts in Uganda.
3. To identify factors for inter and intra-district variations of adult HIV prevalence in Uganda.

4. To assess whether the change in adult HIV prevalence represents change in the factors associated with HIV prevalence between 2011 and 2016.

1.4 Justification/Significance of the study

Uganda was the world leader in boldly facing and tackling the HIV epidemic. In the early phase of the epidemic, early 1990s, the country realized a drastic decline in HIV prevalence of 70%, a decline that has not been observed elsewhere in the world [35,36]. Key pillars to this reduction were messages on positive behavior change such as reduction in sexual partners and increased condom use; high-level political involvement; decentralization/local empowerment of communities and individuals to take action; voluntary counselling and testing and involvement of stakeholders including religious leaders [36–38].

At the turn of the millennium, data from sentinel surveillance sites and targeted population studies across the country found prevalence was no longer falling and was beginning to rise in some sections of the population [15,39], prompting the need for more information to tackle the epidemic.

While the country has relied on antenatal sentinel surveillance data to monitor trends in the epidemic up to this point, results from the general population sero-surveys showed that, the estimates from sentinel sites overestimated prevalence in the general population [40]. This further prompted a search for statistical adjustment methods to generate estimates that accurately measure the level of the epidemic in the general population [41].

Modelling approaches involve the use of available data to obtain indicator estimate, additional data are incorporated in the models at no additional costs for data collection. Some of these methods have not yet been applied in the Ugandan setting, and it is

important to note that modelling approaches involve specific assumptions. For example, the SPECTRUM software package generates estimates at national and regional levels. SPECTRUM spreadsheets and SAE techniques introduced by the CDC are yet to be rolled out in Uganda [42]. Geo-statistical models such as PrevR and fully Bayesian models do not incorporate uncertainty of the estimates and only use data from the population surveys respectively limiting their application.

This study therefore implements the statistical annealing technique, Hybrid Prevalence Estimation, and Small Area Estimation methods using readily available antenatal/health facility HIV testing data, census data and community Lot Quality Assurance Sampling survey data as supplemental information. The HPE methodology anneals a convenience sample with a small population-based random sample to improve the precision of estimates. The methodology minimizes the biases associated with relying on estimates from a convenience sample only. A large database for population level LQAS surveys has also been put together, pooling data, rich with important risk factor information from the annual surveys in Uganda. This data can be used to generate district level information for planning and decision-making at no additional costs.

Additionally, generating indicator estimates at district levels will reveal district specific patterns, the distribution of the epidemic and factors associated with district level variation. These will facilitate identification of highly affected districts and inform implementation of well-tailored HIV/AIDS interventions in those districts.

The Ugandan government also operates a decentralized system of governance where local governments or districts are mandated to implement and monitor services on behalf of the central government. Development partners such as the US Centers for Disease Control and Prevention (CDC) also support service delivery at the district level. Availability

of accurate information at the district level will therefore augment planning and decision-making.

1.5 Conceptual framework.

The conceptual framework in figure 1 presents the approaches used in this study to obtain HIV prevalence estimates for districts in Uganda. It shows how the different methods and data sources were linked/pooled together to obtain the district level HIV prevalence estimates. The framework was adapted from the SPECTRUM/AIDS impact model for estimation of key HIV/AIDS impact indicators from multiple data sources [43,44].

Objective one: Health facility HIV testing data contains basic demographic characteristics (age and gender) and HIV status of individuals tested at health facilities. Using the HPE methodology, this data was combined/annealed with the UAIS (2011) data to obtain HIV prevalence estimates for districts [23]. Weighted probability of testing for HIV in a health facility was generated from the population survey data and used to combine the Health facility HIV testing data and the population survey data. The denominators of the health facility HIV testing data was also adjusted as described in the methods section.

Objective two: This objective focused on using variables from population census, health facility and community LQAS data sets as predictors in the random effects SAE models. Small area models applied include the Area level and Unit level models. Results from the two approaches were compared with direct survey estimates for efficiency of the standard errors of the estimates. Variables were assessed using regression techniques and those that were significant at 5% level of significance were included in the final model.

Objective three and four: These objectives used 2011 and 2016 population survey data. In objective three, HIV prevalence among individuals tested in health facility and those

tested in a community setting was compared while in objective four, change in HIV prevalence between 2011 and 2016, the contribution of the HIV risk factors to the change in HIV prevalence between the two time points was assessed.

Conceptual framework

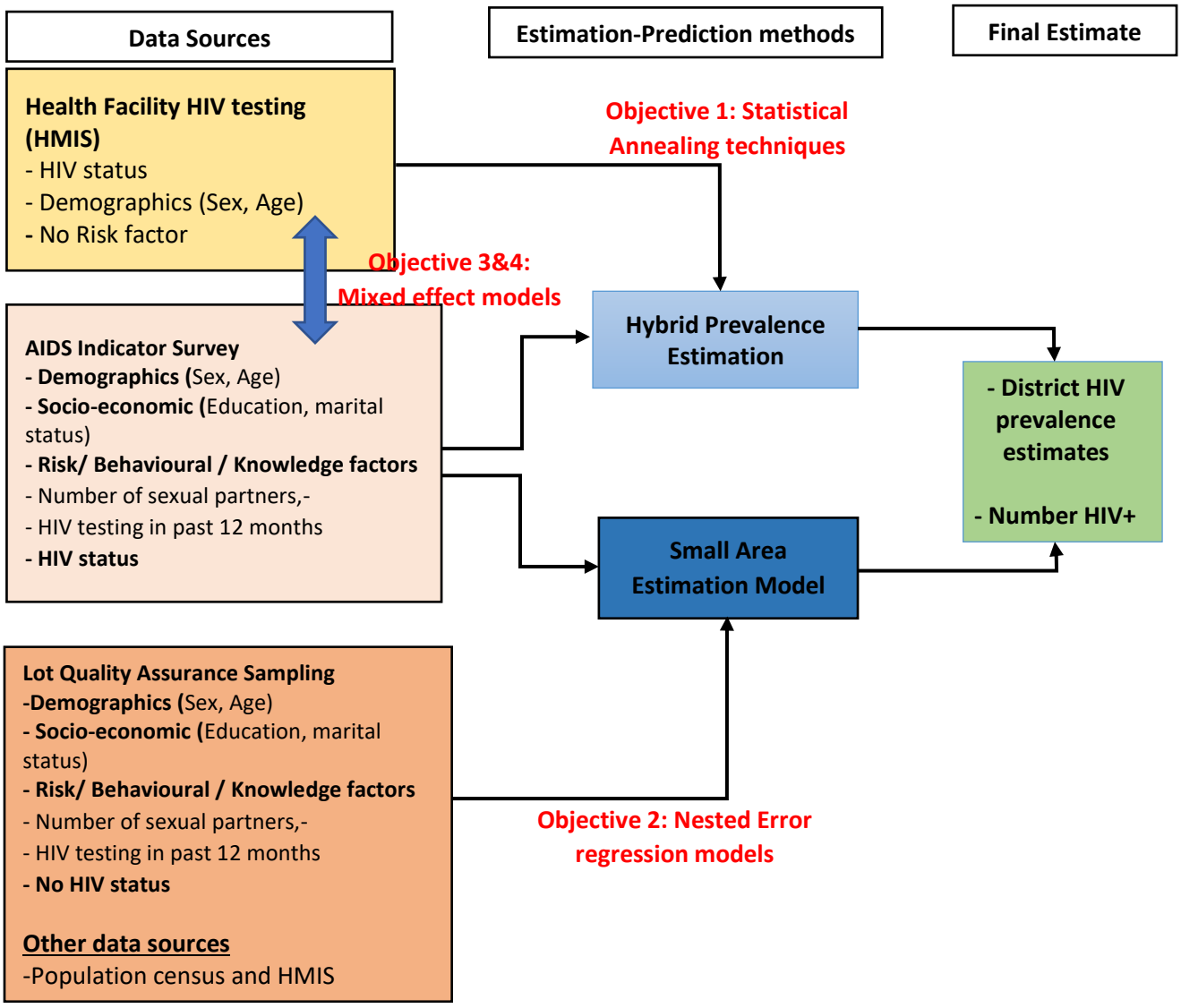


Figure 1: A framework for estimation of HIV prevalence at district levels

Notes: The Hybrid Prevalence Estimation (HPE) and Small Area Estimation (SAE) methods incorporate auxiliary/supplemental information through different approaches. While the HPE methodology combines data from two sources linearly using a weighting

factor that is associated with the outcome of interest, the SAE methods implemented in this research incorporate auxiliary information as predictors at unit and/or target area level through regression modelling. Additionally, SAE models incorporate area-specific variability as part of parameter estimation and a variety of models can be implemented and assessed/validated for accuracy and performance depending on the nature of the outcome.

2.0 CHAPTER TWO: LITERATURE REVIEW

This chapter presents an overview of the literature on methods for obtaining HIV prevalence estimates for the third or lower administrative areas (i.e. districts or sub-district levels); the data sources used and the underlying assumptions. The burden of HIV in Uganda, sources of data and summary of the approaches for estimating prevalence including shortcomings were described in the introduction.

2.1 Introduction

Measuring HIV distribution in a given setting is key to understanding the level of the disease and the type of interventions applicable to curb further spread. HIV prevalence is one of the key indicators used to assess the distribution and impact of the HIV/AIDS epidemic in a given area or subgroups of the population. It is defined as the percentage of people tested in each group who were found to be infected with HIV [45]. It is thus based on a numerator and specific denominator to which the measurement is inferred. Various data sources and methods are used to obtain HIV prevalence estimates each with inherent advantages and limitations.

2.2 Data sources for HIV prevalence estimation

In most of sub-Saharan African, the distribution of HIV in the population is described as generalized [46]. That is, HIV infection is firmly established in the general population with the majority of new infections occurring among individuals in stable sexual relationships/married [47,48]. New infections are also not restricted to a specific high-risk population subgroup.

2.2.1 Use HMIS data for prevalence estimation

HIV testing data from Prevention of Mother to Child Transmission of HIV (PMTCT) programs, and HIV testing data collected among women accessing antennal services are

often used to obtain HIV prevalence estimates [26,49–52]. More recently, WHO/UNAIDS have urged Countries to include HIV testing data for individuals tested as part of health care service at health facilities, to improve HIV prevalence estimates [53]. These sources include only individuals seeking care or maternal and child health care services at health facilities. Seeking care at health facilities is however, known to be associated with several factors ranging from individual, systems to policy level factors [54] some of which influence the risk of HIV infection. For example, shorter distance to the health facility; recent ill health; seeking antenatal, labour and delivery services or care for a sick family member are some of the characteristics of individuals who visit heath facilities and therefore get tested for HIV [54–56]. While district authorities in Uganda and many other settings where facility based HIV testing data is routinely collected, use the data to inform service delivery decisions, the HIV prevalence estimates from this dataset is biased for describing the general population distribution of the epidemic. Many researchers use adjustment methods including combining facility (antenatal testing) based data with other data sources to generate more accurate HIV prevalence estimates [2,57].

2.2.2 Population based random surveys

Population based random surveys with HIV testing component, currently established in most Sub-Saharan Africa (SSA) countries, are a more reliable source of data to describe general population distribution of the HIV epidemic. The surveys are standardized and comparable across time and geographical areas including between countries [46,58,59]. The surveys are however resource intensive, implying they can only be conducted at longer time intervals (e.g. five yearly), when adequate resources are available or have been mobilised [18,60,61]. The surveys are designed to provide accurate estimates at national (first administrative) and regional (second administrative) levels. HIV prevalence

estimates for lower administrative or geographical areas, such as districts (third administrative level), based on survey data are associated with large standard errors, constraining their use for planning and decision-making at the third administrative level. Additionally, the absence of data and/or resources to collect data more frequently implies lack of HIV prevalence information to guide intervention implementation in the period when they are not conducted.

2.2.3 Other sources of data

For settings with low level (concentrated) epidemics [46], sources of data used to obtain HIV prevalence estimates include special studies/surveys targeting specific or key population subgroups who are at greater risk of HIV infection e.g. surveys among People Who Inject Drugs (PWID), Men who have Sex with Men (MSM) or Commercial Sex Workers (CSW). Prevalence estimates from these special surveys cannot be used to describe HIV distribution patterns in the general population as it leads to inaccurate deductions.

2.3 Methods for HIV prevalence estimation

Models that use a single or multiple data sources to predict/establish HIV prevalence estimates for lower administrative levels have been proposed in literature to overcome challenges associated with small sample sizes from population surveys and selection bias associated with health facility data. The models are either imbedded in specific software tools or implemented dynamically to obtain HIV prevalence estimates

2.3.1 Customized Software packages

Modelling approaches imbedded in specific software tools, require specific input variables to obtain HIV prevalence estimates. These extrapolation-modelling approaches have

evolved and become more sophisticated following new knowledge on the nature of the epidemic [62]. The approaches use data including ANC HIV, mortality, and fertility data to estimate the size of the HIV burden at both national and subnational levels.

Their use therefore depends on the availability of the required data. Examples of model-based software packages for the estimation of HIV prevalence include Estimation and Projection Package component of SPECTRUM [16,17] and THEMBISA. SPECTRUM, developed by a consortium of epidemiologists, statisticians, program managers, Country AIDS Control Programmes, incorporates geospatial models as part of EPP package to provide estimates of HIV impact indicators at lower administrative levels [18,63,64].

For SPECTRUM estimation, regional workshops are conducted two yearly, where country teams are trained on the specific methods and tools to produce estimates. The data sources required as inputs to the software include prevalence among women attending antenatal clinics; population based survey data; sentinel surveillance among populations at high risk of infection; Anti-retroviral (ART) uptake; population demographic and HIV incidence data [17,18,64–66]. Estimates from the software tool are however continually updated based on availability of more data and information on HIV distribution in the country.

Thus, every year, new estimates are produced for the current and previous years. Estimates generated for the current year are not however, directly comparable with estimates produced in the previous years. This presents a complexity with the methodology, as estimates generated every year have specific assumptions, are dependent on available data and users have to undergo annual training to use the system efficiently.

2.3.2 Spatial methods

These include geospatial models such as Ordinary/Simple Kriging, Kernel density and Bayesian model based geo-statistics [11,19,21,67]. The geospatial models assume continuous prevalence surfaces and have been applied to predict HIV prevalence in relation to spatial and behavioural determinants in Kenya [20], to predict the spatial distribution of HIV in South Africa [67] and to map change in the spatial distribution of HIV at provincial levels in Zambia [68]. District level prevalence estimates for all districts in Uganda and other countries in SSA have also been generated using Kernel density approach [19,64,69].

The major draw-back of the geo-statistical approaches is uncertainty of estimates in most districts due to small survey sample sizes at the district level. Other challenges include displacement of data points that may affect prevalence estimates for a given district since points near the district boundary may be displaced to another district, use of data layers (climatic/geographical/economic) that are often not readily available and failure by some of the methods to calculate/compute error (uncertainty) around the point estimate.

2.3.3 Model based methods

Statistical approaches require knowledge of the risk factors for HIV infection in the modelling process. Statistical approaches ignore the assumption of spatial dependency, compared to spatial approaches, which is often not achievable in a real setting. The approaches make specific allowance for between area variation beyond that explained by x-variables included in the models. For example, Kondlo and Manda (2011) applied an HIV predictive model to obtain HIV prevalence among adults 15-49 years for local areas/municipalities in South Africa using DHS and Census data [70]. They applied cross validation techniques to obtain an optimum model. Manda et al applied a multi-stage

approach (i.e. Used Inverse probability of testing to predict HIV status for individuals who were not tested during the survey and applied Bayesian models) to generate HIV prevalence for districts in Malawi using population survey and antenatal HIV testing [71].

Small Area Estimation methodology

Statistical Small Area Estimation (SAE) approaches are gaining prominence as tools to obtain parameter estimates for areas with small/reduced survey sample sizes. The approach borrows information from neighbouring areas through linking models to generate parameter estimates. For example, Gutreuter et al. (2019) applied statistical SAE to obtain HIV prevalence for districts in South Africa using antenatal data as the explanatory/predictor variable [22]. The models are generally classified into Area level models and unit level models. The most commonly applied area level model is the Fay-Herriot (FH) model whereas the most commonly applied unit level model is the Battese, Harter and Fuller (BHF). BHF model are however be preferred to the FH models since it incorporates variance of the model explanatory variables.

SAE approaches are however less widely applied due to lack of up to date covariate data. Kondlo and Manda[70] for example used covariates from census data. Census data is however collected 10-yearly rendering it less timely (i.e. not up to date) for monitoring the HIV/AIDS epidemic. HIV transmission pattern change rapidly and therefore up to date information is critical for well-targeted and timely intervention.

Censuses also collect a limited number of covariate data – mainly age and sex, hence are not appropriate for assessing HIV/AIDS dynamics as it HIV incidence and prevalence varies according to knowledge and behavioural risk factors [35,72–74]. This study explored application of the SAE methods, assessing efficiency of the methods based on the level at which covariate data is available.

Hybrid Prevalence Estimation Methodology

Another emerging statistical method in literature to obtain health indicator estimates for areas with small survey sample sizes is the HPE, an annealing methodology [3,75,76]. The methodology combines a small survey sample with routinely collected data to obtain more reliable health indicator estimates [10,77]. While the HPE methodology has not previously been applied to obtain HIV prevalence estimates for districts, work in this doctoral research represents its first assessment and application.

The method applies “Weights” to balance the potential bias of the direct estimators. It also generates estimates with reduced standard errors. The methodology was develop by Hedt and Pagano who used simulations to assess the efficiency of the estimates obtained under different survey designs [10,77] . Their findings show that the methodology generates estimates with narrower confidence intervals (reduced standard errors) compared to direct survey estimates irrespective of the survey design. The reduction in standard errors is mainly attributed to inclusion of more information (samples) to support parameter estimation especially for areas where surveyed samples are markedly reduced or non-existent.

Jeffery et al tested and applied the methodology to obtain estimates of coverage of vitamin A and Polio campaigns in Madagascar and Benin with the weights calculated to minimize the resulting standard errors” [78]. They found that standard errors were reduced by up to 6% [78]. Application to HIV prevalence estimation by Ouma et al, found a 29% reduction in the standard errors of the estimates compared to the standard errors of the direct survey estimates [23]. This methodology provides an appealing and straightforward approach to obtain health indicator estimates for geographical areas with small survey sample sizes

when resources are limited to conduct comprehensive surveys. It optimises use of available data for decision making amidst scarce resources.

2.4 Risk factors for HIV prevalence

Past studies highlight multiple risk factors/predictors for HIV infection. This research work sought to understand and/reassess these factors so as to incorporate them in the models implemented. Although SSA remains the hardest hit with the HIV epidemic, the impact varies across geographical areas and population subgroups even within countries [79,80]. Commonly cited factors for the uneven spread/distribution of the HIV epidemic in SSA include demographic, social-economic, behavioural as well as clinical factors [40,71,81–85]. Others include structural and policy related factors [86].

For example, older age (>24 years), female gender, informal or no education, having been previously married are some of the socio-demographic factors reported to be associated with higher/elevated HIV prevalence [74,87]. Behavioural factors associated with an elevated HIV prevalence include concurrent multiple sexual partners, non-condom use [36,88] while clinical factors include diagnosis of STI's [89] and male circumcision [21,90].

Structural and policy related factors include prohibitive laws that affect how individuals and organisations providing services for people living with HIV work to avert further spread of the epidemic. For example, Uganda passed an anti-LGBTQ law in 2012 prohibiting civil society organisations from engaging in or promoting services for Men who have sex with Men or other LGBT communities [91]. This implied that this vulnerable group of individuals cannot access HIV services freely while practicing their rights.

3.0 CHAPTER THREE: METHODOLOGY

This chapter is divided into Six sections namely the general setting and population of Uganda; data descriptions: sources, strengths and weakness; data and methods for the HPE methodology –objective one; data and methods for the SAE techniques -objective two; data and methods for objective three and data and methods for objective four. This chapter also presents the approach used for analysis for each objective and how missing data was handled in the study. It concludes with ethical considerations, permission obtained to use the different datasets and the assumptions made in the course of the study.

3.1 Setting and population

Uganda is a landlocked country, located in East Africa with a population of 34,634,650 people (Female: 50.7%) in 2014 [92]. The country has one of the highest population growth rates in the world at 3.0% and the population projection for 2019 was estimated to be 40 million. The age group 15-49 years form 44.2% of the population while 47.9% of the population are below the age of 15 years [92]

Administratively, the country is currently divided into 135 districts - the main administrative units mandated to implement and monitor programmes on behalf of the central government. The number of districts has grown over time from 21 in 1969 to 111 districts and one-capital, Kampala in 2014 and later to 135 districts as of July 2020. For this study, analysis was restricted to the administrative boundaries of the 111 districts. The further division of districts is aimed at improving service delivery and easing administration [92,93]. The district populations range from 54,293 in Kalangala to 1,997,418 in Wakiso district. The districts are grouped into 11 geographical regions, although the regions do not have administrative structures to implement and monitor service delivery [92].

The overall HIV prevalence in the country is 6.2% and varies from 3.1% in West Nile region to 8.0% in the Central 1 region [8]. Prevalence also varies by gender (Male – 4.7% and Female – 7.6%) and area of residence (Urban areas - 7.5% and Rural areas - 5.8%) [8]. Over 80% of the PLHIV are aged 15-49 years while Females aged 15-24 years are at higher risk of new infection compared to males of the same age group [94]. Although the number of individuals living with HIV has increased from 850,000 in 1990 to 1,500,000 in 2019, the number of annual new infections dropped significantly from 110,000 to 53,000 cases in the same period [95].

Decrease in the annual infections has been attributed to the aggressive efforts adopted by the country to combat the HIV epidemic. These efforts include adoption of the 90:90:90 UNAIDS targets, implementation of combination treatment strategies that include both behavioural and biomedical interventions. In fact Uganda is among the first countries to adopt the Abstinence, Being faithful and Condom use (ABC) strategy aimed at empowering individuals to adopt safe sexual practices to prevent new HIV infections [96]. This approach contributed to the earlier successes that Uganda registered in reversing the HIV/AIDS epidemic.

3.2 Data Sources: Strengths and Limitations

For this research, data from the District Health Information System (DHIS2), National population surveys conducted in 2011 and 2016 and Community Lot Quality Assurance survey conducted in 2016 were used. A description of each of the datasets is provided below.

3.2.1 District Health Information System

HIV testing data is collected during service delivery at health facilities and is reported to the National District Health Information System version 2 (DHIS2). The system acts as a national data repository for health service delivery data from health facilities across the country. Data is collected using the national Health Management Information System (HMIS) tools and reported as aggregated monthly totals into the DHIS2. The data comprise HIV testing data for pregnant women seeking antenatal care services, their family members, and the general population seeking care at health facilities. Data used for this study comprised aggregated district level data for all health facilities that report to the national HMIS.

Both government supported and some Private Not for Profit (PNFP) health facilities report their service delivery data into the system. In 2018 there were a total of 4092 (Government aided=3118 and PNFP=974) health facilities in the country. The system was setup in 2011 with the core goal of monitoring the National Health Sector Strategic Plans (HSSP). Training workshops are routinely conducted for district health teams, records assistants, Biostatisticians and HMIS focal persons, to empower them to use the system for data management and reporting. Technical agencies including WHO, UNAIDS and PEPFAR support Ministry of Health to develop SOPs and monitor the quality of data in the system through regular Data Quality Assessments (DQAs). Validation checks are also inbuilt in the system to ensure the data reported are of high quality [97,98].

Limitations of District Health Information System data

Although all health facilities offering health services in the country are expected to report their service delivery data into the system, some private health facilities do not report due to limited capacity to compile and submit reports into the system. Reporting rates for some

of the Government aided health facilities especially those located in rural and hard to reach areas is also low. The data is also aggregated by age group and gender and do not contain individual level data such as sexual behaviour and other HIV risk factor variables.

3.2.2 National population surveys

National population surveys are nationally representative, population-based surveys designed to provide accurate estimates at national and regional levels. The surveys include an HIV testing component and therefore aim to describe the status of the HIV epidemic including HIV prevalence in the country. The surveys use a two stage, stratified cluster sampling design. Enumeration areas (EA) (clusters) are selected at the first stage and households at the second stage. Uganda has conducted three population level surveys with the first one conducted in 2004/05. This study used data from the 2011/12 and 2016/17 surveys. During the 2011 and 2016 surveys, the country was stratified into ten geographical regions namely Central 1, Central 2, Kampala, East-Central, Mid-Eastern, North-East, West Nile, Mid-North, Mid-West, and South-West. The regions were created for purposes of the survey and are not administrative units of the country. Other than Kampala, each region comprised of 8 to 15 neighbouring districts that share similar languages and socio cultural characteristics. Uganda Bureau of Statistics (UBOS) provides country level technical and supervisory support for the surveys

Uganda AIDS indicator survey, 2011/12

The Uganda AIDS Indicator Survey (UAIS) was conducted from February to September 2011 [99]. All women and men aged 15-59 years, who slept at the selected household in the night before the survey were eligible for interview and blood draw for HIV testing. Out of 22,357 adults found in the sampled households, 97.2% (21,741) were interviewed, and 91% (19,866) of those interviewed were tested for HIV. Ministry of Health in collaboration

with UBOS implemented the survey while ICF International and the Centers for Disease Control and Prevention (CDC) provided financial and technical support. Details of the survey is available in the survey report [7].

For this study, UAIS 2011 data was analysed at individual level to determine factors associated with HIV positivity and at district level and at district to determine direct HIV prevalence HIV estimates. The HPE methodology was applied to obtain HIV prevalence estimates for districts.

Uganda Population and HIV/AIDS Impact Assessment, 2016/17

The Uganda population and HIV/AIDS Impact Assessment (UPHIA) was conducted countrywide between August 2016 and March 2017. The survey was administered to 12,386 households. In the households surveyed, 30,581 adults aged 15-64 years were eligible to participate. Of these, 96% (29,383) were interviewed, and 99% (29,024) of those interviewed provided blood samples for HIV testing.

The survey was implemented by Ministry of Heath Uganda while technical support was from CDC. Other partners include ICAP at Columbia University, Uganda Virus Research Institute (UVRI). Further details about the survey is available in the survey report [8]

Variables from the population survey data

For both the 2011 and 2016 survey data used for this study, the outcome variable was HIV status. The independent variables were: respondents' age, gender, highest level of education attained, marital status, number of sexual partners in the 12 months preceding the survey, distance to the nearest health facility, employment status, area of residence (urban or rural), district and region of the country.

Limitations of Population survey data

Population surveys are characterised by high levels of missingness and recall bias. It is also not possible to assess causality using census data since individuals are not followed up over time, for example whether the outcome follows an event. The data is a cross-sectional study generating demographic and health profiles of the population at single point in time.

3.2.3 Lot Quality Assurance Sampling surveys

In Uganda, Lot Quality Assurance Sampling (LQAS) surveys have been conducted annually since 2009. The primary objective of the LQAS methodology is for classification of sub project areas as acceptable or unacceptable based on a program coverage. It is applied to identify areas of "high" or "low" program coverage to institute corrective measures or actions including withdrawing resources from well performing areas and re-allocating those to low performing areas. A secondary objective of the methodology is to obtain program area coverage by aggregating sub program area values/responses.

In Uganda, a district represents a program area. During survey implementation, the district is sub-divided into Supervision Areas (SA) comprising of a sub-county or a set of neighbouring sub-counties. A complete listing of sampling units (household) from the most recent population census is used to sample villages and households for the survey.

Sampling is carried out at two stages; at first stage, Survey locations[1] (villages) are selected in each SA using probability proportional to the number of household or individuals in each village. For each SA, a random sample of 19 villages are randomly

[1] In some settings, multiple interview locations are selected for a given village if the number of individuals in the village is greater.

selected for districts with 5-7 SAs while for those with 4 SAs, a sample of 24 villages are randomly selected. A minimum sample of 19 and 24 respondents per SA for districts with 5-7 and 4 SAs respectively are selected so as to maintain the combined SA misclassification errors to less than 15%. These sample sizes also enable the district (program) level coverage to be estimated with an error margin of less than 10% [100].

At the second stage, a reference household is randomly identified using equal probability sampling within the village and the next nearest household from the front exit of the reference household is selected to start the interview. An eligible and consenting adult respondent is then interviewed from the selected household. Parallel sampling is applied to select eligible respondents for the population categories: mothers of children aged 0-11 months, mothers of aged children 12-23 months and mothers of children age 24-59 months, youth aged 15-24 years, women aged 15-49 years, and men aged 15-54 years. The design does not permit selection of youth, men and women from the same household given because similar indicators are assessed in these population subgroups. A total of 34,109 individuals were surveyed for the LQAS survey, hence the total sample size for the LQAS data included in the analysis. More details about the surveys are available in the survey reports [101].

The health areas covered in the surveys include HIV/AIDS, TB, Malaria, Water and sanitation, reproductive health, education and other social health services in the district. Data collection for a district takes between 5 and 7 days depending on the number of SAs within a district, nature of the terrain and the general weather condition. Supervision is carried out at 3 levels namely SA level, district level, and national level to ensure quality and adherence to methodological requirements. The district level supervisors also revisit a sample of the selected households to conduct reliability studies for quality assurance.

Variables from the LQAS data set

Variables extracted from the LQAS dataset include respondents' age, gender, highest level of education attained, marital status, number of sexual partners in the 12 months preceding the survey, area of residence (urban or rural) and district.

Limitation of the LQAS dataset

The survey does not draw blood for HIV testing; therefore, HIV status information for the respondents is unknown. Additionally, during the 2016 survey, only 70 out of the 112 districts in Uganda carried out LQAS surveys. This implies that indicator estimates could not be calculated for those districts that did not conduct LQAS surveys.

3.2.4 National Population and Housing Census

The most recent National Population and Housing Census (NPHC) in Uganda was carried out in 2014 [92]. Censuses are carried out after 10 years. The overall objective of the census is to provide data for planning at all levels including districts and sub-counties. Censuses collects individual level demographics and household level data such a age, gender, education level attained, marital status of household members, housing conditions; household based agricultural activities and deaths in the household in the 12 months preceding the census.

Variables from the NPHC data set

For this study, variables highlighted from literature to have a potential confounding effect on HIV prevalence were extracted from the 2014 NPHC and they include district population density, district population density, urbanicity (defined below) and health facility utilization. These variables were also significantly correlated with direct district HIV prevalence estimates at 5% level of significance during variable selection:

- Urbanicity; measures the level of urbanization of the district. It presents the number of people living in urban areas as a percentage of the total district population.
- Health facility utilization refers to the number of health facilities per 10,000 individuals in the population. It standardizes health facility utilization to enable comparison across districts and evaluates access to health facilities in the district.
- District population density refers to the number of people per square kilometer. This variable takes into account the population distribution in the district.

Limitations of the NPHC data set

The key limitation of the census data is that it may not accurately describe HIV prevalence variation across districts over time since it is collected decennially. It also collects mainly individual demographic information such as age and gender and does not include behaviour characteristics that are important for profiling the HIV epidemic.

3.3 How Missing data was handled in this study

For this research/study, complete case analysis was adopted for all the research objectives. It's noted that this could have affected the information generated making it inferentially biased as cases that had complete information may be different from those who did not respond to specific questions. We attempted to address this by including only variables that had less than 15% level of missingness. This may have had an impact on the information generated by the analysis or models. Information on condom use for example was missing for more than 38.3% of the participants in the LQAS dataset, hence the exclusion of condom use variable in our analysis.

3.4 The role of the PhD Candidate in data collection

The PhD candidate was the principal Statistician for the community LQAS surveys from 2009 to 2016. He participated in designing the surveys, compiling the survey tools and supervision of data collection, data entry and management; conducted all statistical analysis; and supported project and district teams to prepare survey reports. He supported dissemination of results during national information sharing meetings and data use by stakeholders. The PhD candidate extracted data from the different databases, conducted the analysis and wrote the first draft of all the manuscripts submitted to peer reviewed journals for publication as part of the PhD.

3.5 Ethical Consideration

The Human Research Ethics Committee (Medical) of the University of Witwatersrand, Johannesburg, South Africa provided approval to conduct the study (Ethics clearance certificate number **M171053**). The Uganda National Council of Science and Technology (UNCST) provided approval (Registration Number **HS 2366**) to use data collected in Uganda. Uganda Ministry of Health and ICF International provided permission to access the DHIS2 and 2011/12 Uganda AIDS Indicator Survey data respectively. Centres for Disease Control and Prevention, USA and the Ministry of Health, Uganda granted permission to use the 2016/17 Uganda Population HIV Impact Assessment data. Copies of the Ethics clearance certificates and letters of permission are in appendices I-IV

3.6 Data and methods for objective one

This section presents the statistical annealing technique, the HPE methodology, used to obtain HIV prevalence estimates for districts in Uganda. The HPE method combines a small survey sample with routine health facility data, to obtain more accurate indicator estimates for areas with small survey sample sizes. In this study, we combined health facility HIV testing with population survey data after adjusting for the selection bias in health facility data. This led to an overall improvement in the district sample size used for analysis, and therefore overcoming sample size limitation of the direct survey estimates.

3.6.1 Comparability of population survey and health facility testing data

During the 2011 UAIS, respondents were asked whether they had ever tested for HIV, if they answered, “YES”, they were asked follow-up questions regarding the period and the venue where the most recent test was taken. Period of testing was classified as (i) tested in the 12 months preceding the survey, or (ii) test taken more than 12 months before the survey while venue of testing was classified as (i) health facility or (ii) community settings.

Health facility setting comprised Health centres II; III; IV; general district hospitals; private health facilities/clinics and organisations offering HIV care and treatment services while community setting included tests conducted at community events; homes, workplaces, or at standalone VCT centres. More details regarding health facility classifications is given by Ouma et al, (2020) [34].

The 2011 UAIS data was stratified according to venue of the most recent HIV test, i.e. tested in a health facility or tested in a community setting. HIV prevalence ratio and the 95% confidence interval was computed using the KATZ [102] methodology. The data was weighted by the population survey weights and analysis was carried out in Stata Release 15 [103].

3.6.2 Hybrid prevalence Estimation

Assuming that the number of individuals, x, who test in a health facility follow a binomial distribution $x \sim b(n, \pi_c)$, where π_c is the propensity to test in a health facility in a one year period. This will be associated with the outcome, y, HIV prevalence at the district level, which also follows a binomial distribution, $y \sim b(n, p)$, then the HIV Hybrid Prevalence Estimate, $\hat{P}$, for districts was obtained using equation 1:

$$\hat{P} = \hat{\pi}_c P_f + (1 - \hat{\pi}_c)\hat{P}_{\underline{s}} \quad (1)$$

and the variance by

$$Var(\hat{P}) = \frac{1}{n}\left\{\hat{P}_{\underline{s}}(1 - \hat{P}_{\underline{s}})(1 - \hat{\pi}_c) + \hat{\pi}_c(1 - \hat{\pi}_c)\left(P_f - \hat{P}_{\underline{s}}\right)^2\right\} \quad (2)$$

Where:

$\hat{P}$ – HPE (combined estimate), $\hat{\pi}_c$ – the estimated probability of testing for HIV in a health facility, P_f – HIV prevalence for individuals tested at a health facility and $\hat{P}_{\underline{s}}$ – HIV prevalence for individuals tested during the survey and had not tested in a health facility.

A multilevel logistic regression model was fitted to obtain the propensity of testing for HIV in a health facility. HIV prevalence p and the associated individual and district level covariates were modeled using the following logistic regression model with a district random effects component.

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \boldsymbol{X_{ij}^T\beta} + \boldsymbol{u_i} + \boldsymbol{\varepsilon_{ij}} \quad (3)$$

$$u_i \sim i.i.d.Normal(0, \tau^2),\ \varepsilon_{ij} \sim i.i.d.Normal(0, \delta^2),$$

Where

X_{ij} - vector of covariates

β - vector of model parameters

u_i – effect of $i-th$ district

ε_{ij} – is the sampling error.

Variables included in the analysis were (i) at cluster level: area of residence (urban/rural) and region of the country and at (ii) individual level: respondents' gender, marital status, education level attained, number of sexual partners including husband/wife in the 12 months preceding the survey, employment status and distance to nearest health facility.

District level average propensity to test for HIV in a health facility was then obtained and used to combine the health facility and population survey data of non-facility testers to obtain district HIV prevalence estimates for the 111 districts in Uganda using equation 1.

Note: The proportion of individuals tested for HIV in each district can also be used as the propensity to test for HIV in health facility. We conducted this analysis and found negligible differences in the estimates of HIV prevalence.

Percentage change in standard errors for the direct survey and HPEs was computed to assess the improvement in efficiency of the HPEs while consistency between HPEs and the direct survey estimates was assessed using the Bland Altman analysis [104].

3.7 Data and methods for objective two

This objective explored use of other data sources, e.g. routine HIV testing data, as predictors. For this objective, multiple data sources were used according to the model implemented. Details for the data source used for each model is shown in table 1 and also explained under the respective statistical models implemented.

Table 1: Sources of data use for SAE methods/ Objective 2

Method	Source of outcome data	Source of Auxiliary data
Direct Survey estimate	UPHIA 2016/17	none
Fay-Herriot (FH)	UPHIA 2016/17	UPHC & Antenatal HIV testing
Battese, Harter and Fuller (BHF)	UPHIA 2016/17	LQAS 2016/17

3.7.1 Direct survey estimates

Under direct survey estimation, parameter estimates are computed based only on sampled observations from the area of interest.

If y_{ij}, is the HIV status (1-positive and 0-negative) for the $j-th$ individual in the $i-th$ district, p_i the proportion of people living with HIV in district i, the direct estimate of district HIV prevalence, is obtained as follows:

$$\hat{P}_i = A/B \qquad (1.1)$$

Where $A = \sum_i \sum_j w_{ij} y_{ij}$ and $B = \sum_i \sum_j w_{ij}$ and $i = 1,2,\dots,m,\ j = 1,2,\dots,N_i$. For m districts and N_i individuals in district i. and w_{ij} is the sampling weight of individual j from district i.

For the direct survey estimate, only data from the UPHIA 2016/17 survey was utilized.

3.7.2 Fay-Herriot area level model estimation

The most commonly used area level model is the Fay-Herriot (FH) model. It is a two stage hierarchical model, it links the outcome with possible predictors defined at area level [105].

Taking $\theta_i = g(\bar{y}_i)$, a logit function and relating it to $z_i = (z_{1i}, z_{2i}, \dots, z_{pi})$ a vector of p predictor variables from the m districts. A linking model for the area-level parameter θ_i is given in equation 2.1

$$\theta_i = {z_i}^T \beta + v_i \qquad (2.1)$$

Where: $\beta = \left(\beta_1, \beta_2, \ldots, \beta_p\right)^T$ is a $px1$ vector of regression coefficients and $\upsilon_i's$ are area specific random effects.

The $\upsilon_i's$, are assumed to be independent and identically distributed with $E(\upsilon_i) = 0$ and $var\ (\upsilon_i) = \sigma_v^2$. (i.e. $\upsilon_i \sim N(0,\ \sigma_v^2)$). The model variance, σ_v^2, measures homogeneity of the areas after accounting for the effects of the covariates, ${z_i}^T$, included in the model while e_i is the sampling error with known variance, $Var(e_i) = \ \psi_i$ and $E(e_i) = 0$.

Given that θ_i is unobservable; it is estimated based on the sampled observations in the area. The unbiased direct estimator of θ_i is obtained using a multivariate sampling model:

$$\hat{\theta}_i = \theta_i + e_i \qquad (2.2)$$

Where $e_i \sim N(0, \psi_i)$, $i = 1,2, \ldots, m$. for $i = 1,2, \ldots, m$, $e_i \sim N(0, \psi_i)$

Combining 2.1 and 2.2, we obtain the mixed model

$$\hat{\theta}_i = z_i\beta + \upsilon_i + e_i \qquad (2.3)$$

Where $\upsilon_i \sim N(0,\ \sigma_v^2)$, $e_i \sim N(0, \psi_i)$, $i = 1,2, \ldots, m$ υ_i is independent of e_i

The FH estimate is then obtained as a weighted combination of the direct ($\hat{\theta}_i$) and regression-synthetic estimators ($z_i\hat{\beta}$)

$$\hat{\theta}_i^{FH} = \hat{\gamma}_i\hat{\theta}_i + (1 - \hat{\gamma}_i)z_i\hat{\beta} \qquad (2.4)$$

Where: $\hat{\gamma}_i = \frac{\hat{\sigma}_v^2}{\hat{\sigma}_v^2 + \sigma_i^2}$

From 2.4 above, the FH district estimate will tend towards the direct district survey estimates for small sampling errors (large sample size) while it will tend towards the regression synthetic estimates for small sample size in the district (inefficient sample).

Letting $\hat{\tau}_i^{EB}$ to be the target quantity $\hat{\theta}_i^{FH}$

The bootstrap MSE estimator $\hat{\tau}_i^{EB}$ is obtained by

$$mse_B(\hat{\tau}_i^{EB}) = B^{-1}\sum_{b=1}^{B}[\hat{\tau}_i^{EB*}(b) - \tau_i^*(b)]^2 \qquad (2.5)$$

The FH model used data from UPHIA 2016/17, Antenatal HIV prevalence data and UPHC. Details on the implementation of the bootstrap procedure is available in appendix II of the published manuscript attached to this thesis as appendix V of this thesis

3.7.3 Unit-level model estimation

In situations when auxiliary information at individual or unit level is available, the Battese, Harter and Fuller (BHF) unit level model is frequently applied [106].

Letting p_{ij} be the probability of individual j from district i being HIV positive ($p_{ij} = Pr(\, y_{ij} = 1 \mid x_{ij}, \mu_i)$, where y_{ij} corresponds to the HIV status of individual i in district, a logistic regression model with area-level effect is used to estimate p_i

If $X_{ij} = (x_{ij1},\ x_{ij2},\ ,\ x_{ij3},\ \ldots.\ x_{ijq},\ )$ are individual specific characteristics associated with the binary outcome y_{ij}, the nested error logistic regression model 3.1 was applied to obtain parameter estimates

$$\mathrm{logit}(\mathrm{p_{ij}}) = \log\left(\frac{\mathrm{p_{ij}}}{1-\mathrm{p_{ij}}}\right) = \mathrm{X_{ij}^{t}}\beta + \upsilon_{\mathrm{i}} + \mathrm{e_{ij}} \qquad (3.1)$$

And the Empirical Best Predictor (EBP) [105,107] was applied to get district specific estimates of HIV prevalence.

$$\hat{p}_i = \frac{1}{N_i}\left\{\sum_{j\in S} y_{ij} + \sum_{j\in S'} \hat{p}_{ij}\right\} \qquad (3.2)$$

Where:

- $\hat{p}_{ij} = \frac{exp(\hat{\eta}_{ij})}{1+exp(\hat{\eta}_{ij})}$ and $\hat{\eta}_{ij} = \hat{\beta}X^{t}{}_{ij} + \hat{v}_i$
- y_{ij} - takes on the values 0 or 1 depending on whether the $j-th$ individual within the $i-th$ district was tested HIV positive during the survey and
- N_i – is the population size of the $i-th$ district.

Taking $j \in S$ as the sum of n_i values who were tested HIV positive for the sampled individuals from the $i-th$ district and $j \in S'$ of $\hat{p}_{ij}$ as the sum over the estimated probability of infection for the non-sampled individuals in district i.

The mean square error (MSE) of $\hat{p}_i$ as a predictor for p_i was obtained using equation 3.3

$$\widehat{MSE} = \text{var}\left(\frac{\sum_{j\in S'}\hat{p}_{ij}}{N_i}\right) + \frac{\sum_{j\in S'}\hat{p}_{ij}(1-\hat{p}_{ij})}{N_i^2} \qquad (3.3)$$

For the BHF model, data from UPHIA 2016/17 and Community LQAS surveys conducted in 2016 were utilized.

As sensitivity analysis, models were fitted allowing for spatial correlation between the district level random effects.

3.8 Data and methods for objective three

This objective was concerned with assessing the factors associated with HIV positivity within and between districts. It also assessed factors associated with HIV positivity by venue of testing and the factors associated with testing for HIV in a health facility. These findings were critical for understanding the HIV risk profiles for individuals who seek HIV testing services at health facilities compared to those who are testing at community settings including at their homes or work places.

3.8.1 Data source

In this objective, the 2011 UAIS data was used. Analysis was restricted to 11,685 women and men aged 15-49 years, who consented and were tested for HIV during the survey.

3.8.2 Measures/study variables

(1) *Ever tested for HIV:* Survey participants were asked whether they tested for HIV in the 12 months preceding the survey, venue of testing and whether they received their HIV test results. Women who had a birth in the period 2006 to 2011 and were not tested for HIV or had a negative test result during antenatal visits were further asked questions relating to testing during antenatal care attendance. Details of the questions asked are available in the survey report [7].

(2) *Facility testers* were individuals tested for HIV in a health facility in the 12 months preceding the survey and who received their test results. It also includes all women who tested in a health facility during antenatal care attendance for their most recent birth of a child aged less than one year.

(3) *Non-facility testers* were individuals who did not test in a health facility in the 12 months preceding the survey. Those who had never tested for HIV were not included in the analysis.

(4) *Health facilities* included public or private health facilities or non-government or civil society organizations offering HIV testing services.

(5) *Outcome variable* : *HIV positivity*; defined as the proportion of individuals testing HIV positive

(6) *Model explanatory/independent variables were*: area of residence (urban/rural), region of the country, respondents' gender, age (15-19, 20-29, 30-39 and 40-49 years), marital status, highest level of education attained, number of sexual partners including

husband/wife in the 12 months preceding the survey (0,1 and 2 or more), employment status and distance to nearest health facility.

3.8.3 Statistical analysis

We computed the proportion who had ever tested for HIV and HIV prevalence by place of testing. Factors associated with HIV positivity for each testing venue were assessed using multilevel logistic regression. Models were fitted separately for each testing venue (i.e. facility and non-facility testing) and a summary model fitted to identify factors associated with HIV positivity irrespective of the venue of testing. All analysis was weighted using population sampling weights. Analysis was carried out in Stata Release 15 [103].

3.9 Data and methods for objective four

3.9.1 Datasets for the study

This objective used data from the UAIS 2011 and UPHIA 2016. For the 2011 survey, EAs were based on the 2001 NPHC [108] while for the 2016 survey, EAs were based on the 2014 NPHC [109]. Analysis was restricted to data for individuals aged 15-49 years who were tested for HIV during the survey.

3.9.2 Measures/Study variables

(1) *HIV status (*1=positive or 0=negative) was the outcome of interest in this study.

(2) *Model explanatory/independent variables*: gender (1=male, 2=female); age (15-19, 20-24, 25-29, 30-34, 35-39 and 40-49-years); marital status (never, married/cohabiting and previously married-widowed/separated/divorced); highest level of education (none, primary, secondary and tertiary); number of sexual partners including husband/wife in the 12 months preceding the survey (0, 1 and 2 or more); area of residence (1=urban,

2=rural); and region of the country (Central 1, Central 2, Kampala, East-Central, Mid-Eastern, North-East, West Nile, Mid-North, Mid-West, and South-West).

3.9.3 Statistical analysis

Analysis to evaluate participants' distribution by the socio-demographic characteristics was done separately for the UAIS 2011 and UPHIA 2016 surveys. Descriptive analysis was also carried out to assess change in HIV prevalence between the two surveys.

Pooled logistic regression was used to assess factors associated with HIV positivity in the pooled dataset. Multivariable logistic decomposition models were used to determine the contribution of characteristics and coefficients of the characteristics to change in HIV prevalence between the two surveys.

The Blinder-Oaxaca decomposition method [27,28] was used to display the contribution of each variable to the overall difference in characteristics or the effects of characteristics.

Considering an outcome Y as function of a linear combination of predictors and regression coefficients and considering two groups A and B:

$$\mathrm{Y} = \mathrm{F}(X\beta) = \frac{e^{X\beta}}{1+e^{X\beta}}, \tag{4.1}$$

Where Y is an $Nx1$ dependent variable vector, X is NxK matrix of independent variables and β is a $Kx1$ vector of coefficients. The mean difference in Y between the groups A and B is decomposed as follows

$$\begin{aligned} \bar{Y}_A - \bar{Y}_B &= \overline{\mathrm{F}(X_A\beta_A)} - \overline{\mathrm{F}(X_B\beta_B)} \\ &= \underbrace{\left\{\overline{\mathrm{F}(X_A\beta_A)} - \overline{\mathrm{F}(X_B\beta_A)}\right\}}_{E} + \underbrace{\left\{\overline{\mathrm{F}(X_B\beta_A)} - \overline{\mathrm{F}(X_B\beta_B)}\right\}}_{E} \end{aligned} \tag{4.2}$$

Where:

- Component E refers to the part of the differential attributable to differences in endowments or characteristics (explained component/characteristics effects).
- Component C refers to the part of the differential attributable to differences in coefficients (Coefficient effects/unexplained component).

As shown above, the methodology partitions the average difference in an outcome/dependent variable between two groups (or time points) into the part that is due to the differences in the composition of independent variables within the groups and the part that is due to the group differences in the effects of the independent variables[27]. In this study, we applied the methodology to partition the difference in HIV prevalence between the 2011 and 2016 surveys to the part due to the socio-demographic composition of the survey participants in the 2011 compared to the 2016 survey, and the part due to the effects of the socio-demographic characteristics of the survey participants (Coefficient effects). All analyses were weighted using the survey weights and was carried out using Stata Release 15 [103].

4.0 CHAPTER FOUR: RESULTS

For this research, two papers, presenting results of objective one and three, have been published in international peer reviewed journals. A manuscript, presenting results for objective two has been submitted and is under review while the last manuscript, presenting results for objective four is in draft, yet to be submitted for publication. A summary of results from the papers and manuscripts are presented in this thesis according to the objectives of the PhD research. Some results contributing to a particular objective were presented in the two papers and this is highlighted in the respective sections.

4.1 District HIV prevalence based on HPE methodology

This objective was concerned with applying the HPE methodology to obtain HIV prevalence estimates for districts in Uganda. It also assessed the robustness and biases associated with the methodology. Results of this objective have been published in two papers [23,34]

4.1.1 Comparing HIV prevalence by venue of testing

In order to obtain the district HIV prevalence estimates, comparability of health facility and population survey data was evaluated. Methods used to evaluate comparability, compute HIV prevalence, HIV prevalence ratio and the 95% confidence interval are described by Ouma et al (2020) [34].

Generally, testing in a health facility was associated with a higher HIV positivity compared to testing in a community setting (HIV prevalence ratio 1.8, 95% CI: 1.4-2.2). There was a two-fold or higher increase in HIV positivity among health facility testers compared to non-facility testers if they were male; aged 15-19 years, 30-39 years and 40-49 years; had

no formal education; never married or were previously married; reported no sexual partner in the 12 months preceding the survey; not in formal employment and resident in urban areas. Regions that had two-fold or higher increase in HIV prevalence ratio were Kampala, East Central and Northern regions (Table 2).

Table 2: Proportion tested for HIV in a health facility in the 12 months preceding the survey and HIV prevalence by testing venue

Characteristic	Tested in Health facility in 12 months preceding the survey				Tested in community setting		HIV Prevalence Ratio (95% CI)
	Number tested (%)		n#	Weighted Prevalence (95% CI)	n#	Weighted Prevalence (95% CI)	
Total	**6,920**	**59.0**	**722**	**10.9 (10.0, 11.7)**	**100**	**6.2 (5.4, 7.0)**	**1.8 (1.4, 2.2)**
Gender							
Male	1,413	34.9	169	12.6 (10.6, 14.7)	37	4.6 (3.0, 6.2)	2.7 (1.9, 3.9)
Female	5,507	72.7	553	10.4 (9.5, 11.3)	63	7.8 (5.7, 9.9)	1.3 (1.0, 1.7)
Age							
15-19	724	50.1	37	5.4 (3.5, 7.4)	5	1.7 (1.9, 3.3)	3.2 (1.3, 8.0)
20-29	3,221	67.1	225	7.2 (6.2, 8.2)	35	6.7 (4.2, 9.2)	1.1 (0.8, 1.5)
30-39	2,109	62.0	286	14.7 (12.9, 16.4)	30	6.3 (3.9, 8.6)	2.3 (1.6, 3.3)
40-49	866	43.4	174	20.3 (17.4, 23.2)	30	10.3 (6.5, 14.0)	2.0 (1.4, 2.8)
Education Level							
No Education	782	63.7	98	13.3 (10.7, 15.9)	10	5.5 (1.9, 9.1)	2.4 (1.3, 4.5)
Primary	3,917	61.3	446	12.1 (10.9, 13.2)	60	7.5 (5.4, 9.5)	1.6 (1.2, 2.1)
Secondary	2,221	55.5	178	8.0 (6.7, 9.3)	30	4.8 (3.0, 6.6)	1.7 (1.1, 2.4)
Marital status							
Never married	891	42.5	54	6.4 (4.5, 8.2)	11	2.9 (0.8, 5.1)	2.2 (1.2, 4.2)
Married/Cohabiting	5,204	63.7	445	8.8 (7.9, 9.7)	67	6.8 (5.1, 8.5)	1.3 (1.0, 1.7)
Previously married	825	61.5	223	29.7 (26.1, 33.3)	22	13.5 (7.8, 19.1)	2.2 (1.5, 3.3)
Number of sexual partners in 12 months preceding survey							
0	872	45.1	166	20.2 (17.1, 23.4)	20	5.3 (2.9, 7.8)	3.8 (2.4, 6.0)
1	5,585	64.8	499	9.3 (8.4, 10.1)	60	6.3 (4.5, 8.1)	1.5 (1.1, 1.9)
2+	463	43.2	57	12.8 (9.4, 16.3)	19	7.7 (4.2, 11.3)	1.7 (1.0, 2.7)
Currently working							
No	1,820	64.3	158	8.7 (7.2, 10.2)	21	3.9 (2.1, 5.8)	2.2 (1.4, 3.5)
Yes	5,100	58.0	564	11.6 (10.6, 12.6)	79	7.1 (5.4, 8.8)	1.6 (1.3, 2.0)
Distance to nearest HF							
<2	1,493	62.0	171	11.3 (9.4, 13.2)	18	6.2 (2.6, 9.7)	1.8 (1.1, 2.9)
2-5	2,852	59,3	256	9.4 (8.1, 10.6)	34	5.5 (3.5, 7.4)	1.7 (1.2, 2.4)
5+	2,341	58.5	275	12.5 (11.0, 14.0)	44	7.2 (5.0, 9.4)	1.7 (1.3, 2.4)
Don't Know	234	55.8	20	10.0 (4.8, 15.2)	4	6.0 (0.1, 11.8)	1.7 (0.6, 4.7)
Area of residence							
Rural	5,284	59.6	530	10.5 (9.6, 11.4)	77	6.3 (4.8, 7.7)	1.7 (1.3, 2.1)
Urban	1,636	59.2	192	12.1 (10.1, 14.1)	23	6.2 (2.9, 9.5)	2.0 (1.3, 3.0)
Region							
Central 1	662	57.5	95	14.8 (11.7, 17.9)	14	9.6 (4.0, 15.2)	1.5 (0.9, 2.6)
Central 2	713	60.2	81	11.0 (8.5, 13.5)	11	6.1 (2.3, 9.9)	1.8 (1.0, 3.3)
Kampala	803	54.1	81	10.3 (7.8, 12.7)	13	5.1 (2.1, 8.1)	2.0 (1.1, 3.6)
East Central	547	55.0	48	9.8 (7.0, 12.5)	8	4.8 (1.4, 8.3)	2.0 (1.0, 4.2)
Mid-Eastern	544	60.5	31	5.4 (3.5, 7.3)	10	7.1 (2.8, 11.4)	0.8 (0.4, 1.5)
North East	711	63.4	62	7.6 (5.4, 9.8)	11	7.5 (2.9, 12.0)	1.0 (0.5, 1.9)
West Nile	680	54.8	47	7.6 (5.3, 9.8)	9	5.3 (1.8, 8.8)	1.4 (0.7, 2.9)
Mid Northern	886	62.8	106	12.5 (9.8, 15.1)	6	2.6 (0.4, 4.9)	4.8 (2.1,10.8)
South Western	639	62.5	86	13.4 (10.6, 16.2)	9	9.7 (3.2, 16.2)	1.4 (0.7, 2.6)
Mid-Western	735	60.8	85	11.6 (9.1, 14.1)	9	6.3 (2.1, 10.6)	1.8 (0.9, 3.6)

*#- number HIV positive, *HIV prevalence ratio comparing health facility and community testers*

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Levin J. Difference in HIV prevalence by testing venue: results from population level survey in Uganda. AIDS Care [Internet]. 2020;0(0):1–12. Available from: https://doi.org/10.1080/09540121.2020.1734179

4.1.2 Propensity to test for HIV in a health facility

To apply HPE methodology, the probability of testing for HIV in a health facility was computed at individual level. Regional level average propensity to test for HIV in a health facility is presented in Figure 2. It shows that the propensity to test in a health facility was generally higher among females across all regions of the country. It was also higher in Mid-Northern, North East and Kampala regions (Figure 2).

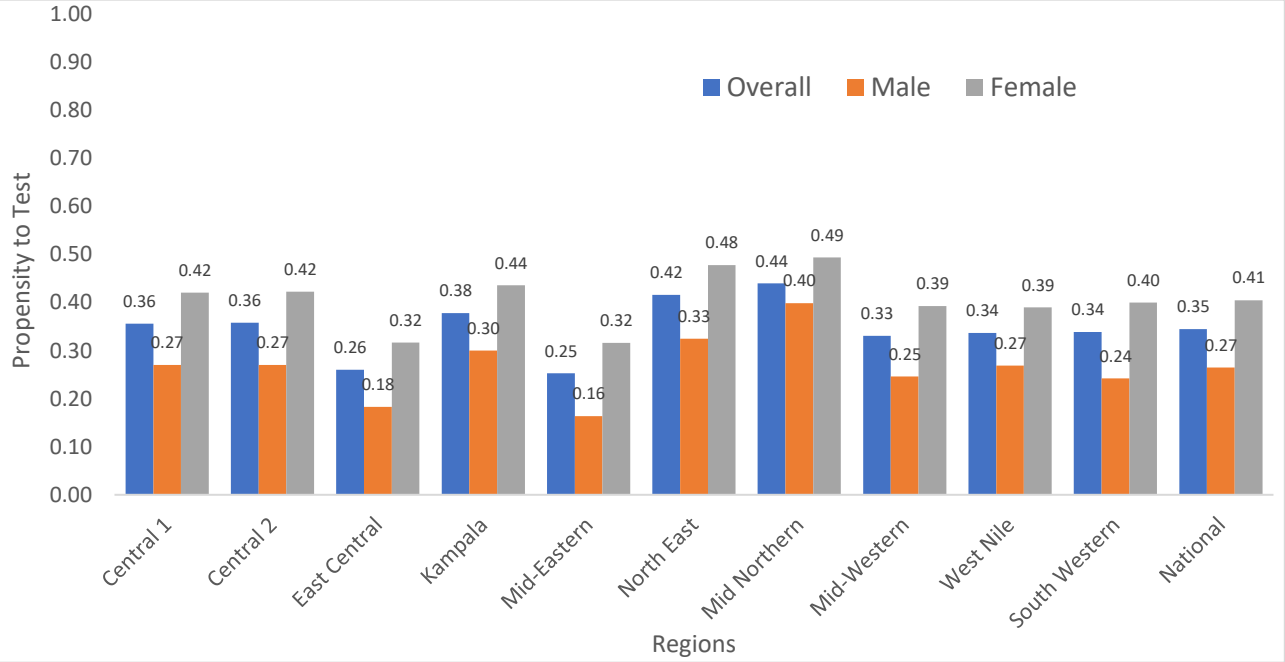


Figure 2: Propensity to test for HIV in a health facility

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020;20(1):1–14.

4.1.3 HIV prevalence estimates based on the HPE methodology

The HIV prevalence estimate for districts based on HPE methodology is presented in Table 3. It shows that the confidence intervals for the HPEs are narrower compared to CIs based on the direct survey estimates. It also shows that Central 1 region had the highest HIV prevalence in the country (0.11) while the Mid-Western and Mid-Eastern regions had the lowest HIV prevalence (0.04) (Table 3).

District level HPEs ranged from 0.01 to 0.18. Districts with the lowest and highest HIV prevalence by region were: Central (Bukomasimbi - 0.06, Masaka - 0.18); Central 2: (Nakaseke - 0.07, Buvuma - 0.17); East Central (Luuka - 0.02, Namayingo - 0.09); Mid-Eastern: (Kween - 0.01, Kapchorwa - 0.09); Mid Northern: (Kole - 0.05, Pader and Otuke - 0.14); Mid-Western: (Bundibujyo - 0.03, Kabarole - 0.16); North East (Kabong-0.04, Amuria - 0.10); South West: (Buhweju and Kabale - 0.04, Rubirizi 0.13) and West Nile (Yumbe and Maracha - 0.03, Nebbi - 0.07). Table 3 also presents direct survey and DHIS2 based HIV prevalence estimates by district.

HIV prevalence was higher in more urbanized districts (for example Kampala, Fort Portal, Mbale, and Jinja) and those at the lakes shores (e.g. Kalangala, Buvuma and Masaka).

District specific estimates for males and females were also computed and compared with the direct survey estimates. Findings showed that HPE estimates were more precise compared to the direct survey estimates. (see Ouma et al [23]).

Table 3: HPE HIV prevalence estimates, (HPE, Survey and unadjusted DHIS2)

Region and District	HPE		Population Survey		Facility Data (unadjusted)
	Estimate (95% CI)		**Estimate** (95% CI)		
Central 1					
Bukomansimbi	0.064	(0.027, 0.101)	0.074	(0.021, 0.127)	0.083
Butambala	0.103	(0.017, 0.190)	0.124	(0.000, 1.000)	0.063
Gomba	0.090	(0.036, 0.144)	0.158	(0.060, 0.257)	0.063
Kalangala	0.138	(0.064, 0.212)	0.195	(0.000, 1.000)	0.135
Kalungu	0.142	(0.065, 0.219)	0.102	(0.025, 0.179)	0.086
Lwengo	0.143	(0.090, 0.195)	0.157	(0.090, 0.223)	0.105
Lyantonde	0.121	(0.012, 0.229)	0.109	(0.000, 1.000)	0.085
Masaka	0.178	(0.129, 0.227)	0.156	(0.096, 0.216)	0.131
Mpigi	0.097	(0.065, 0.129)	0.108	(0.059, 0.158)	0.101
Rakai	0.081	(0.053, 0.109)	0.076	(0.042, 0.109)	0.067
Sembabule	0.072	(0.037, 0.106)	0.067	(0.015, 0.118)	0.078
Wakiso	0.107	(0.090, 0.124)	0.096	(0.070, 0.123)	0.094
Central 2					
Buikwe	0.083	(0.060, 0.106)	0.079	(0.042, 0.116)	0.086
Buvuma	0.170	(0.102, 0.238)	0.185	(0.096, 0.273)	0.085
Kayunga	0.069	(0.041, 0.096)	0.065	(0.030, 0.100)	0.049
Kiboga	0.090	(0.035, 0.145)	0.063	(0.008, 0.117)	0.069
Kyankwanzi	0.111	(0.049, 0.173)	0.128	(0.054, 0.203)	0.045
Luwero	0.094	(0.066, 0.123)	0.080	(0.044, 0.117)	0.075
Mityana	0.136	(0.097, 0.175)	0.104	(0.058, 0.151)	0.125
Mubende	0.077	(0.053, 0.101)	0.094	(0.056, 0.132)	0.057
Mukono	0.090	(0.063, 0.116)	0.092	(0.054, 0.129)	0.069
Nakaseke	0.073	(0.038, 0.109)	0.083	(0.025, 0.141)	0.083
Nakasongola	0.083	(0.041, 0.125)	0.076	(0.018, 0.134)	0.088
East Central					
Bugiri	0.037	(0.019, 0.055)	0.027	(0.008, 0.045)	0.029
Buyende	0.055	(0.020, 0.091)	0.044	(0.006, 0.081)	0.033
Iganga	0.062	(0.037, 0.087)	0.054	(0.025, 0.082)	0.035
Jinja	0.088	(0.064, 0.112)	0.090	(0.056, 0.124)	0.057
Kaliro	0.036	(0.001, 0.071)	0.046	(0.000, 0.098)	0.023
Kamuli	0.051	(0.034, 0.068)	0.058	(0.033, 0.082)	0.036
Luuka	0.023	(0.005, 0.042)	0.018	(0.000, 0.037)	0.018
Mayuge	0.042	(0.023, 0.062)	0.092	(0.043, 0.140)	0.036

Region and District	HPE		Population Survey		Facility Data (unadjusted)
	Estimate (95% CI)		**Estimate** (95% CI)		
Namayingo	0.089	(0.043, 0.135)	0.093	(0.035, 0.151)	0.079
Namutumba	0.037	(0.011, 0.062)	0.038	(0.000, 0.076)	0.030
Kampala	0.076	(0.067, 0.085)	0.071	(0.058, 0.084)	0.098
Mid-Eastern					
Budaka	0.065	(0.018, 0.111)	0.050	(0.007, 0.093)	0.029
Bududa	0.025	(0.000, 0.055)	0.046	(0.006, 0.087)	0.014
Bulambuli	0.026	(0.000, 0.055)	0.023	(0.000, 0.053)	0.040
Busia	0.068	(0.038, 0.097)	0.074	(0.037, 0.112)	0.063
Butaleja	0.026	(0.003, 0.049)	0.026	(0.001, 0.052)	0.018
Kapchorwa	0.092	(0.000, 0.186)	0.097	(0.000, 1.000)	0.021
Kibuku	0.036	(0.000, 0.074)	0.014	(0.000, 0.040)	0.027
Kween	0.013	(0.000, 0.032)	0.007	(0.000, 0.022)	0.022
Manafwa	0.034	(0.014, 0.055)	0.034	(0.006, 0.062)	0.035
Mbale	0.037	(0.022, 0.052)	0.035	(0.015, 0.055)	0.070
Pallisa	0.012	(0.002, 0.022)	0.007	(0.000, 0.019)	0.021
Sironko	0.051	(0.022, 0.079)	0.050	(0.020, 0.080)	0.030
Tororo	0.071	(0.044, 0.097)	0.069	(0.041, 0.098)	0.044
Mid Northern					
Agago	0.080	(0.047, 0.113)	0.046	(0.011, 0.081)	0.071
Alebtong	0.066	(0.035, 0.097)	0.058	(0.016, 0.100)	0.068
Amolatar	0.108	(0.058, 0.159)	0.158	(0.079, 0.237)	0.078
Amuru	0.059	(0.027, 0.091)	0.028	(0.000, 0.061)	0.043
Apac	0.062	(0.043, 0.080)	0.052	(0.022, 0.082)	0.081
Dokolo	0.070	(0.041, 0.099)	0.055	(0.015, 0.094)	0.061
Gulu	0.090	(0.066, 0.113)	0.100	(0.049, 0.151)	0.086
Kitgum	0.081	(0.042, 0.121)	0.083	(0.024, 0.141)	0.072
Kole	0.047	(0.023, 0.070)	0.029	(0.001, 0.057)	0.051
Lamwo	0.088	(0.041, 0.134)	0.121	(0.043, 0.198)	0.071
Lira	0.108	(0.079, 0.138)	0.111	(0.066, 0.157)	0.098
Nwoya	0.123	(0.033, 0.212)	0.163	(0.000, 1.000)	0.052
Otuke	0.136	(0.048, 0.224)	0.128	(0.000, 1.000)	0.075
Oyam	0.066	(0.043, 0.089)	0.068	(0.035, 0.100)	0.058
Pader	0.140	(0.088, 0.192)	0.165	(0.090, 0.239)	0.087
Mid-Western					
Buliisa	0.065	(0.013, 0.117)	0.038	(0.000, 0.785)	0.065

Region and District	HPE		Population Survey		Facility Data (unadjusted)
	Estimate (95% CI)		**Estimate** (95% CI)		
Bundibugyo	0.032	(0.010, 0.054)	0.032	(0.000, 0.071)	0.028
Hoima	0.075	(0.050, 0.100)	0.086	(0.051, 0.121)	0.053
Kabarole	0.159	(0.124, 0.194)	0.137	(0.094, 0.180)	0.099
Kamwenge	0.065	(0.037, 0.093)	0.054	(0.023, 0.086)	0.050
Kasese	0.066	(0.046, 0.085)	0.050	(0.027, 0.073)	0.050
Kibaale	0.068	(0.045, 0.091)	0.079	(0.046, 0.113)	0.049
Kiryandongo	0.059	(0.019, 0.098)	0.071	(0.016, 0.126)	0.048
Kyegegwa	0.118	(0.058, 0.177)	0.127	(0.056, 0.198)	0.046
Kyenjojo	0.077	(0.047, 0.106)	0.121	(0.071, 0.172)	0.070
Masindi	0.063	(0.035, 0.091)	0.060	(0.023, 0.097)	0.062
North East					
Abim	0.066	(0.004, 0.128)	0.098	(0.000, 1.000)	0.031
Amudat	0.090	(0.000, 0.204)	0.049	(0.000, 0.899)	0.022
Amuria	0.102	(0.067, 0.138)	0.102	(0.056, 0.148)	0.039
Bukedea	0.044	(0.014, 0.074)	0.045	(0.002, 0.088)	0.022
Kaabong	0.037	(0.016, 0.057)	0.021	(0.000, 0.044)	0.022
Kaberamaido	0.081	(0.045, 0.117)	0.061	(0.020, 0.102)	0.034
Katakwi	0.074	(0.041, 0.107)	0.080	(0.034, 0.126)	0.035
Kotido	0.052	(0.009, 0.095)	0.046	(0.000, 0.096)	0.017
Kumi	0.040	(0.023, 0.057)	0.025	(0.000, 0.051)	0.024
Moroto	0.041	(0.014, 0.069)	0.065	(0.000, 0.132)	0.023
Nakapiripirit	0.049	(0.002, 0.097)	0.038	(0.000, 0.792)	0.018
Napak	0.075	(0.018, 0.132)	0.073	(0.003, 0.143)	0.023
Ngora	0.046	(0.015, 0.078)	0.061	(0.014, 0.108)	0.020
Serere	0.049	(0.029, 0.069)	0.046	(0.014, 0.077)	0.028
Soroti	0.089	(0.060, 0.118)	0.058	(0.019, 0.097)	0.034
South Western					
Buhweju	0.035	(0.000, 0.077)	0.022	(0.000, 0.592)	0.025
Bushenyi	0.096	(0.060, 0.132)	0.079	(0.036, 0.123)	0.068
Ibanda	0.065	(0.027, 0.102)	0.053	(0.002, 0.104)	0.059
Isingiro	0.102	(0.061, 0.143)	0.112	(0.064, 0.160)	0.045
Kabale	0.043	(0.024, 0.063)	0.037	(0.010, 0.064)	0.035
Kanungu	0.077	(0.042, 0.112)	0.084	(0.036, 0.131)	0.047
Kiruhura	0.095	(0.046, 0.144)	0.086	(0.025, 0.148)	0.072
Kisoro	0.053	(0.015, 0.091)	0.059	(0.012, 0.105)	0.026

Region and District	HPE		Population Survey		Facility Data (unadjusted)
	Estimate (95% CI)		**Estimate** (95% CI)		
Mbarara	0.101	(0.068, 0.134)	0.124	(0.072, 0.177)	0.059
Mitooma	0.097	(0.042, 0.151)	0.076	(0.012, 0.140)	0.069
Ntungamo	0.063	(0.040, 0.085)	0.062	(0.030, 0.095)	0.050
Rubirizi	0.133	(0.067, 0.199)	0.140	(0.044, 0.236)	0.057
Rukungiri	0.082	(0.053, 0.111)	0.076	(0.038, 0.115)	0.064
Sheema	0.095	(0.056, 0.134)	0.124	(0.063, 0.185)	0.069
West Nile					
Adjumani	0.044	(0.020, 0.067)	0.039	(0.012, 0.065)	0.023
Arua	0.043	(0.031, 0.056)	0.049	(0.030, 0.069)	0.035
Koboko	0.048	(0.019, 0.077)	0.063	(0.025, 0.102)	0.025
Maracha	0.029	(0.001, 0.058)	0.007	(0.000, 0.019)	0.014
Moyo	0.033	(0.011, 0.055)	0.037	(0.005, 0.069)	0.019
Nebbi	0.068	(0.045, 0.091)	0.070	(0.042, 0.098)	0.036
Yumbe	0.026	(0.010, 0.042)	0.027	(0.006, 0.048)	0.014
Zombo	0.056	(0.031, 0.082)	0.040	(0.011, 0.069)	0.050

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020;20(1):1–14.

4.1.4 Precision of HIV prevalence estimates

Hybrid Prevalence Estimates had generally narrower confidence intervals compared to direct survey based estimates (Figure 3 and Figure 4). Figure 3 presents the comparison for selected districts while figure 4 presents comparisons for all districts. Of the 111 districts included in the analysis, 105 (95.5%) had narrower confidence intervals compared to confidence intervals from direct survey based estimates. Overall, the HPE standard errors decreased by 28.9% (95% CI: 23.4-34.4) compared to survey-based standard errors.

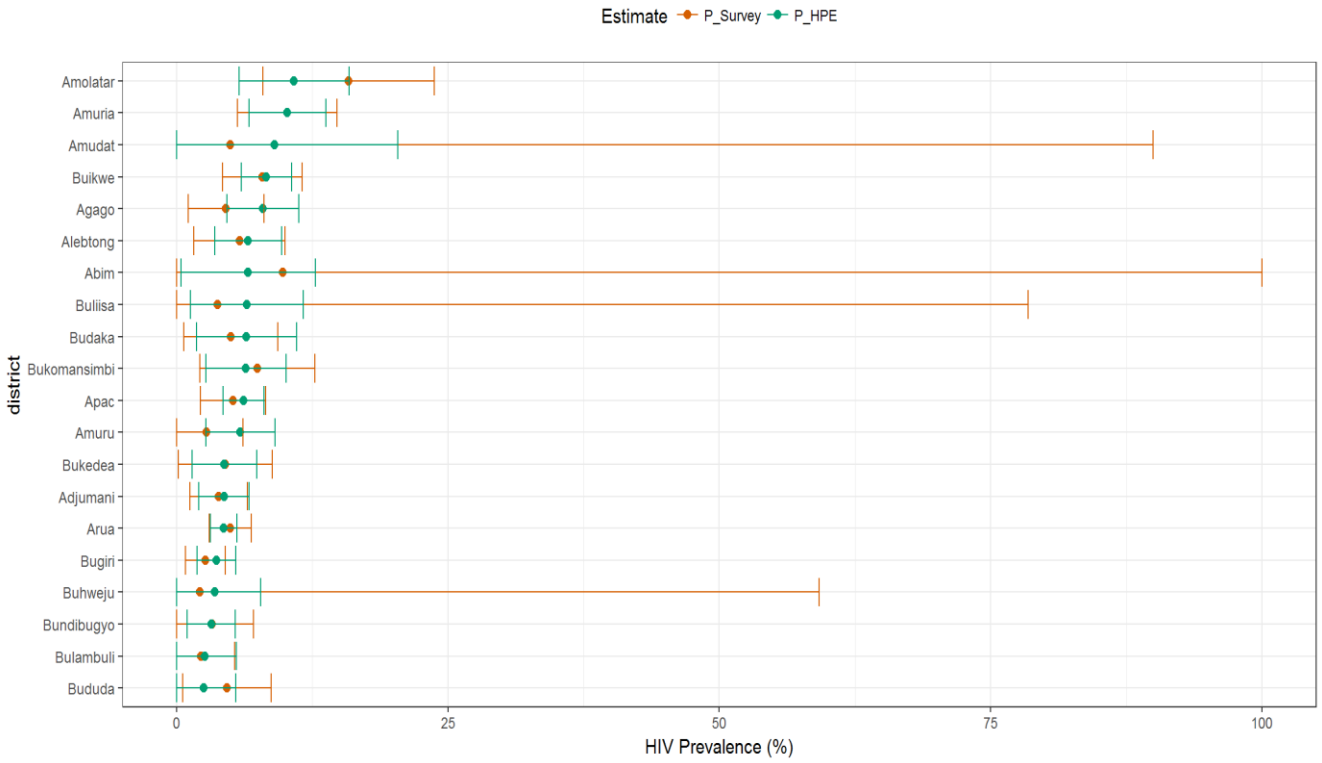


Figure 3: District prevalence estimates from combined and population survey data

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020;20(1):1–14.

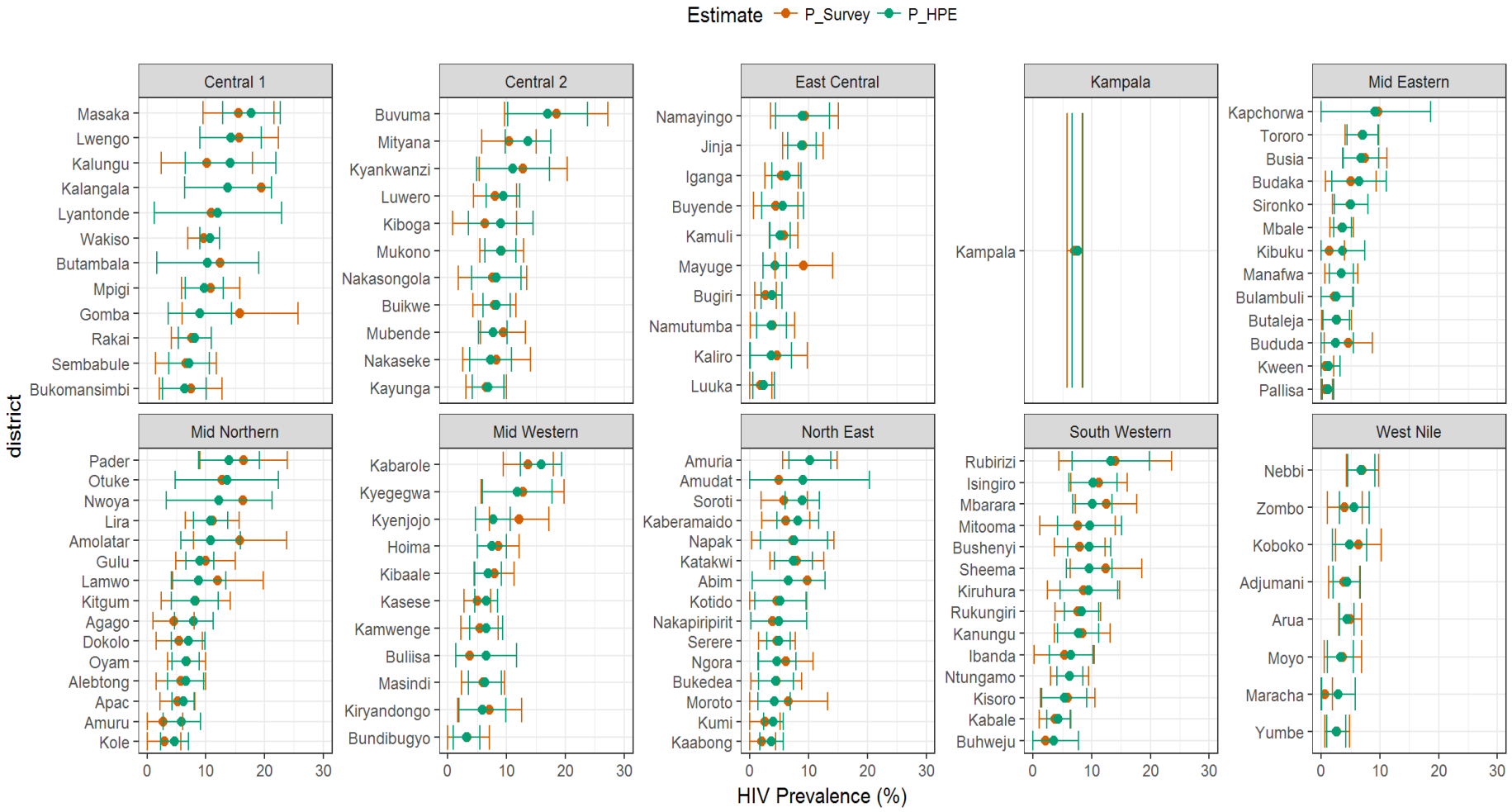


Figure 4: Comparison of district prevalence estimates for the HP and survey-based estimates

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020;20(1):1–14.

4.1.5 Robustness of the HIV prevalence estimates

The potential bias of the HPE methodology in generating accurate HIV prevalence estimates for the districts was assessed using the Bland Altman methodology [104,110]. We found that there was no overall difference between survey and HPE estimates (i.e. average difference was 0.00 units (95% CI: -0.04, 0.04) (Figure 5a). Average difference for males was -0.01 units (95% CI: -0.05, 0.03) while for females was 0.00 units (95% CI: -0.06, 0.06).

While a bias of 0.01 units was observed when assessing the agreement between HPE and survey based estimates for males (Figure 5b), the 95% confidence interval of the difference between the estimates was narrow. Additionally, no systematic pattern was observed as the average of the estimates increased.

The mean difference between the HPE and DHIS2 estimates was 0.01 units (95% CI: -0.05, 0.06) (Figure 6a). Average difference for males was -0.01 units (95% CI: -0.07, 0.06) while for females was 0.02 units (95% CI: -0.05, 0.09), (Figure 6b and 6c respectively).

The bias of the HPE compared to DHIS2 based estimates was more evident from the wider variability of the points about the no-difference (zero) line as the values of the estimates increased (Figure 6a-c). Average bias was 0.02 units and confidence intervals were wider when comparing HPEs and DHIS2 estimates for females (Figure 6c).

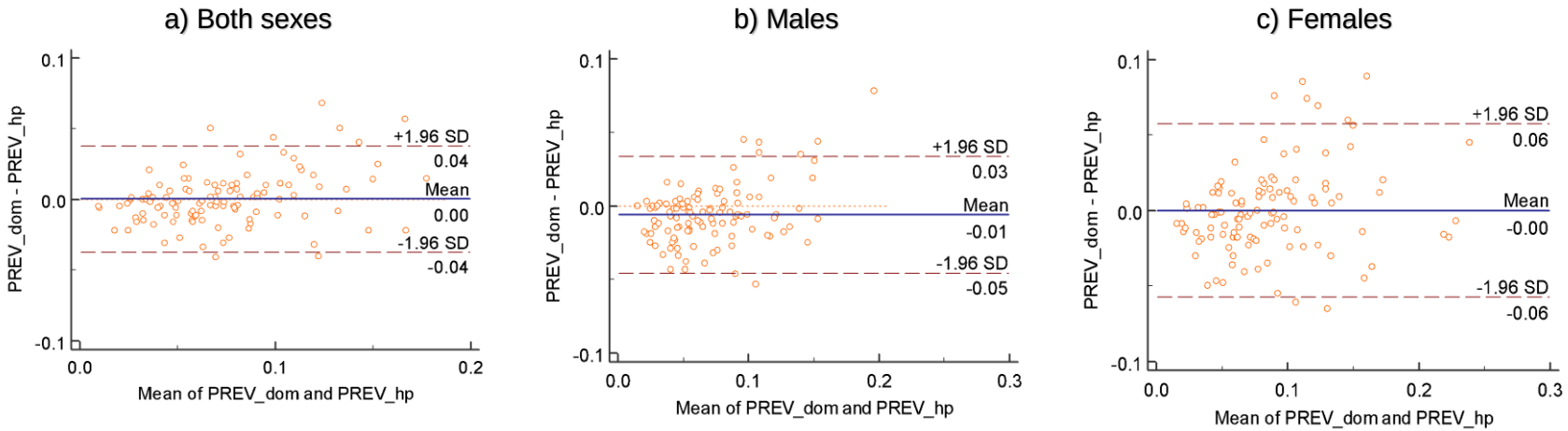


Figure 5: Difference plot comparing HPE and Direct survey estimates

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020;20(1):1–14.

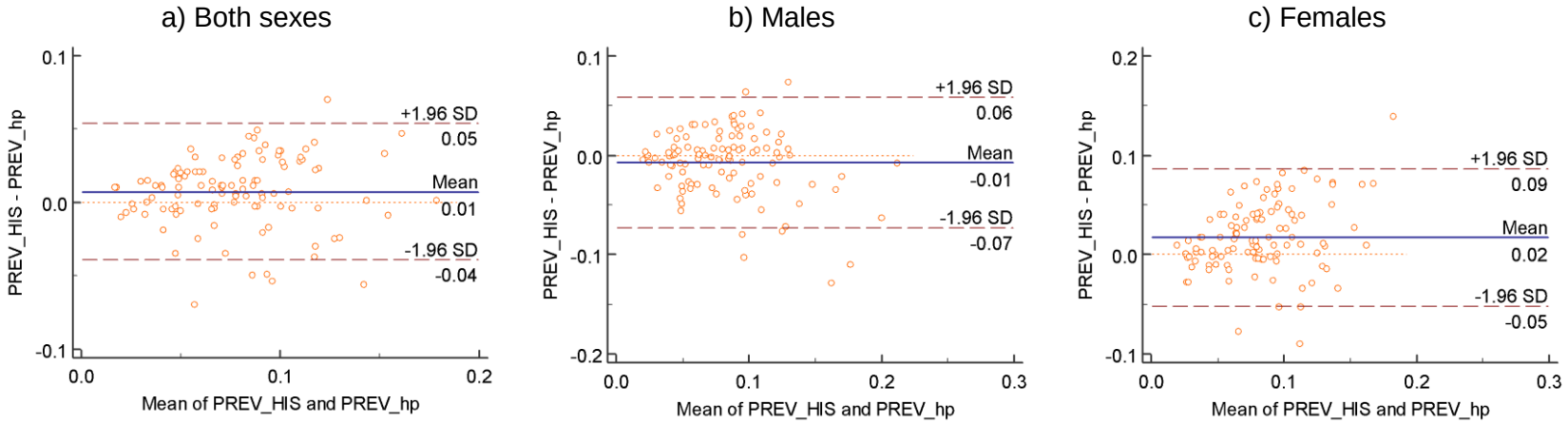


Figure 6: Difference plot comparing HPE and DHIS2 estimates

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020;20(1):1–14.

4.2 HIV Prevalence for districts based on the Small Area Estimation techniques

This objective was concerned with applying SAE techniques to obtain district level HIV prevalence estimates. FH and BHF SAE models were applied to obtain the estimates. The efficiency of the SAE methods in view of the available auxiliary data was also assessed. Auxiliary variables included as predictors was selected via regression modelling. The results of the regression modelling and variable selection were presented published papers attached in appendices V and VI as shown in objective 2. A manuscript addressing this objective was submitted to Plos One and is now published.

4.2.1 HIV prevalence estimates

Table 4 presents a summary of the prevalence estimates obtained using FH, BHF models and the direct survey approaches. It shows that although the estimates were similar across the methods, the direct survey estimate (0.065) had a larger standard error of 0.055 compared to the HF (0.054 (se=0.020)) and the BHF estimates (0.058 (se=0.026)). Results based on the SFH model were very similar to those from the FH model although the standard error of the SFH model was larger than the standard error of the FH model by smaller compared to both standard errors from the BHF and the direct survey model estimate (Table 4). Detailed SFH model results are presented in appendix IV.

Table 4: Summary of HIV Prevalence estimates-pooled results

	Mean (sd)	Minimum	Maximum	First Quartile (Q1)	Second Quartile (Q2)	Third Quartile (Q3)
Direct survey	0.065 (0.055)	0.006	0.464	0.037	0.054	0.083
FH model	0.054 (0.020)	0.007	0.096	0.039	0.055	0.068
BHF model	0.058 (0.026)	0.004	0.142	0.042	0.054	0.074
SFH (Spatial model)	0.054 (0.021)	0.006	0.096	0.039	0.055	0.068

Notes: After omitting the Kalangala district (with outlier HIV prevalence) from the analysis, there was no impact on the model-based estimates. The direct survey estimates changed as follows: Mean (SD) =0.061 (0.036), maximum=0.189 and third quartile=0.082 (other summary results did not change).

Detailed HIV prevalence estimates for all districts are presented in Table 5. Districts with highest and lowest HIV prevalence estimates by region were: Central I: Masaka & Kalangala and Ssembabule; Central II: Luwero and Kyankwanzi; East Central: Busia and Namutumba, Eastern: Katakwi and Butaleja; Karamoja: Abim and Nakapiripirit; Northern: Gulu and Amuru; South Western: Sheema and Kisoro; Western: Kabarole and Bundibugyo and in West Nile: Nebbi and Maracha respectively. Findings were consistent across all methods except for districts where estimates could not be computed due to unavailability of LQAS and population survey data.

Table 5: District HIV prevalence estimates

Region/ District		Direct estimates (95% CI)			FH model estimates (95% CI)			BHF model estimates (95% CI)		
	Survey Sample	Estimate	LCL	UCL	Estimate	LCL	UCL	Estimate	LCL	UCL
Central 1										
Bukomansimbi	116	0.094	0.041	0.147	0.091	0.064	0.119	NA	NA	NA
Butambala	45	0.062	0.000	1.000	0.083	0.054	0.113	NA	NA	NA
Gomba	46	0.104	0.000	1.000	0.060	0.032	0.089	0.072	0.031	0.112
Kalangala	50	0.464	0.000	1.000	0.096	0.065	0.127	NA	NA	NA
Kalungu	86	0.129	0.057	0.200	0.084	0.056	0.112	NA	NA	NA
Lwengo	131	0.106	0.052	0.161	0.077	0.050	0.104	0.092	0.068	0.116
Lyantonde	57	0.089	0.000	1.000	0.075	0.045	0.104	NA	NA	NA
Masaka	176	0.143	0.092	0.194	0.096	0.070	0.122	0.123	0.101	0.146
Mpigi	139	0.127	0.073	0.182	0.083	0.057	0.109	0.102	0.075	0.128
Rakai	307	0.081	0.050	0.111	0.082	0.059	0.104	NA	NA	NA
Ssembabule	135	0.039	0.008	0.070	0.054	0.031	0.076	0.041	0.016	0.066
Wakiso	1384	0.053	0.041	0.064	0.056	0.044	0.067	NA	NA	NA
Central 2										
Buikwe	235	0.081	0.047	0.115	0.069	0.044	0.093	NA	NA	NA
Buvuma	59	0.137	0.000	1.000	0.066	0.037	0.095	NA	NA	NA
Kayunga	233	0.094	0.057	0.132	0.064	0.041	0.088	0.084	0.065	0.102
Kiboga	147	0.066	0.028	0.104	0.073	0.048	0.098	NA	NA	NA
Kyankwanzi	149	0.068	0.029	0.108	0.063	0.038	0.088	NA	NA	NA
Luwero	199	0.091	0.052	0.131	0.075	0.049	0.100	0.082	0.059	0.106

Mityana	233	0.049	0.023	0.076	0.065	0.042	0.087	0.052	0.034	0.071
Mubende	346	0.075	0.047	0.102	0.069	0.048	0.090	0.074	0.058	0.089
Mukono	457	0.062	0.041	0.084	0.066	0.046	0.085	NA	NA	NA
Nakaseke	128	0.108	0.056	0.161	0.070	0.044	0.097	NA	NA	NA
Nakasongola	53	0.050	0.000	0.902	0.076	0.047	0.105	0.049	0.011	0.087
East Central										
Bugiri	292	0.021	0.004	0.037	0.028	0.011	0.044	0.026	0.010	0.042
Busia	324	0.090	0.060	0.120	0.069	0.047	0.091	0.087	0.069	0.104
Buyende	308	0.013	0.002	0.025	0.016	0.005	0.026	NA	NA	NA
Iganga	343	0.028	0.012	0.045	0.034	0.017	0.052	0.033	0.021	0.046
Jinja	494	0.043	0.026	0.059	0.050	0.033	0.066	NA	NA	NA
Kaliro	169	0.041	0.012	0.069	0.034	0.013	0.055	0.051	0.028	0.073
Kamuli	366	0.036	0.018	0.055	0.039	0.022	0.056	0.044	0.030	0.059
Luuka	126	0.057	0.014	0.100	0.048	0.024	0.072	NA	NA	NA
Mayuge	495	0.057	0.036	0.079	0.053	0.036	0.071	0.061	0.051	0.070
Namayingo	202	0.189	0.136	0.242	0.067	0.041	0.094	NA	NA	NA
Namutumba	184	0.018	0.000	0.036	0.026	0.008	0.044	0.027	0.006	0.048
Eastern										
Amuria	333	0.047	0.025	0.070	0.051	0.031	0.071	0.053	0.037	0.069
Budaka	146	0.050	0.016	0.084	0.039	0.016	0.063	0.047	0.024	0.071
Bududa	179	0.049	0.017	0.080	0.039	0.017	0.062	0.044	0.020	0.067
Bukedea	253	0.062	0.034	0.090	0.039	0.016	0.062	NA	NA	NA
Bukwo	110	0.044	0.006	0.083	0.037	0.012	0.061	0.048	0.023	0.073
Bulambuli	177	0.082	0.042	0.123	0.054	0.029	0.078	0.068	0.045	0.091
Butaleja	227	0.017	0.000	0.033	0.016	0.000	0.032	0.020	0.000	0.042
Kaberamaido	279	0.031	0.012	0.050	0.036	0.021	0.052	NA	NA	NA
Kapchorwa	100	0.019	0.000	0.558	0.049	0.020	0.078	0.016	0.000	0.048
Katakwi	222	0.092	0.055	0.128	0.066	0.042	0.089	0.095	0.073	0.117
Kibuku	179	0.033	0.009	0.057	0.032	0.012	0.051	0.037	0.014	0.060
Kumi	336	0.052	0.030	0.075	0.042	0.021	0.063	0.058	0.043	0.074
Kween	130	0.043	0.009	0.076	0.063	0.038	0.088	0.049	0.025	0.074
Manafwa	448	0.024	0.010	0.038	0.030	0.015	0.044	0.027	0.015	0.038
Mbale	874	0.053	0.039	0.068	0.053	0.040	0.067	0.052	0.044	0.060
Ngora	215	0.016	0.000	0.032	0.018	0.004	0.031	NA	NA	NA
Pallisa	337	0.037	0.017	0.057	0.033	0.015	0.051	0.036	0.020	0.052
Serere	418	0.023	0.008	0.037	0.027	0.010	0.043	NA	NA	NA
Sironko	298	0.054	0.028	0.080	0.051	0.029	0.072	0.046	0.030	0.063
Soroti	441	0.037	0.021	0.054	0.042	0.026	0.057	NA	NA	NA

Tororo	685	0.072	0.053	0.091	0.064	0.047	0.082	0.077	0.069	0.086
Kampala	2289	0.069	0.058	0.079	0.069	0.060	0.079	NA	NA	NA
Karamoja										
Abim	102	0.069	0.022	0.115	0.056	0.030	0.081	NA	NA	NA
Amudat	73	0.055	0.000	0.115	0.053	0.026	0.081	NA	NA	NA
Kaabong	231	0.037	0.013	0.061	0.045	0.025	0.064	NA	NA	NA
Kotido	193	0.010	0.000	0.403	0.048	0.019	0.076	0.004	0.000	0.025
Moroto	108	0.017	0.000	0.040	0.019	0.001	0.038	0.020	0.000	0.047
Nakapiripirit	171	0.006	0.000	0.017	0.007	0.000	0.018	NA	NA	NA
Napak	136	0.007	0.000	0.020	0.008	0.000	0.020	NA	NA	NA
Northern										
Agago	191	0.070	0.036	0.105	0.067	0.044	0.090	0.067	0.044	0.090
Alebtong	133	0.092	0.044	0.140	0.058	0.031	0.084	0.078	0.053	0.103
Amolatar	150	0.038	0.008	0.067	0.049	0.030	0.069	0.035	0.011	0.058
Amuru	68	0.037	0.000	0.079	0.044	0.020	0.068	0.037	0.003	0.071
Apac	336	0.086	0.057	0.115	0.069	0.048	0.090	0.089	0.075	0.104
Dokolo	150	0.078	0.037	0.119	0.062	0.037	0.088	0.073	0.049	0.096
Gulu	191	0.118	0.073	0.163	0.080	0.054	0.106	0.093	0.069	0.118
Kitgum	134	0.058	0.021	0.096	0.061	0.038	0.084	0.061	0.035	0.086
Kole	184	0.058	0.026	0.091	0.054	0.031	0.077	0.059	0.036	0.082
Lamwo	101	0.095	0.038	0.151	0.065	0.039	0.092	0.072	0.039	0.105
Lira	226	0.039	0.015	0.063	0.053	0.032	0.074	0.044	0.024	0.064
Nwoya	44	0.127	0.000	1.000	0.062	0.034	0.091	0.083	0.042	0.124
Otuke	87	0.049	0.000	0.893	0.056	0.027	0.084	0.044	0.010	0.078
Oyam	243	0.069	0.038	0.100	0.059	0.036	0.082	0.062	0.042	0.082
Pader	97	0.080	0.026	0.134	0.077	0.050	0.104	0.060	0.033	0.087
South Western										
Buhweju	60	0.031	0.000	0.713	0.053	0.024	0.081	0.017	0.000	0.054
Bushenyi	128	0.056	0.018	0.094	0.067	0.043	0.092	0.055	0.026	0.084
Ibanda	147	0.062	0.025	0.100	0.065	0.040	0.090	0.060	0.033	0.087
Isingiro	166	0.047	0.015	0.079	0.052	0.030	0.075	0.045	0.019	0.071
Kabale	258	0.089	0.054	0.124	0.067	0.043	0.090	0.069	0.051	0.087
Kanungu	199	0.124	0.075	0.172	0.076	0.051	0.102	0.096	0.072	0.120
Kiruhura	169	0.044	0.014	0.075	0.057	0.035	0.080	0.051	0.028	0.073
Kisoro	149	0.036	0.008	0.065	0.029	0.009	0.050	0.033	0.007	0.059
Mbarara	264	0.104	0.068	0.141	0.087	0.062	0.112	0.103	0.085	0.120

Mitooma	113	0.084	0.034	0.134	0.062	0.036	0.088	NA	NA	NA
Ntungamo	225	0.080	0.044	0.115	0.067	0.044	0.091	0.061	0.038	0.084
Rubirizi	57	0.145	0.000	1.000	0.079	0.049	0.108	0.099	0.060	0.139
Rukungiri	181	0.074	0.037	0.112	0.074	0.049	0.099	0.068	0.045	0.091
Sheema	180	0.092	0.051	0.133	0.082	0.056	0.107	0.053	0.000	0.112
West Nile										
Adjumani	186	0.018	0.000	0.039	0.033	0.014	0.053	NA	NA	NA
Arua	1410	0.035	0.025	0.044	0.034	0.024	0.044	0.042	0.037	0.047
Koboko	390	0.023	0.009	0.038	0.027	0.011	0.042	NA	NA	NA
Maracha	245	0.011	0.000	0.022	0.010	0.000	0.021	NA	NA	NA
Moyo	200	0.025	0.005	0.045	0.037	0.020	0.053	NA	NA	NA
Nebbi	587	0.043	0.027	0.059	0.043	0.028	0.058	0.047	0.038	0.057
Yumbe	583	0.024	0.012	0.035	0.023	0.010	0.036	NA	NA	NA
Zombo	384	0.045	0.025	0.065	0.040	0.021	0.058	NA	NA	NA
Western										
Buliisa	50	0.014	0.000	0.483	0.049	0.020	0.078	NA	NA	NA
Bundibugyo	112	0.009	0.000	0.026	0.016	0.001	0.031	NA	NA	NA
Hoima	283	0.062	0.035	0.089	0.059	0.038	0.081	0.062	0.046	0.079
Kabarole	270	0.148	0.107	0.190	0.089	0.063	0.114	0.142	0.126	0.158
Kamwenge	163	0.022	0.001	0.044	0.035	0.017	0.052	0.027	0.005	0.050
Kasese	419	0.019	0.007	0.032	0.023	0.010	0.037	0.023	0.010	0.036
Kibaale	456	0.053	0.033	0.073	0.054	0.037	0.072	NA	NA	NA
Kiryandongo	141	0.049	0.016	0.083	0.050	0.027	0.073	NA	NA	NA
Kyegegwa	162	0.058	0.023	0.094	0.062	0.038	0.085	NA	NA	NA
Kyenjojo	266	0.079	0.048	0.110	0.070	0.046	0.093	0.083	0.062	0.103
Masindi	154	0.043	0.012	0.074	0.055	0.033	0.076	NA	NA	NA
Ntoroko	28	0.133	0.000	1.000	0.078	0.048	0.107	NA	NA	NA

NA- No LQAS survey data available for the district; LCL-Lower Confidence Limit; UCL Upper Confidence Limit

4.2.2 Model diagnostics

To assess bias and reliability of the model based estimates compared to the direct survey estimates, linear regression analysis was conducted regressing model-based estimates against direct survey estimates as shown in Figure 7. Figure 7a presents the regression and the scatter plot of the FH model estimates compared to the direct survey estimate while Figure 7b presents the regression and scatter plot of BHF estimates compared to the directs survey estimates. It shows that the BHF model estimates were more strongly correlated with the direct survey estimates ($\beta_1 = 0.73$, r^2=0.891) than the FH model estimates ($\beta_1 = 0.27$, r^2=0.477).

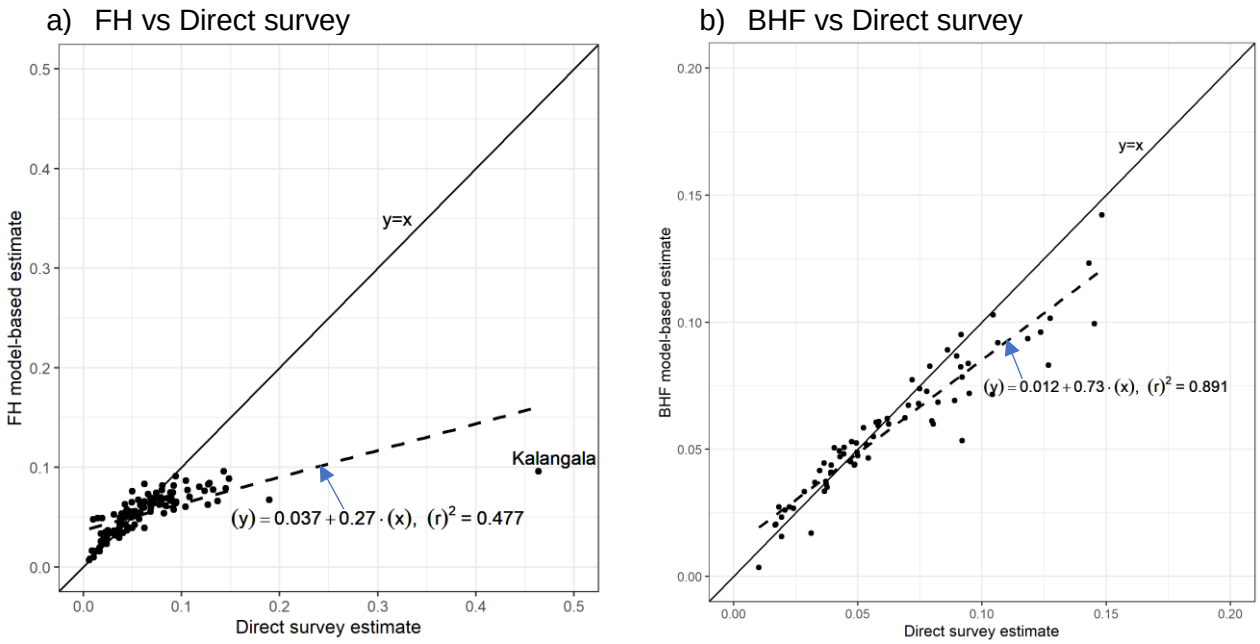


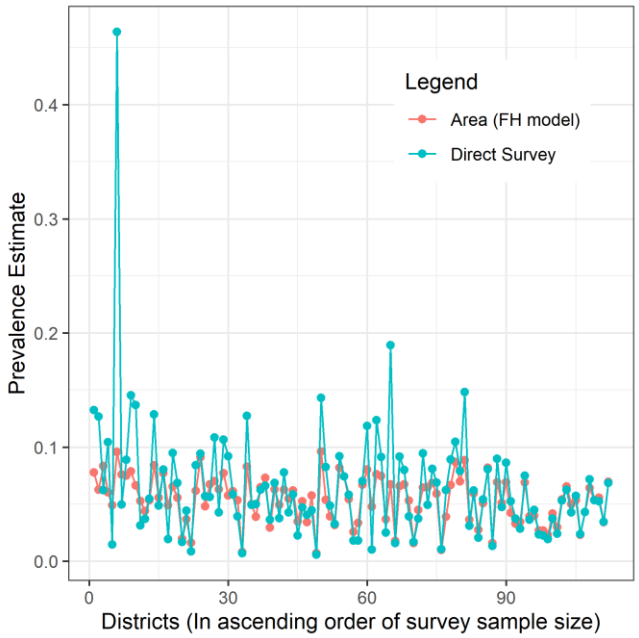
Figure 7: Correlation of Model and Direct Survey estimates
Notes: Regression and scatter plot of a) FH model-based estimates compared to direct survey estimates and b) BHF model-based estimates compared to direct survey estimates.

4.2.3 Precision and consistency of the HIV prevalence estimates

Graphical techniques were used to assess the consistency of the model based estimates compared to the direct survey estimates (Figure 8). It shows that the model-based estimates were similar to the direct survey estimates but had smaller variation around the point estimate compared to the direct survey estimates.

The direct survey point estimates were either higher or lower than model-based point estimates for small survey sample sizes Figure 8, demonstrating the strong shrinkage of the model-based estimates towards the point estimates. Model and survey based estimates were however similar for large district survey sample sizes (Figure 8).

a). HIV prevalence (Direct vs FH model)



b). HIV prevalence (Direct vs BHF model)

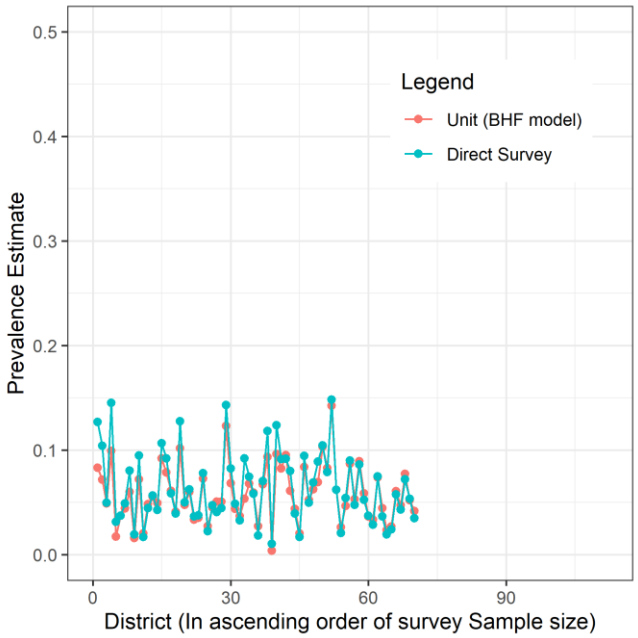


Figure 8: HIV prevalence estimates (Comparing Direct and model based estimates)
Notes: a) direct survey compared to area-level FH model-based estimates (in 111 districts). b) direct survey compared to unit-level BHF model-based estimates (70 districts) BHF model estimates were computed for only districts that had LQAS survey data.

Nwoya district with a survey sample size of 44 individuals had an HIV prevalence estimate of 0.13 (SE=0.68); 0.062 (0.014) and 0.083 (0.021) while Mbale district with a survey sample size of 874 individuals had estimates of 0.053 (0.007); 0.053 (0.007); and 0.052 (0.004) for direct survey, FH and BHF models respectively (Table 5).

Overall improvement in the precision of the estimates, were 37.5% and 33.1% for the BHF and FH model estimates respectively compared to the direct survey estimates.

The coefficient of variation was used to assess variation of the estimates around the point estimates, a summary of the coefficients of variation are presented in Table 6. It shows that coefficient of variation of the direct survey estimates (Mean=136.3, SD=327.7) were generally larger than the coefficient of variation of the FH model estimates (Mean=23.8, SD=11.2) and the BHF model estimates (Mean=28.7, SD=37.9).

Table 6: Summary of Coefficient of variation of HIV prevalence estimates

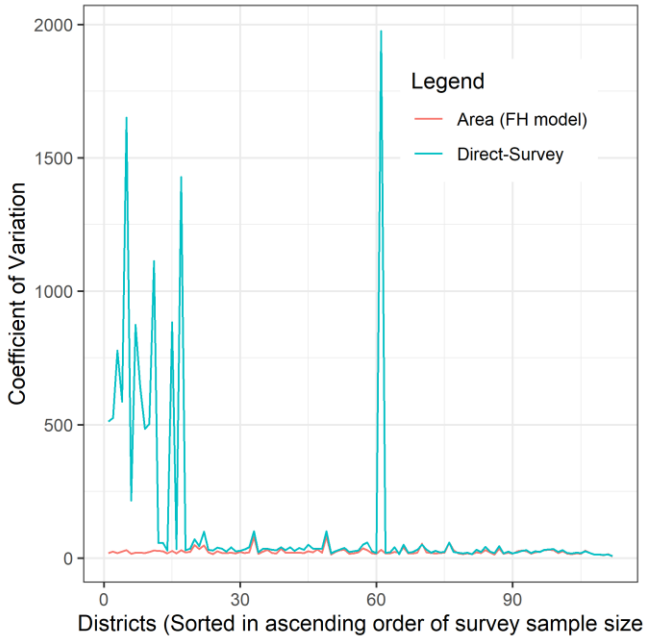
	Mean (SD)	Minimum	Maximum	First Quartile (Q1)	Second Quartile (Q2)	Third Quartile (Q3)
Direct survey	136.3 (327.7)	7.5	1974.8	22.4	30.4	42.8
FH model	23.8 (11.2)	7.1	80.2	17.5	20.2	26.9
BHF model	28.7 (37.9)	5.6	311.9	13.3	20.1	28.7

Notes: After omitting the outlier district (Kalangala) from the analysis, there was no impact on both the model-based and the direct survey estimates (Results not presented).

Graphical techniques were further used to assess the similarities in the coefficients of variation of the model based estimates compared to direct survey based estimates. The results are presented in Figure 9. It shows that direct survey estimates had higher coefficients of variation compared to the model-based estimates irrespective of the survey sample size in the districts (Figure 9).

Assuming that estimates are considered as reliable for decision making if the CV is less than 20% [111]. Only 21 (18%) of the districts would have reliable information for decision making based on direct survey estimates while 53 (47.3%) and 35 (50%) of the districts would have reliable information for decision making based on the FH and BHF models respectively. We also note that the BHF model could be fitted for only 70 districts that had conducted LQAS surveys in 2016.

a). Direct vs area FH model



b). Direct vs unit BHF model

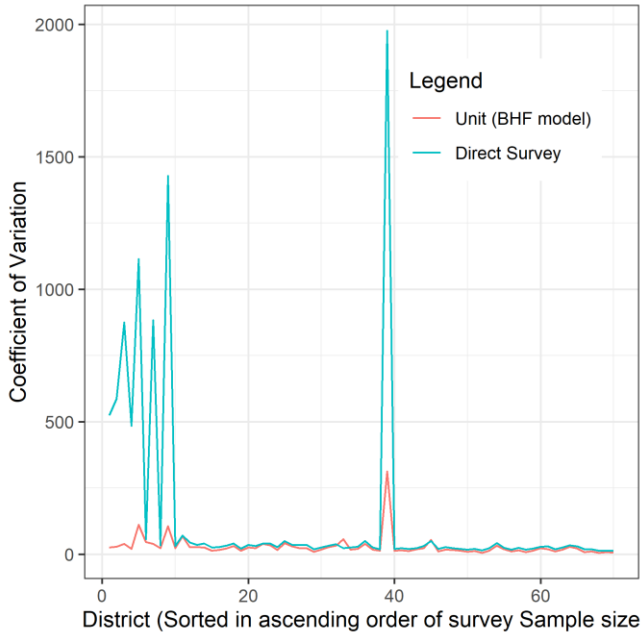


Figure 9: Coefficient of variation of direct survey and model-based estimates with the districts sorted in ascending order of the survey sample sizes

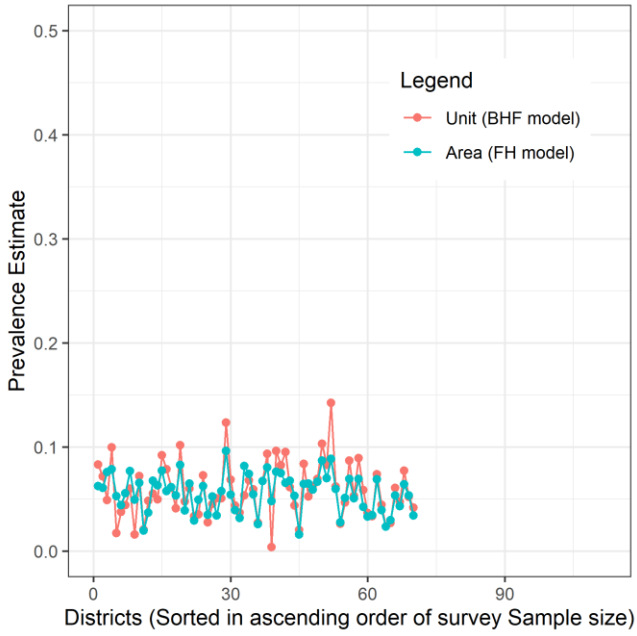
Notes: a) direct survey compared to area-level FH model-based CVs (in 111 districts), b) direct survey compared to unit-level BHF model-based CV (70 districts). BHF model estimates were computed for only districts that had LQAS survey data.

4.2.4 Comparison of the model-based estimates (FH and BHF)

Assessing consistency of the BHF model estimates compared to the FH model estimates, the findings is presented in Figure 10. It shows that the unit level model estimates of HIV prevalence in districts had a larger variation around the point estimate compared to the area-level model estimates (Figure 10a). This is further demonstrated by the comparison of the coefficient variation presented in Figure 10b.

The coefficient of variation of the BHF model estimates were consistently larger for districts with small survey sample sizes and consistently smaller for districts with larger survey sample sizes implying that the BHF model estimates appears to converge more rapidly with increase in survey sample size to the point estimate than the FH model estimates.

a). HIV prevalence (FH vs BHF model)



b). Coefficient of variation

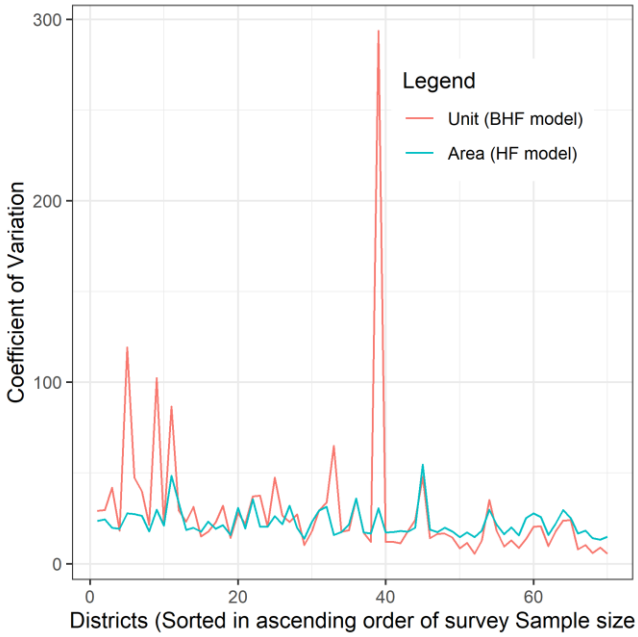


Figure 10: Area-level and unit-level model-based estimates of district HIV prevalence
Notes: HIV prevalence estimates (on the left (a)) and the coefficient of variation (on the right (b)) area and unit-level model-based estimates with the districts sorted in ascending order of the survey sample sizes.

4.3 Factors associated with inter and intra-district variation in adult HIV prevalence in Uganda

This objective assessed the factors associated with HIV positivity. Factors associated with HIV positivity by venue of testing were also explored. Results of this objective have been published [34].

4.3.1 Socio-demographic characteristics of study participants

Overall, there was no difference in the composition of the individuals who had tested in a facility compared to those who had tested in a community setting except by gender (Table 7). Facility testers comprised more females (70.9%), while community testers had more males (54.7%).

Table 7: Socio-demographic characteristics of study participants

Characteristic	Overall n=11,685 (%)	Tested in Health Facility n=8,978 (%)	Tested in Community n=2,707 (%)
Gender			
Male	4,038 (34.9)	2,582 (29.1)	1,456 (54.7)
Female	7,647 (65.1)	6,396 (70.9)	1,251 (45.3)
Age			
15-19	1,498 (12.8)	953 (10.8)	545 (19.5)
20-29	4,789 (41.3)	3,894 (43.5)	895 (33.9)
30-39	3,422 (29.3)	2,716 (30.3)	706 (26.1)
40-49	1,976 (16.6)	1,415 (15.4)	561 (20.5)
Education Level			
No Education	1,216 (9.9)	1,007 (10.6)	209 (7.7)
Primary	6,406 (55.1)	5,003 (55.9)	1,403 (52.2)
Secondary+	4,063 (35.0)	2,968 (33.5)	1,095 (40.1)
Marital status			
Never married	2,159 (18.6)	1,301 (14.8)	858 (31.3)
Married/Cohabiting	8,169 (70.3)	6,609 (73.7)	1,560 (59.0)
Previously married	1,357 (11.1)	1,068 (11.5)	289 (9.8)
Number of sexual partners in 12 months preceding survey			
0	1,957 (16.2)	1,252 (13.7)	705 (24.5)
1	8,649 (74.1)	6,980 (77.4)	1,669 (62.8)
2+	1,079 (9.7)	746 (8.9)	333 (12.8)
Currently working			
No	2,894 (24.2)	2,219 (24.3)	675 (23.8)
Yes	8,791 (75.8)	6,759 (75.7)	2,032 (76.2)
Distance to nearest HF			
<2	2,441 (21.7)	1,920 (22.2)	521 (20.0)
2-5	4,823 (40.7)	3,696 (40.6)	1,127 (40.8)
5+	4,009 (33.7)	3,040 (33.3)	969 (35.2)
Don’t Know	412 (3.9)	322 (3.9)	90 (3.9)
Area of residence			
Rural	8,870 (75.8)	6,845 (76.0)	2,025 (75.4)
Urban	2,815 (24.1)	2,133 (24.0)	682 (24.6)
Region			
Central 1	1,151 (12.2)	889 (12.3)	262 (12.2)
Central 2	1,177 (10.6)	861 (10.0)	316 (12.8)
Kampala	1,464 (9.0)	1,068 (8.4)	396 (11.0)
East Central	1,013 (8.8)	722 (8.2)	291 (10.5)
Mid-Eastern	901 (7.6)	732 (8.0)	169 (6.3)
North East	1,134 (9.2)	875 (9.0)	259 (10.0)
West Nile	1,220 (6.4)	896 (6.0)	324 (7.6)
Mid Northern	1,394 (12.5)	1,107 (12.9)	287 (11.3)
South Western	1,024 (11.3)	853 (12.2)	171 (8.0)
Mid-Western	1,207 (12.4)	975 (12.9)	232 (10.5)

Note: Percentages are weighted by population survey weights

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Levin J. Difference in HIV prevalence by testing venue : results from population level survey in Uganda. AIDS Care [Internet]. 2020;0(0):1–12. Available from: https://doi.org/10.1080/09540121.2020.1734179

4.3.2 Factors associated with HIV positivity by venue of testing

Overall (both testing venues combined), HIV positivity was significantly higher among females compared to males, among older age groups compared to those aged 20-29 years, among individuals previously married compared to those who were married at the time of the survey, those with 2 or more sexual partners compared to those with one sexual partner (Table 8). HIV positivity was significantly lower among individuals aged 15-19 years compared to those aged 20-19 years, those with no education or those with secondary level of education compared to those with primary level of education, those who had never married compared to those who were married at the time of the survey and among those living in rural areas compared to those resident in urban areas.

Among those tested in a facility setting, HIV positivity was significantly higher among older age groups compared to those aged 20-29 years, previously married compared to currently married and among those who reported no sexual partner compared to those who reported one sexual partner at the time of the survey. HIV positivity was significantly lower if individuals were female compared to males, aged 15-19 years compared to those aged 20-29 years and among those living in rural areas compared to those resident in urban areas.

For those tested in a community setting, HIV positivity was significantly higher among females compared to males, among older age groups (30+ years) compared to those aged 20-29 years, among previously married compared to currently married and among those with 2 or more sexual partners compared to those with one sexual partner. HIV positivity was significantly lower among those aged 15-19 years compared to 20-29 years.

Table 8: Factors associated with HIV prevalence by venue of testing

	Overall				Tested in Health facility		Tested in Community	
Characteristic	**cOR**	**(95% CI)**	**aOR**	**(95% CI)**	**aOR**	**(95% CI)**	**aOR**	**(95% CI)**
Gender (Reference group: Male)								
Female	1.25	(1.09, 1.44)*	1.21	(1.03, 1.43)*	0.78	(0.62, 0.97)*	1.62	(1.21, 2.15)*
Age (Reference group: 20-29)								
15-19	0.48	(0.35, 0.66)*	0.53	(0.38, 0.75)*	0.67	(0.45, 0.99)**	0.34	(0.17, 0.70)*
30-39	1.86	(1.58, 2.18)*	1.70	(1.43, 2.03)*	1.96	(1.60, 2.42)*	1.24	(0.89, 1.74)
40-49	2.28	(1.90, 2.72)*	1.89	(1.55, 2.31)*	2.92	(2.28, 3.73)*	1.17	(0.80, 1.69)
Education Level ((Reference group: Primary)								
No Education	1.09	(0.86, 1.38)	0.86	(0.68, 1.10)	0.91	(0.69, 1.19)	0.70	(0.41, 1.19)
Secondary	0.61	(0.52, 0.72)*	0.72	(0.60, 0.86)*	0.64	(0.52, 0.80)*	0.85	(0.61, 1.18)
Marital status (Reference Group: Married/Living together								
Never married	0.49	(0.39, 0.63)*	0.80	(0.58, 1.1)	0.93	(0.64, 1.37)	0.74	(0.43, 1.27)
Previously married	3.33	(2.81, 3.96)*	2.76	(2.27, 3.35)*	3.25	(2.66, 4.12)*	2.05	(1.44, 2.92)*
Number of sexual partners in 12 months preceding survey (Reference Group: 1)								
0	1.52	(1.27, 1.81)*	1.32	(1.07, 1.63)*	1.55	(1.20, 2.00)*	1.01	(0.67, 1.53)
2+	1.39	(1.11, 1.74)*	1.43	(1.12, 1.82)*	1.07	(0.77, 1.50)	1.95	(1.37, 2.76)*
Currently working (Reference Group: Not employed)								
Employed	1.34	(1.13, 1.61)*	1.02	(0.84, 1.24)	1.03	(0.83, 1.28)	0.99	(0.65, 1.50)
Distance to nearest HF (Reference Group: <2Km)								
2-5	0.80	(0.66, 0.98)**						
5+	1.02	(0.83, 1.25)						
Don't Know	0.79	(0.53, 1.19)						
Area of residence (Reference Group: Urban)								
Rural	0.84	(0.69, 1.03)	0.62	(0.47, 0.82)*	0.59	(0.43, 0.80)*	0.75	(0.46, 1.23)
Region (Reference Group: Central 1)								
Central 2	0.75	(0.56, 1.02)	0.74	(0.53, 1.02)	0.75	(0.51, 1.10)	0.71	(0.41, 1.21)
Kampala	0.60	(0.43, 0.84)*	0.49	(0.32, 0.75)*	0.53	(0.33, 0.86)*	0.47	(0.23, 0.93)*
East Central	0.45	(0.31, 0.66)*	0.46	(0.31, 0.66)*	0.58	(0.37, 0.90)*	0.30	(0.15, 0.60)*
Mid-Eastern	0.40	(0.25, 0.64)*	0.44	(0.27, 0.71)*	0.37	(0.23, 0.62)*	0.53	(0.24, 1.14)
North East	0.54	(0.38, 0.77)*	0.58	(0.40, 0.82)*	0.56	(0.37, 0.84)*	0.53	(0.30, 0.95)*
West Nile	0.38	(0.26, 0.56)*	0.36	(0.23, 0.56)*	0.38	(0.22, 0.64)*	0.33	(0.18, 0.63)*
Mid Northern	0.72	(0.52, 1.01)	0.80	(0.56, 1.13)	0.78	(0.53, 1.16)	0.59	(0.33, 1.11)
South western	0.82	(0.60, 1.14)	0.87	(0.62, 1.22)	0.93	(0.62, 1.38)	0.65	(0.38, 1.14)
Mid-western	0.73	(0.53, 102)	0.78	(0.55, 1.11)	0.82	(0.54, 1.24)	0.69	(0.40, 1.17)

**p-value<0.01, **P-value<0.05, cOR- unadjusted Odds Ratio, aOR-Adjusted Odds Ratio*

Source: Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Levin J. Difference in HIV prevalence by testing venue : results from population level survey in Uganda. AIDS Care [Internet]. 2020;0(0):1–12. Available from: https://doi.org/10.1080/09540121.2020.1734179

4.4 Change in adult HIV prevalence between 2011 and 2016

This objective was concerned with evaluating change in adult HIV prevalence between 2011 and 2016 and the contribution of the characteristics and coefficient of the characteristics to change in HIV prevalence between the surveys. Results of this objective are presented in a manuscript still in draft, yet to be submitted for publication.

4.4.1 Socio-demographic characteristics

Table 9 presents the socio-demographic characteristics of the survey participants in the two surveys. It shows that 10.2% of the survey participants had no formal education and 6.1% had tertiary level of education in 2011 while in 2016, only 6.0% had no formal education and 11.4% had tertiary level of education. Similarly, 25% and 10.0% had zero and two or more sexual partners respectively in 2011 while in 2016, 12.6% had zero and 16.2% had two or more sexual partners. The proportion of respondents resident in urban areas, those with one and those with two or more sexual partners increased in the 2016 survey compared to 2011 (Table 9).

Table 9: Background characteristics (weighted) of study participants

Characteristic	Survey period	
Age (Years)	**% 2011 (n=19,475)**	**% 2016 (n=25,420)**
15-19	22.8	25.9
20-24	17.9	21.0
25-29	16.6	16.3
30-34	13.2	12.8
35-39	12.5	9.9
40-49	17.0	14.0
Gender		
Male	44.4	47.5
Female	55.6	52.5
Education level		
None	10.2	6.0
Primary	58.2	55.4
Secondary	25.5	27.2
Tertiary+	6.1	11.4
Marital Status		
Never Married	29.6	35.7
Married	60.7	52.4
Widowed	2.3	1.7
Divorced/Separated	7.4	10.3
Number of sexual partners		
0	25.6	12.6
1	64.4	71.2
2+	10.0	16.2
Residence		
Rural	79.5	70.3
Urban	20.5	29.7
Region		
Central 1	11.2	14.0
Central 2	10.3	12.0
Kampala	7.8	6.6
East Central	10.4	9.8
Eastern	10.5	9.5
North East	8.1	7.4
West Nile	6.3	7.2
North Central	10.3	9.9
South West	12.0	11.0
Western	13.2	12.6

4.4.2 Trends in HIV prevalence between the two time points

Table 10 presents the trends in HIV prevalence by selected socio-demographic characteristics. It shows that HIV prevalence dropped by 1.3 percentage points between 2011 and 2016. A greater reduction in HIV prevalence was observed among Males (1.8); individuals aged 20-24 years (2.0); individuals with primary level of education (1.4); individuals with two or more sexual partners in the 12 months preceding the survey (2.7); and individuals who were divorced/separated (3.5).

Other population sub groups with higher than average reduction in HIV prevalence between the two surveys were those resident in rural areas (1.6); age groups 30-34 years and regions (Central 1(3.0), Central 2 (1.6), North East (1.9), West Nile (2.0), and Western (2.8)). The lowest reduction in HIV prevalence between the two time points was observed among individuals aged 35-39 years (0.4) and those who were married (0.8). An increase in HIV prevalence between the 2011 and 2016 was observed among individuals with no formal education (0.3), individuals with no sexual partners in the 12 months preceding the survey (0.3) and among individuals aged 40-49 years (1.6).

Table 10: Trends in Weighted HIV prevalence between 2011 and 2016

Characteristic	HIV prevalence (weighted)		Percentage point difference in HIV prevalence
	UAIS, 2011	UPHIA, 2016	
Over all	**7.3**	**6.0**	**-1.3**
Age (Years)			
15-19	2.4	1.2	-1.2
20-24	5.3	3.3	-2.0
25-29	7.4	6.3	-1.1
30-34	10.2	8.6	-1.6
35-39	11.6	11.2	-0.4
40-49	10.7	12.3	1.6
Gender			
Male	6.1	4.3	-1.8
Female	8.3	7.5	-0.8
Education level			
None	9.2	9.5	0.3
Primary	8.0	6.6	-1.4
Secondary	5.8	4.8	-1.0
Tertiary+	4.4	3.6	-0.8
Marital Status			
Never Married	2.8	1.6	-1.2
Married	7.4	6.6	-0.8
Widowed	32.6	31.9	-0.7
Divorced/Separated	16.9	13.4	-3.5
# of sexual partners			
0	6.3	6.6	0.3
1	7.5	6.6	-0.9
2+	9.2	6.5	-2.7
Residence			
Rural	7.0	5.4	-1.6
Urban	8.7	7.2	-1.5
Region			
Central 1	10.6	7.6	-3.0
Central 2	9.0	7.4	-1.6
Kampala	7.1	6.6	-0.5
East Central	5.7	4.4	-1.3
Eastern	4.1	4.8	0.7
North East	5.3	3.4	-1.9
West Nile	4.8	2.8	-2.0
North Central	8.3	7.0	-1.3
South West	8.0	7.7	-0.3
Western	8.3	5.5	-2.8

4.4.3 Factors associated with HIV positivity

Pooled multivariable logistic regression was applied to assess the factors associated with HIV positivity across the surveys and the results are presented in Table 11. It shows that there was a significant reduction in HIV prevalence between 2011 and 2016 surveys (aOR=. 0.84, 95% CI: 0.76, 0.92). That is, after adjusting for other factors included in the model, the odds of being HIV-positive were 16% lower in 2016 compared to 2011.

The odds of HIV positivity among the age group 15-19 years was significantly lower compared to the age group 20-24 years while age groups 25+ had significantly higher odds of testing HIV positive compared to the age group 20-24 years (Table 11).

Other characteristics with higher odds of testing HIV positive were females compared to males (aOR=1.37, 95% CI: 1.20, 1.55), divorced/separated aOR=2.27, 95% CI: 1.99, 2.59) and widowed aOR=4.62, 95% CI: 3.84, 5.56) compared to those who were married at the time of the survey. Those who had two or more sexual partners compared to those who had one (aOR=1.29, 95% CI: 1.14, 1.46) and those who were resident in urban areas compared to those resident in rural (aOR=1.58, 95% CI: 1.28, 1.94).

Significantly lower odds of testing HIV positive was among individuals with secondary or higher level of education compared to those with primary level of education and among those never married compared to those who were married at the time of the survey (Table 11).

Table 11: Logistic regression analysis of factors associated with HIV positivity

Characteristic	HIV prevalence (weighted)	
	aOR (95%CI)	
Survey year		
2011	1.00	
2016	0.84	(0.76, 0.92)***
Age (Years)		
15-19	0.54	(0.44, 0.68)***
20-24	1.00	
25-29	1.54	(1.30, 1.83)***
30-34	2.06	(1.72, 2.46)***
35-39	2.41	(2.01, 2.88)***
40-49	2.34	(1.98, 2.75)***
Gender		
Male	1.00	
Female	1.37	(1.20, 1.55)***
Education level		
None	0.85	(0.72, 0.99)
Primary	1.00	
Secondary	0.85	(0.74, 0.99)*
Tertiary+	0.51	(0.41, 0.62)***
Marital Status		
Never Married	0.82	(0.69, 0.97)*
Married	1.00	
Widowed	4.62	(3.84, 5.56)***
Divorced/Separated	2.27	(1.99, 2.59)***
Number of sexual partners		
0	0.96	(0.85, 1.08)
1	1.00	
2+	1.29	(1.14, 1.46)***
Residence		
Rural	1.00	
Urban	1.58	(1.28, 1.94)***
Region		
Central 1	1.00	
Central 2	0.89	(0.66, 1.22)
Kampala	0.68	(0.46, 1.00)
East Central	0.56	(0.36, 0.86)**
Eastern	0.53	(0.37, 0.75)***
North East	0.54	(0.37, 0.79)**
West Nile	0.39	(0.28, 0.55)***
North Central	0.92	(0.67, 1.26)
South West	0.90	(0.64, 1.28)
Western	0.79	(0.48, 1.28)

* = significant at 0.05; ** = significant at 0.01; *** = Significant at <0.001

4.4.4 Components of Change in HIV Prevalence between 2011 and 2016

Table 12 presents results from the multivariable decomposition logistic regression analysis. Results show that, although change in the population characteristics led to a (54%) decline in the likelihood of testing HIV positive between 2011 and 2016, this was offset by the change in the effects (coefficients) of the characteristics.

The change in effects/coefficients contributed to 154% change in the likelihood of testing HIV positive between the survey points. A larger percentage of the positive change in the effects/coefficients attributable to the constant term in the model (421%) – which is an unexplained portion of the differentials in HIV prevalence between the two time points.

The unexplainable portion of the differential include the possible effect of the unobservable variables not included in the model. This implies that a large portion of the change in HIV prevalence is explained by variables not included in the model. Ignoring the constant would results in 321.4% contribution to the decline in HIV prevalence between the two time points.

It can be noted that although the changes in the effects/coefficients contributed to a larger reduction in HIV prevalence, the contribution of most of the variables was not statistically significant (Table 12).

The 0.8% point reduction in HIV prevalence among females contributed to 67% of the overall reduction in HIV prevalence between the two time points. Over 60% of the reduction was due to the coefficient effects while only 7% was due to differences in the characteristics of women in 2016/17 compared to 2011/12 (Table 12).

Table 12: Decomposition changes in HIV prevalence for the period 2011 to 2016

Characteristic	Due to differences in Characteristics (E)		Due to differences in Coefficients (C)		Total amount attributable
	Coefficient	Percent	Coefficient	Percent	
Age (Years)					
15-19	0.00216***	-34.296	-0.000792	12.572	-21.724
20-24	0.00		0.00		
25-29	0.000459***	-7.2861	0.0011549	-18.315	-25.601
30-34	0.000655***	-10.393	0.0010974	-17.402	-27.795
35-39	-0.000379***	6.0247	0.0018557*	-29.428	-23.403
40-49	0.000145***	-2.3027	0.0037359**	-59.245	-61.548
Gender					
Male	0.00		0.00		
Female	0.000453***	-7.1859	0.0038112	-60.438	-67.624
Education level					
None	0.000124	-1.9685	0.00038862	-6.1628	-8.131
Primary	0.00		0.00		
Secondary	0.0000563	-0.89271	0.0011594	-18.386	-19.279
Tertiary+	-0.0014889***	23.611	0.00022552	-3.5763	20.035
Marital Status					
Never Married	0.000403	-6.3837	0.0014417	-22.862	-29.246
Married	0.00		0.00		
Widowed	-0.000132***	2.1033	-0.0000793	1.2586	3.362
Divorced/Separated	0.001613***	-25.592	-0.00013664	2.1669	-23.425
# of sexual partners					
0	0.000532	-8.4421	-0.0012578	19.947	11.505
1	0.00		0.00		
2+	0.000465*	-7.3845	-.00058452	9.2694	1.885
Residence					
Rural	0.00		0.00		
Urban	0.001482***	-23.507	-.0010425	16.532	-6.975
Region					
Central 1	0.00		0.00		
Central 2	-1.1422e-06	0.01811	0.0008337	-13.222	-13.204
Kampala	0.0003067*	-4.8554	0.0012513	-19.844	-24.699
East Central	0.0000431	-0.6836	0.0008742	-13.864	-14.548
Eastern	-0.0009428***	14.952	0.0016567	-26.271	-11.319
North East	-0.0011876***	18.834	-0.0002359	3.7419	22.576
West Nile	-0.0015796***	25.049	0.0000610	-.96834	24.081
North Central	-0.0000336	0.53415	0.000918	-14.558	-14.024
South West	0.0000111	-.1762	0.00069642	-11.044	-11.220
Western	0.0002785*	-4.4175	-0.00021104	3.3467	-1.071
Subtotal	0.0034456***	-54.64	-0.009751***	-266.751	-321.392
Constant			-0.026573	421.39	
Total		-54.64	-0.009751***	154.64	100.000

* = significant at 0.05; ** = significant at 0.01; *** = significant at <0.001

5.0 CHAPTER FIVE: DISCUSSION

5.1 Introduction

HIV/AIDS indicator Information is needed for critical decision making including resource allocation and to evaluate the impact of interventions implemented to avert or minimise further spread of the HIV virus. The granularity of available information and its precision however, determine the level of details available for policy makers and program managers to make decisions as well as the confidence with which they make the decisions. The more granular and precise the information is, the greater the level of detail that is available. This study aimed at generating more precise HIV prevalence estimates at district level in Uganda for HIV/AIDS disease monitoring using advanced statistical techniques.

The HPE and SAE methods were applied to obtain the HIV prevalence estimates for districts. The methods were applied to population survey data (2011 and 2016); routine health facility data; LQAS data and; census data to obtain the estimates. To apply the HPE methodology, health facility data was adjusted based using the population survey design for its selection bias. Health facility data comprise individuals who access health facilities, facilities that report to the national DHIS2 and those in urban and easily accessible location and therefore may not represent the general population. A multivariable logistic regression model was applied to population survey data to establish the factors associated with HIV positivity while the Blinder-Oaxaca decomposition method was used to display the contribution of survey participants' characteristics and the coefficients of the characteristics to the change in HIV prevalence between 2011 and 2016.

A discussion of the findings of this study, implications for existing related literature and future research are presented under the following subheadings: i) Comparability of routine

health facility HIV testing and population survey data; ii) District HIV prevalence estimates based on: (a) HPE, (b) SAE methodologies and; iii) Contributing factors to change in HIV prevalence between 2011 and 2016.

5.2 Comparability of Routine Health Facility and Population Survey data

Population surveys and health facility HIV testing data are often used complimentarily to monitor the HIV/AIDS epidemic. Due to their inherent limitations, adjustment methods are applied to combine the data sets to obtain HIV/AIDS indicator estimates including HIV prevalence. Combining the datasets however, requires understanding the differences or comparability of the data sources.

This study computed HIV prevalence ratio among individuals tested in a health facility compared to those tested in a community setting. Findings show that HIV prevalence among health facility testers was 1.8 times higher compared to those tested in a community setting. Prevalence ratio was more than 2-fold higher among males; individuals aged 15-19 and 40-49 years; never married; and those who reported no sexual partners in the 12 months preceding the survey. Other population subgroups who had more than two fold HIV prevalence ratio were 30-39 years, previously married, no education and urban residence. In terms of socio-demographic characteristics, none of the population subgroups tested in a health facility had lower HIV prevalence compared to those tested in a community setting.

These findings are consistent with findings from many studies in Sub-Saharan Africa that have found higher HIV prevalence among individuals tested in health facilities compared to those tested in community setting [112–116]. Earlier studies attributed higher HIV prevalence among health facility testers to: ill health [112,115–117]; recent risky sexual

behaviour [117] among individuals seeking health care; and fear of health workers not respecting privacy and confidentiality of personal information [118,119]. For example, the higher prevalence among males tested in health facilities compared to females may be attributed to the higher proportion of the women tested in a health facility being those seeking antenatal services, who we would not expect to differ from the community testers in terms of HIV prevalence. While many males would be at the health facility for reasons of ill-health, some of which would be HIV-related.

The differences in HIV prevalence estimates among individuals tested in a health facility compared to those tested in a community setting therefore implies that to combine the datasets, the methodology applied needs to take into account these differences. Zaba et al proposed use of relative HIV prevalence ratios based on fertile and infertile women to adjust HIV prevalence estimates for women from general population surveys [41]. Other methods proposed in literature include adjusting for the differences in age specific fertility by HIV sero-status [57] and adjusting for differences in HIV prevalence by fertility risk category and parity [57]. Use of age specific fertility however led to underestimation of HIV prevalence among women in the general population in an earlier study [57]. This was attributed to the fact that sentinel sites for ANC data excluded health facilities offering only family planning services. To overcome these limitations, inclusion of facilities offering family planning services as part of the sentinel surveillance sites is recommended [57].

Other approaches used to obtain HIV prevalence, such as the SPECTRUM software package, are continuously updated incorporating emerging information on HIV/AIDs dynamics from participating countries [18,66]. Recent updates include use of PMTCT data to adjust for indicator estimates and matching it with previous estimates that only used ANC data to adjust the HIV/AIDS indicator estimates [120]. The software provides age

and gender adjusted estimates [17,18]. The major limitation of the SPECTRUM software however is the comparability of estimates with those produced in the previous frameworks/years especially before the inputs changed.

Findings from this study support taking into account the socio-demographic differences in HIV prevalence among individuals tested in a facility setting. That is, in addition to age and gender differences, sexual activity/experiences including the number of concurrent sexual partners and marital status need to be incorporated in the models to obtain more reliable HIV prevalence estimates.

5.3 Application of the Hybrid Prevalence Estimation Methodology

In this study, the HPE methodology was applied to the 2011 UAIS and the DHIS2 data to obtain district HIV prevalence estimates. The HPE methodology uses a weighting factor to combine a small population survey sample with routine health facility data to obtain a point estimate and the 95% confidence interval [10,23,78]. The methodology is applicable to overcome the limitations and biases associated with using population survey and DHIS2 data independently to obtain estimates for decision making at district levels.

Propensity to test for HIV generated using multivariable logistic regression was used to combine the population survey and DHIS2 prevalence data [23]. To combine the datasets, DHIS2 data was adjusted using the population survey design to obtain larger denominators for each district [23]. The comparability of the estimates with population survey and DHIS2 data was assessed using the Bland Altman analysis [104,110].

The estimates obtained from the HPE methodology were similar and consistent with survey-based estimates but had narrower confidence intervals compared to estimates from population survey and DHIS2 data [23]. Overall, standard errors of the district

prevalence estimates were reduced by 29% compared to standard errors of the direct population survey estimate. Application of the HPE methodology to obtain immunization coverage for districts that utilized LQAS surveys in Benin and Madagascar found up to 6% reduction in the standard errors [78]. These findings imply that the methodology can be used to provide policy makers and program managers with more reliable and efficient estimates for decision-making.

The HPE methodology combines a small survey sample with routinely collected service delivery data to obtain indicator estimates. Low resourced settings have challenges of conducting large and regular population surveys to obtain data for planning and decision-making, especially at lower administrative levels [121–123]. Applying this methodology therefore will bridge this gap of unavailability of information amidst scarce resources.

The methodology uses three parameters: 1) HIV prevalence among individuals tested in a health facility; 2) Prevalence among those tested in a community setting during the population surveys and 3) an annealing factor, propensity to test for HIV in a health facility, all of which may vary from district to district. The knowledge and variation of these quantities given the available data, determines the quality and variance associated with the estimates and hence the efficiency of the estimates as explained by Hedt et al [10]. For example, for districts with high coverage of HIV testing in a health facility, HIV prevalence from the population survey had no impact on the HPE estimates whereas for those with lower health facility HIV testing coverage, HPEs were dependent and very similar to population survey estimates irrespective of the survey sample size.

To generate the propensity to test for HIV in a health facility, multi-level logistic regression model was applied. The model incorporates individual and district level contextual effects as well as any correlation effects [124–126]. This ensures that both individual and district

level contextual factors for HIV variation are included in the model. Studies elsewhere have found neighbourhood as well as contextual level factors, including the level of urbanisation of the area, to influence HIV distribution [80,127,128]. It is also important to note that factors that influence health facility access such as education level and ill health are also risk factors for HIV positivity [129–132].

Furthermore, to obtain the HPEs, the denominators for health facility testing data were adjusted based on the design of the population survey data. Other studies have used population census [126] and data from routine services such as child immunization which have a near universal coverage to define denominators for population level parameter estimates [133]. Population censuses are however, conducted after ten years and therefore may not reflect HIV dynamics in the period between the censuses. The poor documentation and reporting of routine data in low resource settings implies that routine service delivery data even when coverage appears to be universal may not be appropriate for computation of denominators even for interventions with near universal coverage.

In summary, this study presents a robust, simple and easy to apply annealing methodology to obtain more precise and efficient HIV prevalence estimates for districts in Uganda. The procedure is applied at no additional data collection costs other than the effort to clean and manage the data. The methodology can therefore be applied to combine population survey data such as the DHS and routine service delivery data to obtain health indicator estimates for other social services.

5.4 Application of the Small Area Estimation methodology

The SAE methodology was applied to obtain HIV prevalence estimates for districts using auxiliary information from antenatal HIV testing and community LQAS surveys. The FH

model was applied to obtain HIV prevalence estimates for districts using antenatal HIV prevalence as a predictor, while the BHF model was applied to obtain HIV prevalence estimates, using community LQAS survey data as auxiliary predictors. The SAE techniques borrow information from neighbouring areas through linking models [134–137].

Overall, the FH model and BHF Model estimates had average standard errors reduced by 37.5% and 33.1% respectively. The estimates were similar to the direct survey estimates although the BHF model estimates were more highly correlated with the direct survey estimates ($\beta_1 = 0.73$, r^2=0.891) compared to the correlation between the FH model and the direct survey estimates ($\beta_1 = 0.27$, r^2=0.477). This finding was contrary to the theory on SAE methodology [105]. Regional estimates for both the HF and BHF were generally similar to regional direct survey prevalence estimates [8]. The lower correlation of the FH model estimates with direct survey estimates compared to the BHF model estimates versus direct survey estimates may be attributed to the aggregated nature of the area-level covariates which mask any variations or heterogeneity of units within the district. The heterogeneous spread of HIV disease is well documented in many studies [8,22,34].

The coefficient of variation of both the BHF and FH model estimates were computed and used to compare variation of the estimates. The CV of the BHF estimates were larger than the CV of the FH model estimates implying that FH model estimates were more precise compared to the BHF model estimates. Coefficient of variation expresses the sampling variability of the model estimates from the point estimate [138]. Covariates for the unit level model comprise units across the district that are documented to be heterogeneous [22]. BHF model would therefore be preferred to the FH models since it incorporates variance of the model explanatory variables.

In Uganda, districts plan, implement and monitor interventions on behalf of the central government. However, information for planning is sparse as there are insufficient resources to collect representative data for monitoring social services at the district-level. This study demonstrates that it is possible to apply the SAE methodology, using available sources of data with no additional costs of data collection, to generate more precise estimates at district-levels consistent with findings from earlier studies [22,139,140].

Although antenatal HIV prevalence is often applied to monitor HIV/AIDS impact in settings with generalised epidemics, it contains information on: health facilities located in urban areas [4]; government aided facilities [3]; younger and more educated women [1,2]; and excludes non-pregnant women and men [41]. These limitations imply that antenatal survey data may not accurately be used to explain the general population HIV prevalence distribution.

District HIV prevalence estimates obtained using the SAE methods were consistent with regional HIV prevalence estimates based on direct survey estimates. That is, districts with higher HIV prevalence were those from regions with higher HIV prevalence. For example Kaborale district in Western region; Masaka and Mpigi in central 1 region; and Mbarara in south western region [8]. These districts are more urban and are also major transport corridors for truck drivers. Masaka district is also inhabited by fishing communities, who are documented to have elevated HIV prevalence compared to other population subgroups [14,90].

Overall, the district HIV prevalence estimates obtained in this study were more precise compared to the direct survey estimates. The standard errors were improved by 37.5%, 33.1% and 29% for the FH model, BHF Model and HPE methodologies respectively compared to the direct survey based estimates. While the regression based methods lead

to larger improvement in the standard errors, the HPE methodology is equally appealing and may be applicable in the absence of predictor data. Our findings further show that the availability of individual level predictor data provides more precise and reliable estimates for decision-making.

5.5 Contribution of Factors to Change in HIV Positivity between 2011 and 2016.

Using Blinder-Oaxaca decomposition method, the contribution of survey participants' characteristics (endowments) as well as the coefficients of the characteristics to change in HIV prevalence between 2011 and 2016 surveys were assessed. Descriptive analysis was used to assess the composition of survey participants and their HIV prevalence in the 2011 and 2016 surveys. Factors associated with HIV positivity were assessed using the pooled logistic regression.

The analysis results show that HIV prevalence reduced from 7.3% in 2011 to 6.0% in 2016 among individuals aged 15-49 years representing an 18% decline in HIV prevalence. Prevalence decreased in all population subgroups except among individuals who had no formal education; individuals with no sexual partners in the 12 months preceding the survey and among individuals aged 40-49 years.

Factors associated with a significant reduction in HIV prevalence include never being married compared to those who were married, secondary and tertiary level of education compared to primary level of education and being in the age group 15-19 years compared to the age 20-24 years. These socio-demographic sub groups represent individuals who learn and adopt new practices/behaviours, contributing to significant reductions in HIV prevalence.

Both the characteristics (endowments) of individuals surveyed and the coefficients of the characteristics had a significant contribution on the reduction in HIV prevalence between the two surveys. Changes in the population endowments (composition) that led to significantly larger change in HIV prevalence include young age (15-19 years), tertiary level of education, being previously married and urban residence.

While the coefficients contribute up to 154.6% of the increase in HIV prevalence between the two survey time points, the characteristics of the survey participants contribute 54% reduction in HIV prevalence between the two survey time points. The change in the characteristics ("endowments") of the survey participants refers to the differences in the composition of survey participants in the 2011 survey compared to the participants in the 2016 survey. Coefficient effects refers to change in HIV prevalence in 2016 if the effects of the coefficients in 2011 were maintained [28]. This finding implies that if the effects of the socio-demographic characteristics in 2011 were maintained there would be an overall increase HIV prevalence. Reduction in HIV prevalence may therefore be attributed to the HIV prevention, care and treatment interventions having direct impact on the characteristics of the survey participants in 2011.

In Uganda, a decline in HIV prevalence has been observed in earlier studies among women accessing ANC services [37,128] and among rural communities surveyed in western Uganda [39,141]. Key factors highlighted in these studies for the decline in HIV prevalence include change in sexual behaviour such as reduction in the number of sexual partners [90]. In this study, although the proportion of individuals with two or more sexual partners increased, there was a significant reduction in HIV prevalence among this population category. Studies elsewhere have attributed reducing or minimizing multiple sexual relationships to the adoption of acceptable and desirable behaviours following new

information on the realities of the HIV epidemic including the modes of transmission and prevention [35,142].

Additionally, effective interventions such as male circumcision were rolled out in Uganda in 2010, following studies that showed its effectiveness against HIV incidence [143] as well its acceptability as an HIV prevention tool [144,145]. Another key intervention was the test and treat policy rolled out in 2013 that recommended initiation of ART irrespective of CD4 count for individuals who tested HIV positive. This followed evidence of minimized chances of passing on the virus to un-infected sexual partner if the HIV positive individual was taking ART optimally [15,79,146–150]. The rollout of these interventions following the 2011 survey may therefore explain the reduction in HIV prevalence in the general population observed in the 2016 survey.

Increase in HIV prevalence on the other hand among individuals aged 40-49 years is attributable to cumulative cohort effect given that this group comprise both new incident cases as well as individuals infected earlier in their lives. This is compounded further by the effectiveness of the HIV treatment interventions that has enabled HIV infected individuals to live longer with the infection [15,53,79].

HIV prevalence among young people aged 15-19 years is argued to provide a better reflection of the nature of the HIV/AIDS epidemic in any population [79,151]. Prevalence in this age group is therefore often used as an indicator of the level of new infections since this population sub-group is newly exposed to risky behaviours such as initiation of sex [152–157]. In this study, although the proportion of young people (15-25 years) participating in the survey increased from 40% in 2011 to 48% in 2016, HIV prevalence decreased significantly from 2.4% to 1.2% and 5.3% to 3.3% among individuals aged 15-19 years and 20-24 years respectively. Additionally, the change in composition of young

people was significantly associated with change in HIV prevalence. Individuals aged 15-19 years contributed to 37% reduction in HIV prevalence between 2011 and 2016.

The large reduction in HIV prevalence among young people and the significant contribution to overall change in HIV prevalence has been attributed to both increased sensitivity to new information and the higher likelihood to adopt or change behaviour [1,41,158]. For example, DREAMs – an HIV intervention program targeting young people aged 15-24 years was rolled out in many countries in sub Saharan Africa including Uganda due to the evidence of high infection rates in the subgroup but also their higher likelihood to adapt to new preventive behaviours [159].

There was an increase in HIV prevalence among individuals with no formal education and those who reported no sexual partner in the 12 months preceding the survey, although the composition of these population categories decreased between the two surveys. Increased HIV prevalence may be attributed to lack of specific programs that target them. More specifically individuals who report no sexual relationships may be perceived to be at low risk of new infection yet they just concealing these relationship. Earlier studies have found higher prevalence of sexual relationships among individuals who report lack of any [35,73].

5.6 Strengths of the Study

Multilevel logistic regression and SAE models applied in this study have several advantages over classical logistic regression models. These include adjusting for cluster level correlation through the random effects component and inclusion of predictors at both individual and cluster levels. This enables more explicit modelling of the available data to obtain the prevalence estimates for districts; factors associated with HIV positivity and the

contribution of the factors to change in positivity. The weighting factor, propensity to test for HIV in a health facility, that was used to combine the population survey and health facility HIV testing data was also generated using multilevel logistic regression model.

The study applied the HPE and SAE methodologies using available data with no additional costs for data collection to obtain more precise HIV prevalence estimates for districts to aid planning and decision-making. These methodologies can therefore be applied in low resource settings where such information and resources to collect data is a challenge.

5.7 Study Limitations

The study applied multilevel logistic regression and SAE models, these models assume independence of the random area effects. This is unlikely to hold since district boundaries are set arbitrarily without any regard to the availability of services within the district. Additionally, individuals tend to seek health care services in places that are convenient or offer better quality services. This implies that estimates within a district may not be a true reflection of the HIV prevalence situation in the district.

This study used population survey data that are prone to missingness due to refusal to participate/respond. This may have affected the precision of the results since the effective sample size used to estimate HIV prevalence for each district is substantially reduced. Additionally, the results may also be biased since those who refuse to participate may have characteristics different from those who participated in the survey.

DHIS2 data includes individuals who may have tested multiple times which can lead to use of incorrect denominators for individuals tested at health facilities thus increasing bias of the estimates obtained. Additionally, some health facilities, especially those that are

privately owned and those located in hard to reach areas, do not report their data to the national DHIS2.

Population survey data consists of self-reported, stigmatizing, and sensitive information such as *“previously tested for HIV”, “previous test result”* and *“number of sexual partners”*. There is a likelihood of underreporting this information as well as preference by respondents to report only desirable experiences. For example, individuals who received a positive HIV test result or had previously tested multiple times may not want to share such information.

The reasons for not testing at a health facility was not captured during the survey, it was therefore not possible to ascertain the reasons for low coverage of facility HIV testing observed in the surveys especially among men and young people aged 15-24 years.

The study used data from two time points (2011 and 2016 surveys) only. Availability of more data points would enable validation of the trend and assessing the contribution of each of the factors to change in prevalence between 2011 and 2016 observed in the study.

6.0 CHAPTER SIX: CONCLUSIONS, IMPLICATIONS AND AREAS FOR FURTHER RESEARCH

6.1 Conclusions

Information for decision making for districts in Uganda are sparse, resulting in inefficient allocation of resources. This is largely due to the lack of resources to conduct large and regular surveys to monitor the status of interventions. Through this study, use of available data to generate estimates to inform planning and decision making was explored. The findings show that it's feasible to use existing population survey and routine facility HIV testing data either combined through the HPE method or as auxiliary variables in the SAE models to obtain more precise health indicator estimates for district level planning and decision making.

Additionally, while health facility HIV testing data is biased, i.e. contains information on only individuals who visit health facilities for services, population survey sample size at the district level is not adequate. The study therefore explored the comparability of these two dataset to inform using them complementarily [34]. The study found a two-fold increase in HIV prevalence among facility testers compared to those tested in a community setting. The prevalence ratio was higher among some population categories such as male, age groups 15-19, never married and those who had no sexual partners in the 12 months preceding the survey. These findings imply that to use the datasets to complement each other, there is need to take into account these differences.

This study further demonstrates that using facility and other available data as auxiliary information increases the precision of the district level estimates by up to 38% compared to estimates obtained based on only individuals surveyed in the district. Availability of individual/unit level auxiliary information such as those obtained using LQAS surveys as

was the case in this study leads to greater increase in the precision and reliability of the estimates compared to estimates based on aggregate district level auxiliary information.

The HPE methodology presents another appealing approach for obtaining estimates when survey sample sizes are small. Application of the methodology in this study presents two benefits: i) estimates with improved precision (up to 29% reduction in standard errors) compared to direct survey estimates and ii) computation of the standard errors of routine health facility data whose denominator is directly difficult to ascertain.

Although there have been extensive efforts and investments in HIV prevention based on knowledge of risk factors for new HIV infection and therefore prevalence, more research on the contribution of the factors/characteristics and the coefficient of the characteristics to change in HIV prevalence are needed. This knowledge will enable focusing in areas where the most impact can be realised. In this study, use of the decomposition analysis demonstrated that it's possible to isolate the contribution of the characteristics as well as the coefficient of the characteristics to identify those with the highest contribution to change in HIV prevalence. Young age group (15-19 years), individuals with higher level of education, those who were previously married and those who had two or more sexual partners had a larger contribution to the reduction in HIV prevalence observed. Similarly, the effects of the coefficient of the age group 40-19 years had a larger contribution to the reduction in HIV prevalence.

The SAE models and all the statistical methods implemented in this research can be carried out in freely available software such as R and STATA. The methods can directly be implemented in any software that is able to handle analysis of complex survey design data. To implement the methodologies, advanced or routine training may not be required.

6.2 Implication of the findings to HIV prevention policy

It is established from the findings in this research that pooling of multiple sources of data leads to gains in accuracy of estimates (>28% reduction in standard errors of the estimates compared to direct survey estimates), implying more accurate information on the level of HIV spread at districts available at no additional data collection costs. Many districts in Uganda are short of resources to assess the impact of social service interventions in their settings yet they are mandated to monitor service delivery on behalf of the central government including HIV services. While routine data is available to support monitoring, the limitations as highlighted in this and earlier research imply that they cannot be used to inform optimal monitoring of social service delivery impacts. This research therefore brings to the fore tools that both the central government (Ministry of Health) and District health managers can use to evaluate the impact of HIV prevention interventions. Multiple surveys are carried out by donor agencies in conjunction with the districts. The surveys may contain richer HIV risk factor variables that can be used with the tools tested and used in this research to improve monitoring of social services.

Furthermore, this research also provides population level estimates of HIV prevalence and it shows that prevalence varies significantly across districts in Uganda. It therefore presents critical information for HIV prevention prioritization in the country. The current country HIV prevention strategies emphasize concentrating prevention and treatment interventions to areas with higher HIV prevalence and among most at risk populations. For example, the country has been stratified into 3 categories (Sustained, Scale-up and Centrally supported) by development partners, USAID and CDC, in conjunction with government. More resources and support is being offered to districts categorised as scale-up and less into those that are only receiving central support. The levels of support within

each category is also defined to be similar. However, within the categories, there is limited information on level of HIV spread, thus hindering prioritization of the highly impacted districts within the same category. Information on HIV prevalence from this research can therefore be used to prioritize and expand interventions in districts with higher HIV prevalence.

6.3 Areas for further Research

Results from this doctoral work demonstrate that, it is feasible to obtain more reliable information for decision-making at lower administrative levels at no additional data collection costs. To extend this work, the following areas are proposed:

a) The only disaggregation considered before applying the HPE methodology was gender. Findings from objective 3 show that there were differences in HIV prevalence between facility and non-facility testers by among others age, gender, education level and marital status. It is therefore prudent to consider further disaggregation while combining the datasets.
b) Differences in HIV prevalence by venue of testing was established, this could be incorporated as prior information while combining the data sets. Applying the HPE methodology incorporating Bayesian methods therefore needs to be considered to assess whether it has any impact in the estimates obtained.
c) Due to missing data limitations, we were unable to explore the contribution of more critical HIV risk factors such as condom use during last high risk sex and age at sexual debut or marriage. Application of missing data methods needs to be considered to overcome the missing data problem to include more HIV related auxiliary information in the SAE models as well increase the effective analysis sample size for the HPE methodology. This will also enable an assessment of the contribution of these risk

factors to change in HIV prevalence over time. Additionally, missing methods are widely developed, integrating these methods within the Hybrid Prevalence Estimation (HPE) and Small Area Estimation (SAE) methods needs to be explored.

d) The Hybrid Prevalence Estimation (HPE) and Small Area Estimation (SAE) methods incorporate auxiliary/ supplemental information through different approaches. While the HPE methodology combines the data sources linearly, SAE methods implemented in this research incorporate auxiliary information as predictors at unit and/or target area level. A formal assessment and comparison of the methods via a simulation study would highlight differences in the estimates obtained from the methodologies.

References

1. Gregson S, Terceiria N, Kakowa M, Mason PR, Anderson RM, Chandiwana SK CM. Study of bias in antenatal clinic HIV-1 surveillance data in a high contraceptive prevalence population in sub-Saharan Africa. AIDS. 2002;16(4):643–52.

2. Saphonn V, Hor LB, Ly SP, Chhuon S, Saidel T, Detels R. How well do antenatal clinic (ANC) attendees represent the general population? A comparison of HIV prevalence from ANC sentinel surveillance sites with a population-based survey of women aged 15-49 in Cambodia. Int J Epidemiol [Internet]. 2002;31(2):449–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11980815

3. Giorgi E, Sesay SSS, Terlouw DJ, Diggle PJ. Combining data from multiple spatially referenced prevalence surveys using generalized linear geostatistical models. J R Stat Soc Ser A Stat Soc. 2015;178(2):445–64.

4. UNAIDS/WHO. Guidelines for Conducting HIV Sentinel Serosurveys among Pregnant Women and Other Groups. 2003.

5. Uganda Bureau of Statistics. Uganda Bureau of Statistics Statistical. 2014;1–305. Available from: http://www.ubos.org/onlinefiles/uploads/ubos/statistical_abstracts/Statistical Abstract 2014.pdf. Accessed: August 15, 2016

6. Uganda Bureau of Statistics (UBOS) and ICF. Uganda Demographic and Health Survey 2016 [Internet]. Kampala Uganda and Rockville, Maryland, USA; 2016. Available from: https://www.ubos.org/onlinefiles/uploads/ubos/pdf documents/Uganda_DHS_2016_KIR.pdf

7. Ministry of Heath and ICF international. Uganda AIDS Indicator Survey (AIS) 2011 [Internet]. Kampala Uganda and Rockville, Maryland, USA; 2012. Available from: http://health.go.ug/docs/UAIS_2011_REPORT.pdf

8. Ministry of Health Uganda. Uganda Population-based HIV Impact Assessment (UPHIA) 2016-2017: Final Report. [Internet]. Kampala, Uganda; 2019. Available from: https://phia.icap.columbia.edu/wp-

content/uploads/2019/07/UPHIA_Final_Report_Revise_07.11.2019_Final_for-web.pdf

9. Uganda Ministry of Health. Uganda HIV/AIDS Sero-Behavioural Survey 2004-05. 2004;36; 107. Available from: http://pdf.usaid.gov/pdf_docs/Pnadg508.pdf

10. Hedt BL, Pagano M. Health indicators: Eliminating bias from convenience sampling estimators. Stat Med. 2011;30(5):560–8.

11. Gregson, S. Dharmayat K, Pereboom M, Takaruza A, Mugurungi O, Schur N, Nyamukapa CA. Do HIV prevalence trends in antenatal clinic surveillance represent trends in the general population in the antiretroviral therapy era? The case of Manicaland, East Zimbabwe. AIDS. 2015;29(14):1845–53.

12. Musinguzi J, Kirungi W, Opio A, Montana L, Mishra V, Madraa E, et al. Comparison of HIV Prevalence Estimates From Sentinel Surveillance and a National Population-Based Survey in Uganda, 2004-2005. JAIDS J Acquir Immune Defic Syndr. 2009;51(1):78–84.

13. Kapesa A, Basinda N, Nyanza EC, Mushi MF, Jahanpour O, Ngallaba SE. Prevalence of HIV infection and uptake of HIV/AIDS services among fisherfolk in landing Islands of Lake Victoria, north western Tanzania. BMC Health Serv Res [Internet]. 2018 Dec 18 [cited 2019 Apr 2];18(1):980. Available from: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-018-3784-4

14. Opio A, Muyonga M, Mulumba N. HIV Infection in Fishing Communities of Lake Victoria Basin of Uganda – A Cross-Sectional Sero-Behavioral Survey. PLoS One. 2013;8(8).

15. Johnson LF, May MT, Dorrington RE, Cornell M, Boulle A, Egger M, et al. Estimating the impact of antiretroviral treatment on adult mortality trends in South Africa: A mathematical modelling study. PLoS Med. 2017;14(12):1–17.

16. Avenir Health. Avenir Health: One Health Tool [Internet]. [cited 2021 Jan 23]. Available from: https://avenirhealth.org/software-spectrum.php

17. Stover J, Brown T, Puckett R. Updates to the Spectrum / Estimations and Projections Package model for estimating trends and current values for key HIV

indicators. 2017;0(October 2016):5–11.

18. Stover J, Andreev K, Slaymaker E, Gopalappa C, Sabin K, Velasquez C, et al. Updates to the Spectrum model to estimate key HIV indicators for adults and children. Aids [Internet]. 2014;28 Suppl 4(September):S427-34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25406748

19. Larmarange J, Bendaud V. HIV estimates at second subnational level from national population-based surveys. AIDS. 2014;28(Suppl 4):S469-2476.

20. Hoshi T, Fuji Y, Nzou SM, Tanigawa C, Kiche I, Mwau M, et al. Spatial distributions of HIV infection in an endemic area of Western Kenya: Guiding information for localized HIV control and prevention. PLoS One. 2016;11(2):1–14.

21. Ngesa O, Mwambi H, Achia T. Bayesian spatial semi-parametric modeling of HIV variation in Kenya. PLoS One. 2014;9(7).

22. Gutreuter S, Igumbor E, Wabiri N, Desai M, Durand L. Improving estimates of district HIV prevalence and burden in South Africa using small area estimation techniques. PLoS One [Internet]. 2019;14(2):1–14. Available from: http://dx.doi.org/10.1371/journal.pone.0212445

23. Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020;20(1):1–14.

24. Leon N, Mathews C, Lewin S, Osler M, Boulle A, Lombard C. A comparison of linkage to HIV care after provider-initiated HIV testing and counselling (PITC) versus voluntary HIV counselling and testing (VCT) for patients with sexually transmitted infections in Cape Town, South Africa. BMC Health Serv Res. 2014 Dec 18;14(1):350.

25. Roura M, Watson-jones D, Kahawita TM, Ferguson L, Ross DA. Provider-initiated testing and counselling programmes in sub-Saharan Africa : a systematic review of their operational implementation. AIDS. 2013;(27):617–26.

26. Wilson KC, Mhangara M, Dzangare J, Eaton JW, Hallett TB, Mugurungi O, et al.

Does nonlocal women's attendance at antenatal clinics distort HIV prevalence surveillance estimates in pregnant women in Zimbabwe? AIDS. 2017;31(Suppl 1):S95–102.

27. Blinder AS. Wage Discrimination : Reduced Form and Structural Estimates. J Hum Resour. 1973;8(4):436–55.

28. Powers DA, Yoshioka H, Yun M-S. Mvdcmp: Multivariate Decomposition for Nonlinear Response Models. Stata J Promot Commun Stat Stata. 2011;11(4):556–76.

29. UNAIDS. Global HIV & AIDS statistics — 2020 fact sheet _ UNAIDS. 2020.

30. Nishiura H. Estimating the incidence and diagnosed proportion of HIV infections in Japan: A statistical modeling study. PeerJ. 2019;2019(1).

31. Baggaley R. WHO Consolidated Guidelines on HIV Testing Services What you need to know The aims of the new WHO HTS Guidelines. Who. 2015;(July 2015).

32. Wong GY, Mason WM. The Hierarchical Logistic Regression Model for Multilevel Analysis. J Am Stat Assoc. 1985;80(391):513–24.

33. Goldstein H. Extensions to the basic multilevel model. Multi-level Stat Model. 1995;64–88.

34. Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Levin J. Difference in HIV prevalence by testing venue : results from population level survey in Uganda survey in Uganda. AIDS Care [Internet]. 2020;0(0):1–12. Available from: https://doi.org/10.1080/09540121.2020.1734179

35. Fylkesnes K, Musonda RM, Sichone M, Ndhlovu Z, Tembo F, Monze M. Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population-based surveys. AIDS. 2001;15(7):907–16.

36. Green EC, Halperin DT, Nantulya V, Hogle JA. Uganda's HIV prevention success: The role of sexual behavior change and the national response. AIDS Behav. 2006;10(4):335–46.

37. Kirungi WL, Musinguzi J, Madraa E, Mulumba N, Callejja T, Ghys P. Trends in

antenatal HIV prevalence in urban Uganda associated with uptake of preventive sexual behaviour. Sex Transm Infect [Internet]. 2006;82(suppl 1 (i36-i41)). Available from: www.stijournal.com

38. Stoneburner RL, Low-Beer D. Population-Level HIV Declines and Behavioral Risk Avoidance in Uganda. Science (80-) [Internet]. 2004 Apr 30;304(5671):714 LP – 718. Available from: http://science.sciencemag.org/content/304/5671/714.abstract

39. Shafer LA, Biraro S, Nakiyingi-Miiro J, Kamali A, Ssematimba D, Ouma J, et al. HIV prevalence and incidence are no longer falling in southwest Uganda: Evidence from a rural population cohort 1989-2005. Aids. 2008;22(13):1641–9.

40. Judith RG, Anne B, Michel C, Rosemary MM, Maina K, Isaac M, Francis T LZ. Factors influencing the difference in HIV prevalence between antenatal clinic and general population in sub- Saharan Africa. AIDS. 2001;15:1717–25.

41. Zaba BW, Carpenter LM, Boerma JT, Gregson S, Nakiyingi J, Urassa M. Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa. AIDS. 2000;14(17):2741–50.

42. CDC. HEALTH INFORMATICS, DATA MANAGEMENT AND STATISTICS: SUB-NATIONAL ESTIMATION OF HIV INDICATORS [Internet]. [cited 2021 Jan 31]. Available from: https://www.cdc.gov/globalhivtb/who-we-are/resources/Health-Informatics-Data-Management-and-Statistics-Sub-National-Estimation-of-HIV-Indicators.pdf

43. Bärnighausen T, Tanser F. Rethinking the role of the local community in HIV epidemic spread in sub-Saharan Africa : a proximate-determinants approach. HIV Ther. 2009;3(5):435–45.

44. UNAIDS. Methodology – Understanding the HIV estimates. Produced by the Strategic Information and Monitoring Division. 2013;(November):11. Available from: http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/20131118_Methodology.pdf

45. UNAIDS, Joint Programme on HIV/AIDS with Unicef, UNDP, UNFPA, UNDCP,

UNESCO, WHO WB. Report on the global HIV/AIDS epidemic. Geneva, biennial. AIDS epidemic update. 2002.

46. UNAIDS/WHO. AIDS Epidemic Update 2007. UNAIDS WHO [Internet]. 2007;1. Available from: http://data.unaids.org/pub/EPISlides/2007/2007_epiupdate_en.pdf

47. Baryarama F, Bunnell R, McFarland W, Hudes ES, Neilands TB, Ransom RL, et al. Estimating HIV incidence in voluntary counseling and testing clients in Uganda (1992-2003). J Acquir Immune Defic Syndr. 2007;44(1):99–105.

48. Baryarama F, Bunnell RE, Ransom RL, Ekwaru JP, Kalule J, Tumuhairwe EB, et al. Using HIV voluntary counseling and testing data for monitoring the uganda HIV epidemic, 1992-2000. J Acquir Immune Defic Syndr. 2004;37(SUPPL. 1):1180–6.

49. Lohr S, Rao JNK. Estimation in Multiple-Frame Surveys. J Am Stat Assoc [Internet]. 2006;101(475):1019–30. Available from: http://www.tandfonline.com/doi/abs/10.1198/016214506000000195

50. Sheng B, Marsh K, Slavkovic AB, Gregson S, Eaton JW, Bao L. Statistical models for incorporating data from routine HIV testing of pregnant women at antenatal clinics into HIV/AIDS epidemic estimates. AIDS. 2017;31(Suppl 1):S87–94.

51. Group UR, Boston P. Method Development for the UNAIDS Estimates : June 2015 REPORT & RECOMMENDATIONS. 2015;(June):2–4.

52. Makinde O a, Oyediran K a. Rethinking HIV prevalence determination in developing countries. AIDS Care [Internet]. 2014;0121(September):1–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25174731

53. Yotebieng M, Brazier E, Addison D, Kimmel AD, Cornell M, Keiser O, et al. Research priorities to inform “Treat All” policy implementation for people living with HIV in sub-Saharan Africa: a consensus statement from the International epidemiology Databases to Evaluate AIDS (Ie DEA). J Int AIDS Soc. 2019;22(1):e25218.

54. Mekonnen ZA, Lerebo WT, Gebrehiwot TG, Abadura SA. Multilevel analysis of individual and community level factors associated with institutional delivery in Ethiopia. BMC Res Notes. 2015;1–9.

55. Sprague DA, Jeffery C, Crossland N, House T, Roberts GO, Vargas W, et al. Assessing delivery practices of mothers over time and over space in Uganda, 2003-2012. Emerg Themes Epidemiol. 2016;13(1).

56. Paulin HN, Blevins M, Koethe JR, Hinton N, Vaz LME, Vergara AE, et al. HIV testing service awareness and service uptake among female heads of household in rural Mozambique: Results from a province-wide survey. BMC Public Health. 2015;15(1):1–11.

57. Fabiani M, Fylkesnes K, Nattabi B, Ayella EO, Declich S. Evaluating two adjustment methods to extrapolate HIV prevalence from pregnant women to the general female population in sub-Saharan Africa. AIDS. 2003;17(3):399–405.

58. Paper W, Application A. A New Approach to Producing Geographic Profiles of HIV Prevalence An Application to Malawi. 2010;(February).

59. WHO. Guidelines for measuring national HIV prevalence in population-based surveys Guidelines for measuring national HIV prevalence in population-based surveys [Internet]. WHO Technical Reports. 2005. Available from: http://data.unaids.org/pub/manual/2005/20050101_gs_guidemeasuringpopulation_en.pdf

60. Mishra V, Barrere B, Hong R, Khan S, V. M, B. B, et al. Evaluation of bias in HIV seroprevalence estimates from national household surveys. Sex Transm Infect [Internet]. 2008 Aug 1;84 Suppl 1(Supplement 1):i63–70. Available from: http://sti.bmj.com/content/84/Suppl_1/i63.short%5Cnhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed8&NEWS=N&AN=2008380126%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2569831&tool=pmcentrez&rendertype=abstract

61. Hogan DR, Salomon JA, Canning D, Hammitt JK, Zaslavsky AM, Barnighausen T, et al. National HIV prevalence estimates for sub-Saharan Africa: controlling selection bias with Heckman-type selection models. Sex Transm Infect [Internet]. 2012;88 Suppl 2(Ci):i17-23. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3512441&tool=pmcentrez&rendertype=abstract

62. Brookmeyer R. Measuring the HIV/AIDS epidemic: Approaches and challenges. Epidemiol Rev. 2010;32(1):26–37.

63. Grantham KL, Kerr CC, Wilson DP. Local responses to local epidemics for national impact need advanced spatially explicit tools. AIDS [Internet]. 2016;30(9):1481–2. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201606010-00019

64. Anderson S-J, Consortium M. Evaluation of geospatial methods to generate subnational HIV prevalence estimates for local level planning. Aids [Internet]. 2016;30(9):1467–74. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201606010-00017

65. Brown T, Bao L, Eaton JW, Hogan DR, Mahy M, Marsh K, et al. Improvements in prevalence trend fitting and incidence estimation in EPP 2013. Vol. 28, AIDS. 2014. p. S415–25.

66. Stover J, Brown T, Puckett R, Sheng B, Marsh K, Slavkovic AB, et al. Estimation in Multiple-Frame Surveys. AIDS. 2015;31(1):1689–99.

67. Kleinschmidt I, Pettifor A, Morris N, MacPhail C, Rees H. Geographic distribution of human immunodeficiency virus in South Africa. Am J Trop Med Hyg. 2007;77(6):1163.

68. Ngianga-Bakwin Kandala, Paul Brodish, Bates Buckner, Susan Foster and NM. Millennium development goal 6 and HIV infection in Zambia: what can we learn from successive household surveys? NIH Public Access. 2012;76(October 2009):211–20.

69. UNAIDS. Developing Subnational Estimates of HIV Prevalence and the Number of People Living with HIV. 2014.

70. Kondlo L, Manda S. Small Area Estimation of HIV Prevelance using National Survey data in South Africa. Int Stat Inst Proc 58th World Stat Congr 2011, Dublin (Session CPS045). 2011;(2007):5002–12.

71. Manda S, Masenyetse L, Cai B, Meyer R. Mapping HIV prevalence using population and antenatal sentinel-based HIV surveys: a multi-stage approach. Popul Health Metr. 2015;13:22.

72. Mullens AB, Kelly J, Debattista J, Phillips TM, Gu Z, Siggins F. Exploring HIV risks, testing and prevention among sub-Saharan African community members in Australia. Int J Equity Health [Internet]. 2018 Dec 21 [cited 2019 Apr 2];17(1):62. Available from: https://equityhealthj.biomedcentral.com/articles/10.1186/s12939-018-0772-6

73. Nalugoda F, Guwatudde D, Bwaninka JB, Makumbi FE, Lutalo T, Kagaayi J, et al. Marriage and the risk of incident HIV infection in Rakai, Uganda. J Acquir Immune Defic Syndr [Internet]. 2014 Jan 1;65(1):91–8. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24419066

74. Amornkul PN, Vandenhoudt H, Nasokho P, Odhiambo F, Mwaengo D, Hightower A, et al. HIV prevalence and associated risk factors among individuals aged 13-34 years in rural Western Kenya. PLoS One. 2009;4(7).

75. Bao L, Raftery AE, Reddy A. Estimating the sizes of populations at risk of HIV infection from multiple data sources using a Bayesian hierarchical model. Stat Interface [Internet]. 2015;8(2):125–36. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4442027&tool=pmcentrez&rendertype=abstract%5Cnhttp://www.intlpress.com/site/pub/pages/journals/items/sii/content/vols/0008/0002/a001/

76. Ntirampeba D, Neema I, Kazembe LN. Joint spatial modelling of disease risk using multiple sources : an application on HIV prevalence from antenatal sentinel and demographic and health surveys in Namibia. 2017;1–16.

77. Hedt BL. Novel methods for efficient surveillance and monitoring. Diss Abstr Int Sect B Sci Eng [Internet]. 2008;69(4-B):2034. Available from: http://search.ebscohost.com.proxy-ub.rug.nl/login.aspx?direct=true&db=psyh&AN=2008-99200-569&site=ehost-live&scope=site

78. Jeffery C, Pagano M, Hemingway J, Valadez JJ. Hybrid prevalence estimation :

Method to improve intervention coverage estimations. PNAS. 2018;115(51):13063–8.

79. Frank TD, Carter A, Jahagirdar D, Biehl MH, Douwes-Schultz D, Larson SL, et al. Global, regional, and national incidence, prevalence, and mortality of HIV, 1980-2017, and forecasts to 2030, for 195 countries and territories: A systematic analysis for the Global Burden of Diseases, Injuries, and Risk Factors Study 2017. Lancet HIV. 2019;6(12):e831–59.

80. Hajizadeh M, Sia D, Heymann SJ, Nandi A. Socioeconomic inequalities in HIV/AIDS prevalence in sub-Saharan African countries: Evidence from the Demographic Health Surveys. Int J Equity Health [Internet]. 2014;13(1):1–22. Available from: International Journal for Equity in Health

81. Nutor JJ, Duah HO, Agbadi P, Duodu PA, Gondwe KW. Spatial analysis of factors associated with HIV infection in Malawi : indicators for effective prevention. 2020;1–14.

82. Chinomona A, Mwambi HG. Estimating HIV prevalence in Zimbabwe using population-based survey data. PLoS One. 2015;10(12):1–17.

83. Asiedu C, Asiedu E, Owusu F. The Socio-Economic Determinants of HIV/AIDS Infection Rates in Lesotho, Malawi, Swaziland and Zimbabwe. Dev Policy Rev. 2012;30(3):305–26.

84. Piot P, Kazatchkine M, Dybul M, Lob-Levyt J. AIDS: lessons learnt and myths dispelled. Lancet [Internet]. 2009;374(9685):260–3. Available from: http://dx.doi.org/10.1016/S0140-6736(09)60321-4

85. Fox A. The HIV-poverty thesis re-examined: poverty, wealth or inequality as a social determinant of HIV infection in sub-Saharan Africa? Vol. 44, Journal of biosocial science. 2012. 459–480 p.

86. Thamattoor U, Thomas T, Banandur P, Rajaram S, Duchesne T, Abdous B, et al. Multilevel analysis of the predictors of HIV prevalence among pregnant women enrolled in annual HIV sentinel surveillance in four states in Southern India. PLoS One [Internet]. 2015;10(7):1–11. Available from:

http://dx.doi.org/10.1371/journal.pone.0131629

87. Hegdahl HK, Fylkesnes KM, Sandøy IF. Sex differences in HIV prevalence persist over time: Evidence from 18 countries in Sub-Saharan Africa. PLoS One [Internet]. 2016;11(2):1–17. Available from: http://dx.doi.org/10.1371/journal.pone.0148502

88. Valadez JJ, Jeffery C, Davis R, Ouma J, Lwanga SK, Moxon S. Putting the C back into the ABCs: A multi-year, multi-region investigation of condom use by Ugandan youths 2003-2010. PLoS One. 2014;9(4).

89. Grosskurth TH, Todd J, Mwijarubi E, Mayaud P, Nicoll FRCP A, ka-Gina G, et al. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomised controlled trial Summary. Vol. 346, African Medical and Research Foundation (AMREF). 1995.

90. Mafigiri R, Matovu JKBB, Makumbi FE, Ndyanabo A, Nabukalu D, Sakor M, et al. HIV prevalence and uptake of HIV/AIDS services among youths (15-24 Years) in fishing and neighboring communities of Kasensero, Rakai District, South Western Uganda. BMC Public Health [Internet]. 2017 Dec 14 [cited 2019 Apr 2];17(1):251. Available from: http://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-017-4166-2

91. Human Rights Watch. Curtailing Criticism: Intimidation and Obstruction of Civil Society in Uganda [Internet]. 2012. Available from: https://www.hrw.org/sites/default/files/reports/uganda0812ForUpload.pdf

92. Uganda Bureau of Statistics 2016. The National Population and Housing Census 2014 – Main Report. Kampala, Uganda; 2016.

93. Ralf M, Burgard JP, Florian, Ertz; Simon, Lenau; Julia MHM. Guidelines on small area estimation for city statistics and other functional geographies. 2019.

94. UGANDA AIDS COMMISSION. UGANDA HIV/AIDS COUNTRY PROGRESS REPORT [Internet]. 2016 [cited 2019 Apr 1]. Available from: http://www.unaids.org/sites/default/files/country/documents/UGA_2018_countryreport.pdf

95. UNAIDS. AIDSinfo | UNAIDS [Internet]. [cited 2019 Mar 31]. Available from:

http://aidsinfo.unaids.org/

96. Cohen J, Tate T. The Less They Know, the Better: Abstinence-Only HIV/AIDS Programs in Uganda. 2006;8080.

97. Kalibala S. Monitoring and Evaluation of the Emergency Plan Progress (MEEPP): End-of-Project Evaluation. 2010;(November).

98. Kiberu VM, Matovu JK, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. BMC Med Inform Decis Mak. 2014;14(1):40.

99. Uganda Bureau of Statistics (UBOS) and ICF International Inc. Uganda Demographic and Health Survey 2011. Kampala, Uganda; 2012.

100. Robertson SE, Valadez JJ. Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004. Soc Sci Med. 2006;63(6):1648–60.

101. Assessment HF. STAR-E LQAS Key results of the 2013 LQAS community surveys & Health Facility Assessment. 2013;1–27.

102. Azen S.P. PM. KDBJ. Obtaining Confidence Intervals for the Risk Ratio in Cohort Studies. Int Biometric Soc. 1978;34(3):469–74.

103. StataCorp. Stata Statistical Software: Release 15 [Internet]. College Station, TX: StataCorp LLC; 2017. Available from: https://www.stata.com/

104. Bland JM, Altman DG. Statistical methods for assessing agreement betweeen two methods of clinical treatement. (fig 1):1–9.

105. Rao JNK, Molina I. Small Area Estimation [Internet]. Wiley; 2015. (In Wiley Online Library: Books). Available from: https://books.google.co.za/books?id=i1B_BwAAQBAJ

106. Battese GE, Harter RM, Fuller WA. An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data. J Am Stat Assoc [Internet]. 1988;83(401):28–36. Available from: http://www.jstor.org/stable/2288915?origin=crossref%5Cnhttp://www.jstor.org/stabl

e/2288915

107. Molina I, Rao JNK. Estimation of Poverty Measures in Small Areas. (1984):1–4.

108. Uganda Bureau of Statistics. 2002 Uganda Population and Housing Census Administrative Report. 2007.

109. UBOS. National housing and population census 2014-Area Specific Profiles, Wakiso District. 2017.

110. Bland U, Giavarina D. Lessons in biostatistics. 2015;25(2):141–51.

111. Molina I, Marhuenda Y. sae : An R Package for Small Area Estimation. 2015;7(June):81–98.

112. Chirawu P, Langhaug L, Mavhu W, Pascoe S, Dirawo J, Cowan F, et al. Acceptability and challenges of implementing voluntary counselling and testing (VCT) in rural Zimbabwe: evidence from the Regai Dzive Shiri Project. AIDS Care. 2010;22(1):81–8.

113. Govindasamy D, Ferrand RA, Wilmore SM, Ford N, Ahmed S, Afnan-Holmes H, et al. Uptake and yield of HIV testing and counselling among children and adolescents in sub-Saharan Africa: A systematic review. J Int AIDS Soc [Internet]. 2015 Jan 1 [cited 2019 Apr 1];18(1):no pagination. Available from: http://www.jiasociety.org/index.php/jias/article/view/20182/pdf_1%5Cnhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed13&NEWS=N&AN=2015488483

114. Lugada E, Levin J, Abang B, Mermin J, Mugalanzi E, Namara G, et al. Comparison of home and clinic-based HIV testing among household members of persons taking antiretroviral therapy in Uganda: results from a randomized trial. J Acquir Immune Defic Syndr. 2010 Oct 1;55(2):245–52.

115. Sharma M, Ying R, Tarr G, Barnabas R, Division ID, Hutchinson F. Systematic review and meta-analysis of community and facility-based HIV testing to address linkage to care gaps in sub-Saharan Africa. Nature. 2016 Dec 2;528(7580):1–26.

116. Silvestri DM, Modjarrad K, Blevins ML, Halale E, Vermund SH, Mckinzie JP. A

comparison of HIV detection rates using routine opt-out provider-initiated HIV testing and counseling versus a standard of care approach in a rural African setting. J Acquir Immune Defic Syndr. 2011;56(1).

117. Ssebunya RN, Wanyenze RK, Namale L, Lukolyo H, Kisitu GP, Nahirya-Ntege P, et al. Prevalence and correlates of HIV testing among adolescents 10–19 years in a post-conflict pastoralist community of Karamoja region, Uganda. BMC Public Health [Internet]. 2018 Dec 10 [cited 2019 Apr 2];18(1):612. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-018-5544-0

118. Nannozi V, Wobudeya E, Matsiko N, Gahagan J. Motivators of couple HIV counseling and testing (CHCT) uptake in a rural setting in Uganda. BMC Public Health [Internet]. 2017;17(1):1–6. Available from: http://dx.doi.org/10.1186/s12889-017-4043-z

119. Ramirez-Avila L, Nixon K, Noubary F, Giddy J, Losina E, Walensky RP, et al. Routine HIV Testing in Adolescents and Young Adults Presenting to an Outpatient Clinic in Durban, South Africa. PLoS One. 2012;7(9):5–9.

120. Mahiane SG, Marsh K, Grantham K, Crichlow S, Caceres K, Stover J. Improvements in Spectrum's fit to program data tool. Aids [Internet]. 2017;31(November 2016):S23–30. Available from: http://insights.ovid.com/crossref?an=00002030-201704001-00004

121. Chan M, Kazatchkine M, Lob-levyt J, Obaid T, Schweizer J, Veneman A, et al. Meeting the Demand for Results and Accountability : A Call for Action on Health Data from Eight Global Health Agencies. PLoS Med. 2010;7(1):5–8.

122. UNAIDS. Ten targets: 2011 United Nations Political Declaration on HIV and AIDS. 2015;53-pp. Available from: http://www.unaids.org/en/resources/documents/2015/ten_targets

123. Boerma T, Eozenou P, Evans D, Evans T, Kieny MP, Wagstaff A. Monitoring Progress towards Universal Health Coverage at Country and Global Levels. PLoS Med. 2014;11(9).

124. Carle AC. Fitting multilevel models in complex survey data with design weights:

Recommendations. BMC Med Res Methodol. 2009;9(1):1–13.

125. Kovačević MS, Rai SN. A pseudo maximum likelihood approach to multilevel modelling of survey data. Commun Stat - Theory Methods. 2003;32(1):103–21.

126. Rabe-Hesketh S, Skrondal A. Multilevel modelling of complex survey data. J R Stat Soc Ser A Stat Soc. 2006;169(4):805–27.

127. Lakew Y, Benedict S, Haile D. Social determinants of HIV infection, hotspot areas and subpopulation groups in Ethiopia: evidence from the National Demographic and Health Survey in 2011. BMJ Open. 2015;5(11):e008669.

128. Asiimwe-Okiror G, Opio AA, Musinguzi J, Madraa E, Tembo G, Caraël M. Change in sexual behaviour and decline in HIV infection among young pregnant women in urban Uganda. Aids. 1997;11(14):1757–63.

129. Asaolu IO, Gunn JK, Center KE, Koss MP, Iwelunmor JI, Ehiri JE. Predictors of HIV testing among youth in sub-Saharan Africa: A cross-sectional study. PLoS One [Internet]. 2016;11(10):1–12. Available from: http://dx.doi.org/10.1371/journal.pone.0164052

130. Facha W, Kassahun W, Workicho A. Predictors of provider- initiated HIV testing and counseling refusal by outpatient department clients in Wolaita zone , Southern Ethiopia : a case control study. BMC Public Health. 2016;1–8.

131. Kiros G, Workagegn F, Gebretsadik LA. Predictors of HIV-test utilization in PMTCT among antenatal care attendees in government health centers: institution-based cross-sectional study using health belief model in Addis Ababa, Ethiopia, 2013. HIV/AIDS - Res Palliat Care. 2015;215.

132. Turan JM, Bukusi EA, Onono M, Holzemer WL, Miller S, Cohen CR. HIV/AIDS stigma and refusal of HIV testing among pregnant women in rural Kenya: Results from the MAMAS study. AIDS Behav. 2011;15(6):1111–20.

133. Maina I, Wanjala P, Soti D, Kipruto H, Boerma T. Using health-facility data to assess subnational coverage of maternal and child health indicators, Kenya. Bull World Health Organ. 2017;95:683–94.

134. J.N.K. Rao. Small Area Estimation: Methods, Applications and New Developments. Vol. 1. 2013. p. 80.

135. Farrell PJ, Macgibbon B, Tomberlin TJ, Farrell PJ. Small-Area Estimation Estimation Using Using Empirical Logistic Regression Models and and Summary Summary Logistic Statistics Statistics. 2011;15(1):101–8.

136. Rao JNK. Some Recent Advances in Model- Based Small Area Estimation. Surv Methodol [Internet]. 1999;25(2):175–86. Available from: http://www.statcan.gc.ca/pub/12-001-x/1999002/article/4880-eng.pdf

137. Tzavidis N, Zhang LC, Luna A, Schmid T, Rojas-Perilla N. From start to finish: a framework for the production of small area official statistics. J R Stat Soc Ser A Stat Soc. 2018;181(4):927–79.

138. Raykov T, Marcoulides GA, Akaeze HO. Comparing Between- and Within-Group Variances in a Two-Level Study: A Latent Variable Modeling Approach to Evaluating Their Relationship. Educ Psychol Meas. 2017;77(2):351–61.

139. Johnson FA, Chandra H, Brown JJ. District-level Estimates of Institutional Births in Ghana : Application of Small Area Estimation Technique Using Census and DHS Data. 2010;26(2):341–59.

140. Johnson FA, Padmadas SS, Chandra H, Matthews Z, Madise NJ, Amoako F, et al. Estimating unmet need for contraception by district within Ghana : An application of small-area estimation techniques. 2012;4728(May).

141. Kamali A, Carpenter LM, Whitworth JAG, Pool R, Ruberantwari A, Ojwiya A. Seven-year trends in HIV-1 infection rates, and changes in sexual behaviour, among adults in rural Uganda. Aids. 2000;14(4):427–34.

142. Schaefer R, Thomas R, Maswera R, Kadzura N, Nyamukapa C, Gregson S. Relationships between changes in HIV risk perception and condom use in East Zimbabwe 2003-2013: Population-based longitudinal analyses. BMC Public Health. 2020;20(1).

143. Gray R, Kigozi G, Kong X, Ssempiija V, Makumbi F, Wattya S, et al. The effectiveness of male circumcision for HIV prevention and effects on risk behaviors

in a posttrial follow-up study. Aids. 2012;26(5):609–15.

144. Albert LM, Akol A, L'Engle K, Tolley EE, Ramirez CB, Opio A, et al. Acceptability of male circumcision for prevention of HIV infection among men and women in Uganda. AIDS Care - Psychol Socio-Medical Asp AIDS/HIV. 2011 Dec;23(12):1578–85.

145. Odoch WD, Kabali K, Ankunda R, Zulu JM, Tetui M. Introduction of male circumcision for HIV prevention in Uganda: analysis of the policy process. Heal Res Policy Syst [Internet]. 2015 Jun 20 [cited 2021 Feb 13];13(1):31. Available from: https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-015-0020-0

146. Keith A. SEARCH study exceeds 90-90-90 targets after 2 years of "test and treat" for HIV in rural East Africa | aidsmap [Internet]. 2013 [cited 2021 Jan 23]. Available from: https://www.aidsmap.com/news/jul-2016/search-study-exceeds-90-90-90-targets-after-2-years-test-and-treat-hiv-rural-east

147. Livia, Montana; Melissa, Neuman; Vnod M. Spatial modelling of HIV prevalence in Kenya. Vol. 0. 2007.

148. Das M, Chu PL, Santos G-M, Scheer S, Vittinghoff E, Mcfarland W, et al. Decreases in Community Viral Load Are Accompanied by Reductions in New HIV Infections in San Francisco. [cited 2021 Feb 13]; Available from: www.plosone.org

149. Abongomera G, Kiwuwa-Muyingo S, Revill P, Chiwaula L, Mabugu T, Phillips AN, et al. Impact of decentralisation of antiretroviral therapy services on HIV testing and care at a population level in agago district in rural Northern Uganda: Results from the Lablite population surveys. Int Health. 2017;9(2):91–9.

150. Johnson LF, Davies MA, Moultrie H, Sherman GG, Bland RM, Rehle TM, et al. The effect of early initiation of antiretroviral treatment in infants on pediatric AIDS mortality in South Africa: A model-based analysis. Pediatr Infect Dis J. 2012 May;31(5):474–80.

151. Granich R, Gupta S, Hersh B, Williams B, Montaner J, Young B, et al. Trends in AIDS Deaths, New Infections and ART Coverage in the Top 30 Countries with the

Highest AIDS Mortality Burden; 1990–2013. Paraskevis D, editor. PLoS One [Internet]. 2015 Jul 6 [cited 2021 Feb 13];10(7):e0131353. Available from: https://dx.plos.org/10.1371/journal.pone.0131353

152. Rodríguez MP, Hayes PR. Reducing HIV prevalence among young people : A review of the UNGASS prevalence goal and how it should be monitored Discussion paper commissioned by WHO October 2002 TABLE OF CONTENTS. Trop Med. 2002;(October).

153. Marsh KA, Nyamukapa CA, Donnelly CA, Garcia-Calleja JM, Mushati P, Garnett GP, et al. Monitoring trends in HIV prevalence among young people, aged 15 to 24 years, in Manicaland, Zimbabwe. J Int AIDS Soc [Internet]. 2011;14(1):27. Available from: http://www.jiasociety.org/content/14/1/27

154. Phimemon RN, Mahande MJ, Ramadhani HO. Factors Associated with Changes in Uptake of HIV Testing among Young Women (age 15-24) in Tanzania from 2003 and 2012. DHS Working Papers No. 119. Rockville, Maryland, USA: ICF International. 2015.

155. Ortblad KF, Musoke DK, Ngabirano T, Salomon JA, Haberer JE, McConnell M, et al. Is knowledge of HIV status associated with sexual behaviours? A fixed effects analysis of a female sex worker cohort in urban Uganda. J Int AIDS Soc. 2019;22(7):1–11.

156. Igulot P, A Magadi M. Social and structural vulnerability to HIV infection in Uganda: A multilevel modelling of AIDS indicators survey data, 2004-2005 and 2011. J Community Med. 2018;1(2):2004–5.

157. Marsh K, Mahy M, Salomon JA, Hogan DR. Assessing and adjusting for differences between HIV prevalence estimates derived from national population-based surveys and antenatal care surveillance, with applications for Spectrum 2013. Aids [Internet]. 2014;28(August):S497–505. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201411004-00011

158. Kilian AHD, Gregson S, Ndyanabangi B, Walusaga K, Kipp W, Sahlmüller G, et al. Reductions in risk behaviour provide the most consistent explanation for declining

HIV-1 prevalence in Uganda. Aids. 1999;13(3):391–8.

159. DREAMS: Partnership to Reduce HIV/AIDS in Adolescent Girls and Young Women | U.S. Agency for International Development [Internet]. [cited 2021 Feb 13]. Available from: https://www.usaid.gov/global-health/health-areas/hiv-and-aids/technical-areas/dreams

Appendix I: Human Research Ethics Clearance WITS

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

R14/49 Mr Joseph Ouma

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M171053

NAME: (Principal Investigator)	Mr Joseph Ouma
DEPARTMENT:	School of Public Health Ministry of Health, Uganda
PROJECT TITLE:	Monitoring HIV/AIDS Dynamics: An Application of Real Time Assessment Methods for Estimating Prevalence at District Levels in Uganda
DATE CONSIDERED:	27/10/2017
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Prof Jonathan Levin
APPROVED BY:	Prof C Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	03/04/2018

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary in Room 301, Third floor, Faculty of Health Sciences, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in October and will therefore be due in the month of October each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix II: Human Research Ethics Clearance UNCST

Uganda National Council for Science and Technology

(Established by Act of Parliament of the Republic of Uganda)

Our Ref: HS 2366 **20th February 2018**

Mr. Joseph Ouma
Principal Investigator
C/o School of Public Health
University of the Witwatersrand
Johannesburg, South Africa

Dear Mr. Ouma,

Re: Research Approval: Monitoring HIV/AIDS Dynamics: An Application of Real Time Assessment Methods for Estimating Prevalence at District Levels in Uganda

I am pleased to inform you that on **06/02/2018**, the Uganda National Council for Science and Technology (UNCST) approved the above referenced research project. The Approval of the research project is for the period of **06/02/2018** to **06/02/2019**.

Your research registration number with the UNCST is **HS 2366**. Please, cite this number in all your future correspondences with UNCST in respect of the above research project.

As Principal Investigator of the research project, you are responsible for fulfilling the following requirements of approval:

1. All co-investigators must be kept informed of the status of the research.
2. Changes, amendments, and addenda to the research protocol or the consent form (where applicable) must be submitted to the designated Research Ethics Committee (REC) or Lead Agency for re-review and approval **prior** to the activation of the changes. UNCST must be notified of the approved changes within five working days.
3. For clinical trials, all serious adverse events must be reported promptly to the designated local IRC for review with copies to the National Drug Authority.
4. Unanticipated problems involving risks to research subjects/participants or other must be reported promptly to the UNCST. New information that becomes available which could change the risk/benefit ratio must be submitted promptly for UNCST review.
5. Only approved study procedures are to be implemented. The UNCST may conduct impromptu audits of all study records.
6. An annual progress report and approval letter of continuation from the REC must be submitted electronically to UNCST. Failure to do so may result in termination of the research project.

Below is a list of documents approved with this application:

	Document Title	Language	Version	Version Date
1.	Research proposal	English	1.0	December 2017

Yours sincerely,

Isaac Makhuwa
For: Executive Secretary
UGANDA NATIONAL COUNCIL FOR SCIENCE AND TECHNOLOGY

Copied to: Chair, Makerere University School of Social Sciences, Research Ethics Committee

LOCATION/CORRESPONDENCE

Plot 6 Kimera Road, Ntinda
P. O. Box 6884
KAMPALA, UGANDA

COMMUNICATION

TEL: (256) 414 705500
FAX: (256) 414-234579
EMAIL: info@uncst.go.ug
WEBSITE: http://www.uncst.go.ug

Appendix II: Letter of permission, MOH

Telephone: General Lines: 256 – 417 - 712260
Permanent Secretary's Office: 256 – 417 - 712221

E-mail: ps@health.go.ug

Website: www.health.go.ug

THE REPUBLIC OF UGANDA

Ministry of Health
P. O. Box 7272
Plot 6, Lourdel Road,
Wandegeya
KAMPALA
UGANDA

IN ANY CORRESPONDENCE ON
THIS SUBJECT PLEASE QUOTE No. **ADM. 130/313/05**

February 23, 2018

Mr. Joseph Ouma,
School of Public Health,
University of Witwatersrand
SOUTH AFRICA

Dear Mr. Ouma

DATA REQUEST FOR PhD STUDIES "MONITORING HIV/AIDS DYNAMICS: AN APPLICATION OF REAL TIME ASSESSMENT METHODS FOR ESTIMATING PREVALENCE AT DISTRICT LEVELS IN UGANDA"

Thank you for showing interest in the data from Facility HIV testing in Uganda and the recently concluded Uganda Population-based HIV Impact Assessment Survey (UPHIA) as per your letter dated 28th Nov 2017 and further addendum dated 12th Jan 2018. The Data Analysis Advisory Committee (DAAC) has reviewed your request and wishes to inform you of the following:

The DAAC is currently reviewing the results of the UPHIA survey and anticipates having the final report ready by July 2018. Thereafter, datasets will be available for public use.

The Facility HIV testing data is currently available and can be accessed upon completion of the data request form available at the Ministry.

You will be able to access both the UPHIA 2016 and the Facility HIV testing data upon presentation of Ethics approval from your University and the Uganda National Council of Science and Technology (UNCST).

We regret any delays this may cause and appreciate your patience.

Yours Sincerely,

Dr. Allan Muruta
FOR: DIRECTOR GENERAL HEALTH SERVICES

Appendix III: Clearance to access UPHIA 2016/17 data

Telephone: General Lines: 256 – 417 – 712260
Permanent Secretary's Office: 256 – 417 – 712221
Toll Free 0800100066
E-mail: ps@health.go.ug
Website: www.health.go.ug

Ministry of Health
P. O. Box 7272
Plot 6, Lourdel Road,
Wandegeya
KAMPALA
UGANDA

REF: ADM.140/269/01

28th October, 2019

Mr. Joseph Ouma,
School of Public Health,
University of Witwatersrand,
Johannesburg,
South Africa.

REQUEST FOR DATA FROM THE 2016/17 UPHIA SURVEY FOR PHD STUDIES "MONITORING HIV/AIDS DYNAMICS: AN APPLICATION OF REAL TIME ASSESSMENT METHODS FOR ESTIMATING HIV PREVALENCE AT DISTRICT LEVELS IN UGANDA".

This is to approve your request to use data from the Uganda population and HIV Impact Assessment survey 2016/17 to investigate methods to obtain HIV prevalence estimates at district levels in Uganda as part of your PhD studies, specifically:

1. To apply and compare small area estimation techniques for estimation of district level HIV prevalence for districts in Uganda.
2. To identify factors for inter and intra-district variations of adult HIV prevalence in Uganda.
3. To assess whether the change in adult HIV prevalence represents change in the risk factors between 2011 and 2016.

We will provide you with data on the following variables from the UPHIA survey: Cluster and household number; Demographics including age, sex, marital status, education level, residence (urban/rural), sub-county, parish, district, region, number of sexual partners in the past 12 months, nearest health facility, testing during antenatal attendance, Human immunodeficiency virus (HIV) status, occupation.

We believe this work will contribute to public health and we do expect Ministry of health and UPHIA survey partners will be acknowledged in all presentations of this work.

We wish you every success in your PhD studies.

Dr. Joshua Musinguzi
PROGRAM MANAGER, AIDS CONTROL PROGRAM/CHAIR UPHIA TECHNICAL WORKING GROUP

Appendix IV: District HIV prevalence estimates allowing for for spatial correlation between the district

Region/ District	Estimate	LCL	UCL
Central 1			
Bukomansimbi	0.091	0.064	0.118
Butambala	0.083	0.054	0.113
Gomba	0.060	0.032	0.089
Kalangala	0.096	0.065	0.126
Kalungu	0.084	0.056	0.112
Lwengo	0.077	0.051	0.103
Lyantonde	0.075	0.046	0.104
Masaka	0.096	0.070	0.122
Mpigi	0.083	0.057	0.109
Rakai	0.082	0.059	0.104
Ssembabule	0.054	0.031	0.076
Wakiso	0.056	0.044	0.068
Central 2			
Buikwe	0.069	0.044	0.093
Buvuma	0.066	0.037	0.095
Kayunga	0.065	0.041	0.089
Kiboga	0.073	0.048	0.098
Kyankwanzi	0.063	0.038	0.088
Luwero	0.075	0.049	0.100
Mityana	0.065	0.042	0.087
Mubende	0.069	0.048	0.090
Mukono	0.065	0.046	0.085
Nakaseke	0.070	0.044	0.097
Nakasongola	0.076	0.046	0.105
East Central			
Bugiri	0.027	0.011	0.044
Busia	0.069	0.047	0.091
Buyende	0.016	0.005	0.026
Iganga	0.034	0.017	0.052
Jinja	0.050	0.033	0.066
Kaliro	0.034	0.013	0.056
Kamuli	0.039	0.022	0.056
Luuka	0.048	0.024	0.072
Mayuge	0.053	0.036	0.071
Namiyango	0.067	0.041	0.094
Namutumba	0.026	0.007	0.044
Eastern			
Amuria	0.051	0.031	0.071
Budaka	0.039	0.015	0.063
Bududa	0.039	0.016	0.062
Bukedea	0.039	0.016	0.062
Bukwo	0.037	0.013	0.060
Bulambuli	0.054	0.030	0.078
Butaleja	0.016	0.000	0.033
Kaberamaido	0.036	0.021	0.052
Kapchorwa	0.049	0.020	0.078
Katakwi	0.066	0.042	0.089
Kibuku	0.032	0.012	0.051
Kumi	0.042	0.021	0.063
Kween	0.063	0.038	0.088
Manafwa	0.030	0.015	0.044
Mbale	0.053	0.039	0.067
Ngora	0.018	0.004	0.031
Pallisa	0.033	0.015	0.051
Serere	0.027	0.010	0.043
Sironko	0.051	0.029	0.072
Soroti	0.042	0.026	0.057
Tororo	0.064	0.046	0.082
Kampala	0.069	0.060	0.079
Karamoja			
Abim	0.056	0.030	0.081
Amudat	0.054	0.026	0.081
Kaabong	0.045	0.026	0.064
Kotido	0.048	0.019	0.077
Moroto	0.020	0.001	0.038
Nakapiripirit	0.007	0.000	0.018
Napak	0.008	0.000	0.020
Northern			
Agago	0.067	0.045	0.090
Alebtong	0.058	0.032	0.084
Amolatar	0.049	0.029	0.069
Amuru	0.044	0.020	0.068
Apac	0.069	0.048	0.090
Dokolo	0.063	0.037	0.088
Gulu	0.080	0.054	0.106

Kitgum	0.061	0.038	0.084
Kole	0.054	0.031	0.077
Lamwo	0.065	0.039	0.092
Lira	0.053	0.032	0.074
Nwoya	0.062	0.034	0.091
Otuke	0.056	0.027	0.084
Oyam	0.059	0.036	0.082
Pader	0.077	0.050	0.104
South Western			
Buhweju	0.053	0.024	0.081
Bushenyi	0.067	0.043	0.092
Ibanda	0.065	0.040	0.089
Isingiro	0.052	0.030	0.075
Kabale	0.066	0.043	0.090
Kanungu	0.076	0.050	0.102
Kiruhura	0.058	0.035	0.080
Kisoro	0.029	0.009	0.049
Mbarara	0.087	0.062	0.112
Mitooma	0.062	0.036	0.088
Ntungamo	0.067	0.044	0.090
Rubirizi	0.079	0.049	0.108
Rukungiri	0.074	0.049	0.099
Sheema	0.082	0.056	0.108

West Nile			
Adjumani	0.033	0.013	0.054
Arua	0.034	0.024	0.044
Koboko	0.027	0.011	0.043
Maracha	0.010	0.000	0.021
Moyo	0.037	0.020	0.054
Nebbi	0.043	0.027	0.058
Yumbe	0.023	0.010	0.036
Zombo	0.040	0.021	0.059
Western			
Buliisa	0.049	0.020	0.078
Bundibugyo	0.016	0.001	0.031
Hoima	0.059	0.038	0.081
Kabarole	0.089	0.063	0.115
Kamwenge	0.035	0.017	0.052
Kasese	0.023	0.010	0.037
Kibaale	0.054	0.037	0.072
Kiryandongo	0.050	0.027	0.073
Kyegegwa	0.062	0.038	0.086
Kyenjojo	0.070	0.046	0.093
Masindi	0.055	0.033	0.076
Ntoroko	0.078	0.048	0.108

Appendix V: Original paper one – AIDS CARE

AIDS CARE
https://doi.org/10.1080/09540121.2020.1734179

Taylor & Francis
Taylor & Francis Group

Difference in HIV prevalence by testing venue: results from population level survey in Uganda

Joseph Ouma [a], Caroline Jeffery [b], Joseph J. Valadez[b], Rhoda K. Wanyenze[c], Jim Todd[d] and Jonathan Levin[a]

[a]Division of Epidemiology and Biostatistics, School of Public Health, University of Witwatersrand, Johannesburg, South Africa; [b]METRe Group, Department of International Health, Liverpool School of Tropical Medicine, Liverpool, UK; [c]Department of Disease Control and Environmental Health, Makerere University School of Public Health, Kampala, Uganda; [d]Department of Population Health, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

ABSTRACT

Growing demand for use of Health Facility (HF) HIV testing data, in addition to other testing data to obtain district level HIV prevalence requires understanding the comparability of these various sources. We analysed the 2011 Uganda AIDS indicator survey data to assess: the proportion of people tested for HIV across Uganda by venue of testing; HIV prevalence ratio for those tested in a HF compared to those tested in community setting; [Katz, D., Baptista, J., Azen, S. P., & Pike, M. C. (1978). Obtaining confidence intervals for the risk ratio in cohort studies. International Biometric Society, 34(3), 469–474. https://doi.org/10.2307/2530610] and factors associated with HIV positivity in each subgroup. Of the 11,685 individuals, 8978 (77.1%) had ever tested for HIV in a HF. Fifty nine per cent tested in a HF in the 12 months preceding the survey (female: 5507, 72.7% versus male: 1413, 34.9%). HIV prevalence ratio was 1.8 times among those tested in a HF compared to those tested at community setting (10.9% [95% CI: 10.0–11.7] versus 6.2% [95% CI: 5.4–7.0]). Among HF testers, older age group, previously married and having no sexual partner was associated with significantly higher HIV prevalence. Using facility testing data for planning and decisions should take into consideration the elevated and varying HIV prevalence among individuals accessing HIV testing services at HFs as well as differences in their social-demographic characteristics.

ARTICLE HISTORY
Received 5 October 2019
Accepted 10 February 2020

KEYWORDS
Population survey; Health Information System; HIV; Prevalence; Testing venue

Introduction

Population sero-surveys and facility-based HIV testing including testing during antenatal care attendance provide the main sources of data for monitoring the HIV/AIDS epidemic. National population surveys are used to generate accurate estimates of population level HIV/AIDS indicators at national or regional levels while antenatal and other health facility-based testing data are used to obtain HIV/AIDS indicator estimates at district level although they possess biases including reporting on only individuals who access health facilities (Eaton et al., 2014; Fabiani et al., 2003; Gregson, S. Dharmayat et al., 2015; Musinguzi et al., 2009; Wilson et al., 2017).

Hybrid approaches use these data sources to complement each other and hence overcome the limitations of using them independently, to obtain HIV prevalence estimates. The use of hybrid approaches however, requires knowledge of the comparability of the data sources. Population-based surveys, contain information on HIV testing including testing venues, that can be analysed to determine how HIV testing data from health facilities can be used to complement data from population survey or other sources to obtain more accurate indicator estimates.

HIV Testing Services (HTS) and approaches have evolved overtime and are broadly categorised as health facility and community-based (Ministry of Health Uganda, 2016; Uganda AIDS Commission, 2017; UNAIDS, September 2016). In Uganda, community-based HIV testing refers to testing offered in homes, social gatherings/events, in educational establishments and at workplaces, to individuals who do not access health facilities whereas health facility-based HIV testing (Provider-Initiated Testing and Counselling (PITC)) is offered at health facilities as part of health care (Ministry of Health Uganda, 2016).

World Health Organization (WHO) developed the first policy to guide health facility based testing in 2007 (World Health Organization, & Joint United Nations Programme on HIV/AIDS, 2007). By 2012, more than 42 (~80%) of the 52 countries in Africa had adopted it in their national HTS policies (Lule et al., 2019). In Uganda, PITC is implemented in all hospitals, Health Centre (HC) IV, HC III and in more than 30% of HC

CONTACT Joseph Ouma oumajosephd@gmail.com or 1608215@students.wits.ac.za

II (Uganda AIDS Commission, 2017). See Appendix 1 for more details regarding health services provided by the different levels of health centres.

Several studies have found higher HIV prevalence and linkage to care among individuals tested at a health facility compared to those tested in a community setting (Govindasamy et al., 2015; Leon et al., 2014; Lugada et al., 2010; Montoy et al., 2016; Roura et al., 2013; Wanyenze et al., 2008, 2009), and others have found similar HIV prevalence estimates for community-based testing during a population surveys and those tested in health facility during antenatal care (Gonese et al., 2010; Judith et al., 2001; Musinguzi et al., 2009; Wilson et al., 2017).

We analysed data from the 2011 UAIS to address two questions. First, what proportion of people have tested for HIV across Uganda, and where have they tested? Second, how does the HIV prevalence compare between those who have tested in a health facility and those who have tested in a community setting. We also investigate factors associated with HIV positivity in each of these subgroups.

Methods

During the UAIS 2011 survey, the country was divided into 10 geographical regions. Survey sample sizes were allocated equally across regions. Clusters were randomly selected from each region with probability proportional to number of households in a cluster (Ministry of Heath and ICF international, 2012). A systematic sample of 25 households was selected from each cluster and all adults present in the selected households who consented to participate in the survey were interviewed and blood drawn for HIV testing (Ministry of Heath and ICF international, 2012). More details about the survey are available from https://dhsprogram.com.

We analysed data of 11,685 individuals aged 15–49 years, who reported having *ever tested* for HIV before the survey.

During the survey, respondents were asked "*Have you ever been tested to see if you have the AIDS virus?*". If they answered "*yes*", additional questions relating to period of the most recent test, and receipt of test results were asked. Women were asked questions relating to antennal attendance; "*Were you offered a test for the AIDS virus as part of your antenatal care?*" If they answered "*yes*", additional questions regarding place of testing and receipt of test results were asked.

Study variables

(1) *HIV status.* HIV testing was conducted for all consenting individuals and results provided to the respondent at home. Only final test results (either 1 = Positive or 0 = Negative) were provided.

(2) *Testing venue (1 = Health facility or 0 = community setting)*:

Tested in a health Facility: Individuals who had ever tested for HIV in a health facility and received test results prior to the survey. This includes women who tested during antenatal care. Health facilities include hospitals; public HC II, III and IV; private hospitals/clinics; and organisations offering HIV/AIDS care and treatment services.

Tested in a community setting: Individuals ever tested during community events, at their homes or work places before the survey. It also includes individuals who were tested in standalone voluntary counselling and testing centres not offering general healthcare.

(3) *Explanatory/independent variables*: area of residence (1 = urban, 2 = rural), gender (1 = male, 2 = female), age group (15–19, 20–29, 30–39 and 40–49-years), marital status (never married, married/cohabiting and previously married-widowed/separated/divorced), highest level of education (none, primary, secondary or higher), number of sexual partners including husband/wife in the 12 months preceding the survey (0, 1 and 2 or more), employment status (employed, not employed), and distance to nearest health facility in kilometres (categorised into <2, 2–5, 5 or more).

Statistical analysis

We computed the proportion tested for HIV; prevalence by venue of testing; and HIV Prevalence Ratio (PR) and the 95% confidence intervals for health facility compared to community testing using the Katz methodology (Appendix 2) (Katz et al., 1978; Koopman, 1984). We also assessed factors associated with HIV positivity in each subgroup using multilevel logistic regression. We further compared HIV prevalence among those tested in a health facility and community; (a) overall and (b) among those tested in the 12 months preceding the survey. All analysis was carried out using Stata release 15 (StataCorp, 2017) and weighted by population sampling weights.

Results

Of the 11685 (female: 7647, 65.1%) individuals, 4789 (41.3%) were aged 20–29 years, 1216 (9.9) % had no formal education, 1375 (11.1%) were previously married and 1097 (9.7%) reported that they had two or more sexual partners in the 12 months preceding the survey.

Of those who tested in a health facility, 6396 (70.9%) were female, 3894 (43.5%) were aged 20–29 years, 6609 (73.7%) were married/cohabiting and 6980 (77.4%) had only one sexual partner. Of the 2707 tested for HIV in a community setting, 1251 (45.3%) were female, 895 (33.9%) were aged 20–29 years, 1560 (59.0%) were married or cohabiting while 1669 (62.8%) had one sexual partner (Table 1).

Prevalence of ever testing by venue of testing

Of the 11,685 individuals, 8978 (77.1%) had tested for HIV in a health facility (female: 6396, 84.0% versus male: 2582, 64.2%) (Appendix 3). Testing in a health facility was higher among individuals aged 20–29 years (3894, 81.2%), no education (1007, 82.3%), married or cohabiting (6609, 80.8%) and among those with one sexual partner in the 12 months preceding the survey (6 980, 80.6%). In contrast, testing in a health facility was lower among males (2582, 64.2%), age group 15–19 years (953, 65.1%), never married (1301, 61.4%) and among those who had no sexual partners in the 12 months preceding the survey (1252, 65.4%) (Appendix 3).

HIV prevalence by venue of testing

HIV prevalence was 9.9% (95% CI: 9.2–10.7) and 5.8% (95% CI: 4.8–6.8) among those tested in a health facility and in a community setting respectively (Table 2). Prevalence among those tested in a health facility was highest in age group 40–49 years (15.0%, 95% CI: 12.9–17.1); previously married 25.7% (95% CI: 22.6, 28.7); and in those who had no sexual partner in the 12 months preceding the survey (15.7%, 95% CI: 13.4–18.1) (Table 2). Prevalence among those tested in a

Table 1. Socio-demographic characteristics of study participants.

	Overall	Tested in health facility	Tested in community
Characteristic	*n* = 11685 (%)	*n* = 8978 (%)	*n* = 2707 (%)
Gender			
Male	4038 (34.9)	2582 (29.1)	1456 (54.7)
Female	7647 (65.1)	6396 (70.9)	1251 (45.3)
Age			
15–19	1498 (12.8)	953 (10.8)	545 (19.5)
20–29	4789 (41.3)	3894 (43.5)	895 (33.9)
30–39	3422 (29.3)	2716 (30.3)	706 (26.1)
40–49	1976 (16.6)	1415 (15.4)	561 (20.5)
Education Level			
No Education	1216 (9.9)	1007 (10.6)	209 (7.7)
Primary	6406 (55.1)	5003 (55.9)	1403 (52.2)
Secondary+	4063 (35.0)	2968 (33.5)	1095 (40.1)
Marital status			
Never married	2159 (18.6)	1301 (14.8)	858 (31.3)
Married/Cohabiting	8169 (70.3)	6609 (73.7)	1560 (59.0)
Previously married	1357 (11.1)	1068 (11.5)	289 (9.8)
Number of sexual partners in 12 months preceding survey			
0	1957 (16.2)	1252 (13.7)	705 (24.5)
1	8649 (74.1)	6980 (77.4)	1669 (62.8)
2+	1079 (9.7)	746 (8.9)	333 (12.8)
Currently working			
No	2894 (24.2)	2219 (24.3)	675 (23.8)
Yes	8791 (75.8)	6759 (75.7)	2032 (76.2)
Distance to nearest HF			
<2	2441 (21.7)	1920 (22.2)	521 (20.0)
2–5	4823 (40.7)	3696 (40.6)	1127 (40.8)
5+	4009 (33.7)	3040 (33.3)	969 (35.2)
Don't Know	412 (3.9)	322 (3.9)	90 (3.9)
Area of residence			
Rural	8870 (75.8)	6845 (76.0)	2025 (75.4)
Urban	2815 (24.1)	2133 (24.0)	682 (24.6)
Region			
Central 1	1151 (12.2)	889 (12.3)	262 (12.2)
Central 2	1177 (10.6)	861 (10.0)	316 (12.8)
Kampala	1464 (9.0)	1068 (8.4)	396 (11.0)
East Central	1013 (8.8)	722 (8.2)	291 (10.5)
Mid-Eastern	901 (7.6)	732 (8.0)	169 (6.3)
North East	1134 (9.2)	875 (9.0)	259 (10.0)
West Nile	1220 (6.4)	896 (6.0)	324 (7.6)
Mid Northern	1394 (12.5)	1107 (12.9)	287 (11.3)
South Western	1024 (11.3)	853 (12.2)	171 (8.0)
Mid-Western	1207 (12.4)	975 (12.9)	232 (10.5)

Note: Percentages are weighted by population survey weights.

Table 2. HIV prevalence distribution overall and by venue of testing.

			Ever tested for HIV in			
	Overall (*N* = 11685)		Health facility		Community	
Characteristic	*n*#	Weighted prevalence (95% CI)	*n*#	Weighted prevalence (95% CI)	*n*#	Weighted prevalence (95% CI)
Total	1002	9.0 (8.4, 9.6)	850	9.9 (9.2, 10.7)	152	5.8 (4.8, 6.8)
Gender						
Male	300	7.8 (6.8, 8.8)	227	9.4 (8.1, 10.8)	73	4.9 (3.7, 6.1)
Female	702	9.6 (8.9, 10.4)	623	10.2 (9.3, 11.0)	79	6.9 (5.2, 8.5)
Age						
15–19	50	3.6 (2.5, 4.8)	42	4.7 (3.1, 6.3)	8	1.7 (0.5, 2.9)
20–29	314	6.9 (6.0, 7.7)	268	7.2 (6.3, 8.2)	46	5.4 (3.7, 7.2)
30–39	382	11.9 (10.6, 13.1)	334	13.2 (11.7, 14.7)	48	6.7 (4.7, 8.8)
40–49	256	13.3 (11.6, 15.0)	206	15.0 (12.9, 17.1)	50	9.0 (6.5, 11.6)
Education Level						
No Education	123	10.4 (8.6, 12.3)	112	11.6 (9.4, 13.8)	11	5.0 (1.9, 8.0)
Primary	614	10.1 (9.3, 11.0)	523	11.0 (10.0, 12.0)	91	6.9 (5.3, 8.4)
Secondary	265	6.8 (5.9, 7.7)	215	7.6 (6.5, 8.8)	50	4.5 (3.2, 5.9)
Marital status						
Never married	89	4.5 (3.4, 5.6)	69	5.6 (4.2, 7.1)	20	2.7 (1.3, 4.2)
Married/Living together	622	7.9 (7.2, 8.6)	529	8.4 (7.6, 9.1)	93	6.0 (4.7, 7.3)
Previously married	291	23.4 (20.8, 26.0)	252	25.7 (22.6, 28.7)	39	14.3 (9.9, 18.8)
Number of sexual partners in 12 months preceding survey						
0	219	11.9 (11.2, 13.5)	189	15.7 (13.4, 18.1)	30	4.5 (2.8, 6.3)
1	672	8.1 (7.5, 8.8)	580	8.7 (7.9, 9.5)	92	5.8 (4.5, 7.1)
2+	111	10.7 (8.6, 12.8)	81	11.8 (9.0, 14.5)	30	8.2 (5.1, 11.3)
Currently working						
No	205	7.2 (6.1, 8.3)	178	8.1 (6.8, 9.4)	27	4.1 (2.5, 5.8)
Yes	797	9.6 (8.9, 10.3)	672	10.5 (9.7, 11.4)	125	6.3 (5.1, 7.5)
Distance to nearest HF						
<2	222	9.3 (7.9, 10.7)	192	10.1 (8.5, 11.7)	30	6.4 (3.7, 9.1)
2–5	374	8.1 (7.3, 9.0)	316	9.0 (8.0, 10.1)	58	5.2 (3.8, 6.6)
5+	375	9.8 (8.8, 10.8)	315	10.9 (9.7, 12.1)	60	6.3 (4.6, 7.9)
Don't Know	31	9.0 (5.2, 12.9)	27	10.5 (5.7, 15.3)	4	4.0 (0.0, 8.1)
Area of residence						
Rural	737	8.7 (8.1, 9.4)	624	9.6 (8.8, 10.3)	113	5.9 (4.7, 7.0)
Urban	265	9.9 (8.5, 11.3)	226	11.2 (9.5, 12.9)	39	5.5 (3.4, 7.7)
Region						
Central 1	144	13.0 (10.7, 15.2)	123	14.3 (11.6, 16.9)	21	8.5 (4.4, 12.5)
Central 2	116	9.9 (8.1, 11.7)	99	11.4 (9.1, 13.7)	17	6.0 (3.1, 8.9)
Kampala	116	7.8 (6.2, 9.3)	94	9.0 (7.0, 11.0)	22	4.4 (2.4, 6.4)
East Central	65	7.1 (5.4, 8.9)	52	8.0 (5.8, 10.2)	13	5.0 (2.2, 7.8)
Mid-Eastern	50	5.4 (3.9, 6.9)	41	5.3 (3.7, 7.0)	9	5.6 (2.0, 9.2)
North East	84	6.7 (5.0, 8.4)	71	7.7 (5.5, 9.8)	13	3.6 (1.5, 5.7)
West Nile	67	6.0 (4.5, 7.5)	54	6.5 (4.7, 8.3)	13	4.8 (2.2, 7.4)
Mid Northern	133	9.7 (7.9, 11.5)	124	11.5 (9.3, 13.7)	9	2.9 (0.9, 4.9)
South Western	110	10.5 (8.5, 12.4)	94	10.9 (8.7, 13.1)	16	8.3 (4.2, 12.4)
Mid-Western	117	9.9 (8.0, 11.7)	98	10.2 (8.1, 12.2)	19	8.6 (4.7, 12.5)

– Number HIV positive, Denominators are presented in Table 1. HIV prevalence estimates are weighted by population survey weights.

community setting was higher among previously married and those who had 2 or more sexual partners in the 12 months preceding the survey (8.2%, 95% CI: 5.1–11.3).

HIV prevalence by testing venue among those who tested in the 12 months preceding survey

Of the 11,685 individuals included in the analysis, 6920 (59.2%) tested in a health facility in the 12 months preceding the survey (female: 5507, 72.7% versus male: 1413, 34.9%). HIV prevalence among health facility testers was 10.9% (95% CI: 10.0–11.7) compared to 6.2% (95% CI: 5.4–7.0) among individuals tested in a community setting (Table 3).

HIV PR was 1.8 (95% CI: 1.6, 2.1) for those tested in a health facility compared to those tested in a community setting. PR was 3.8 (95% CI: 2.4, 6.0) for those who had no sexual partner in the 12 months preceding the survey; 2.7 (95% CI: 1.9, 3.9) among males; 3.2 (95% CI: 1.3, 8.0) for age group 15–19 years; 2.4 (95% CI: 1.3, 4.5) for those with no education; and 2.2 (95% CI: 1.2, 4.2) for never married (Table 3).

Among health facility testers, the odds of testing HIV positive during the survey was significantly higher in age group 40–49 years compared to 20–29 years (aOR; 2.92, 95% CI: 2.28–3.73); previously married versus married/cohabiting (aOR; 3.25, 95% CI: 2.66, 4.12); and for those who had no sexual partner versus those who had one (aOR; 1.55, 95% CI: 1.20, 2.00). HIV positivity was significantly lower among females compared to males (aOR: 0.78, 95% CI: 0.62, 0.97); age group 15–19 compared 20–29 years

Table 3. Proportion tested for HIV in a health facility in the 12 months preceding the survey and HIV prevalence by testing venue.

	Tested in health facility in 12 months preceding the survey				Tested in community setting		
Characteristic	Number tested	(%)	*n#*	Weighted prevalence (95% CI)	*n#*	Weighted prevalence (95% CI)	HIV prevalence ratio (95% CI)*
Total	6920	59.0	722	10.9 (10.0, 11.7)	100	6.2 (5.4, 7.0)	1.8 (1.4, 2.2)
Gender							
Male	1413	34.9	169	12.6 (10.6, 14.7)	37	4.6 (3.0, 6.2)	2.7 (1.9, 3.9)
Female	5507	72.7	553	10.4 (9.5, 11.3)	63	7.8 (5.7, 9.9)	1.3 (1.0, 1.7)
Age							
15–19	724	50.1	37	5.4 (3.5, 7.4)	5	1.7 (1.9, 3.3)	3.2 (1.3, 8.0)
20–29	3221	67.1	225	7.2 (6.2, 8.2)	35	6.7 (4.2, 9.2)	1.1 (0.8, 1.5)
30–39	2109	62.0	286	14.7 (12.9, 16.4)	30	6.3 (3.9, 8.6)	2.3 (1.6, 3.3)
40–49	866	43.4	174	20.3 (17.4, 23.2)	30	10.3 (6.5, 14.0)	2.0 (1.4, 2.8)
Education level							
No Education	782	63.7	98	13.3 (10.7, 15.9)	10	5.5 (1.9, 9.1)	2.4 (1.3, 4.5)
Primary	3917	61.3	446	12.1 (10.9, 13.2)	60	7.5 (5.4, 9.5)	1.6 (1.2, 2.1)
Secondary	2221	55.5	178	8.0 (6.7, 9.3)	30	4.8 (3.0, 6.6)	1.7 (1.1, 2.4)
Marital status							
Never married	891	42.5	54	6.4 (4.5, 8.2)	11	2.9 (0.8, 5.1)	2.2 (1.2, 4.2)
Married/Cohabiting	5204	63.7	445	8.8 (7.9, 9.7)	67	6.8 (5.1, 8.5)	1.3 (1.0, 1.7)
Previously married	825	61.5	223	29.7 (26.1, 33.3)	22	13.5 (7.8, 19.1)	2.2 (1.5, 3.3)
Number of sexual partners in 12 months preceding survey							
0	872	45.1	166	20.2 (17.1, 23.4)	20	5.3 (2.9, 7.8)	3.8 (2.4, 6.0)
1	5585	64.8	499	9.3 (8.4, 10.1)	60	6.3 (4.5, 8.1)	1.5 (1.1, 1.9)
2+	463	43.2	57	12.8 (9.4, 16.3)	19	7.7 (4.2, 11.3)	1.7 (1.0, 2.7)
Currently working							
No	1820	64.3	158	8.7 (7.2, 10.2)	21	3.9 (2.1, 5.8)	2.2 (1.4, 3.5)
Yes	5100	58.0	564	11.6 (10.6, 12.6)	79	7.1 (5.4, 8.8)	1.6 (1.3, 2.0)
Distance to nearest HF							
<2	1493	62.0	171	11.3 (9.4, 13.2)	18	6.2 (2.6, 9.7)	1.8 (1.1, 2.9)
2–5	2852	59,3	256	9.4 (8.1, 10.6)	34	5.5 (3.5, 7.4)	1.7 (1.2, 2.4)
5+	2341	58.5	275	12.5 (11.0, 14.0)	44	7.2 (5.0, 9.4)	1.7 (1.3, 2.4)
Don't Know	234	55.8	20	10.0 (4.8, 15.2)	4	6.0 (0.1, 11.8)	1.7 (0.6, 4.7)
Area of residence							
Rural	5284	59.6	530	10.5 (9.6, 11.4)	77	6.3 (4.8, 7.7)	1.7 (1.3, 2.1)
Urban	1636	59.2	192	12.1 (10.1, 14.1)	23	6.2 (2.9, 9.5)	2.0 (1.3, 3.0)
Region							
Central 1	662	57.5	95	14.8 (11.7, 17.9)	14	9.6 (4.0, 15.2)	1.5 (0.9, 2.6)
Central 2	713	60.2	81	11.0 (8.5, 13.5)	11	6.1 (2.3, 9.9)	1.8 (1.0, 3.3)
Kampala	803	54.1	81	10.3 (7.8, 12.7)	13	5.1 (2.1, 8.1)	2.0 (1.1, 3.6)
East Central	547	55.0	48	9.8 (7.0, 12.5)	8	4.8 (1.4, 8.3)	2.0 (1.0, 4.2)
Mid-Eastern	544	60.5	31	5.4 (3.5, 7.3)	10	7.1 (2.8, 11.4)	0.8 (0.4, 1.5)
North East	711	63.4	62	7.6 (5.4, 9.8)	11	7.5 (2.9, 12.0)	1.0 (0.5, 1.9)
West Nile	680	54.8	47	7.6 (5.3, 9.8)	9	5.3 (1.8, 8.8)	1.4 (0.7, 2.9)
Mid Northern	886	62.8	106	12.5 (9.8, 15.1)	6	2.6 (0.4, 4.9)	4.8 (2.1, 10.8)
South Western	639	62.5	86	13.4 (10.6, 16.2)	9	9.7 (3.2, 16.2)	1.4 (0.7, 2.6)
Mid-Western	735	60.8	85	11.6 (9.1, 14.1)	9	6.3 (2.1, 10.6)	1.8 (0.9, 3.6)

– number HIV positive. *HIV prevalence ratio comparing health facility and community testers.

(aOR: 0.67, 95% CI: 0.45, 0.99); secondary versus primary education (aOR: 0.64, 95% CI: 0.52, 0.80) and rural versus urban residence (aOR; 0.59, 95% CI: 0.43, 0.80) (Table 4).

For those tested in a community setting, the odds of testing HIV positive was significantly higher for females compared to males (aOR; 1.62, 95% CI: 1.21, 2.15); previously married versus married/cohabiting (aOR: 2.05, 95% CI: 1.44, 2.92); and among those who had two or more sexual partners compared to those who had one (aOR; 1.95, 95% CI: 1.37, 2.76). The likelihood of a positive test result was significantly lower among those aged 15–19 years versus age group 20–29 years (aOR: 0.34, 95% CI: 0.17, 0.70) (Table 4). Further analysis comparing HIV prevalence among those tested in a health facility in 12 months preceding the survey with those tested in a community setting are presented in Appendix 4.

Prevalence of testing for HIV in a health facility in the 12 months preceding survey

Prevalence of testing for HIV in a health facility in the 12 months preceding the survey was higher among females (5507, 72.7%); age group 20–29 years (3221, 67.1%); married/cohabiting (5204, 63.7%); and in those who reported one sexual partner in the 12 months preceding the survey (5585, 64.8%). Testing was lower among males (1413, 34.9%); age group 15–19 years (724, 50.1%); age group 40–49 (866, 43.4%); never married (891, 42.5%); and among those who had two or more

Table 4. Factors associated with HIV prevalence by venue of testing.

	Overall				Tested in health facility		Tested in community	
Characteristic	cOR	(95% CI)	aOR	(95% CI)	aOR	(95% CI)	aOR	(95% CI)
Gender (Reference group: Male)								
Female	1.25	(1.09, 1.44)*	1.21	(1.03, 1.43)*	0.78	(0.62, 0.97)*	1.62	(1.21, 2.15)*
Age (Reference group: 20–29)								
15–19	0.48	(0.35, 0.66)*	0.53	(0.38, 0.75)*	0.67	(0.45, 0.99)**	0.34	(0.17, 0.70)*
30–39	1.86	(1.58, 2.18)*	1.70	(1.43, 2.03)*	1.96	(1.60, 2.42)*	1.24	(0.89, 1.74)
40–49	2.28	(1.90, 2.72)*	1.89	(1.55, 2.31)*	2.92	(2.28, 3.73)*	1.17	(0.80, 1.69)
Education Level (Reference group: Primary)								
No Education	1.09	(0.86, 1.38)	0.86	(0.68, 1.10)	0.91	(0.69, 1.19)	0.70	(0.41, 1.19)
Secondary	0.61	(0.52, 0.72)*	0.72	(0.60, 0.86)*	0.64	(0.52, 0.80)*	0.85	(0.61, 1.18)
Marital status (Reference Group: Married/Living together)								
Never married	0.49	(0.39, 0.63)*	0.80	(0.58, 1.1)	0.93	(0.64, 1.37)	0.74	(0.43, 1.27)
Previously married	3.33	(2.81, 3.96)*	2.76	(2.27, 3.35)*	3.25	(2.66, 4.12)*	2.05	(1.44, 2.92)*
Number of sexual partners in 12 months preceding survey (Reference Group: 1)								
0	1.52	(1.27, 1.81)*	1.32	(1.07, 1.63)*	1.55	(1.20, 2.00)*	1.01	(0.67, 1.53)
2+	1.39	(1.11, 1.74)*	1.43	(1.12, 1.82)*	1.07	(0.77, 1.50)	1.95	(1.37, 2.76)*
Currently working (Reference Group: Not employed)								
Employed	1.34	(1.13, 1.61)*	1.02	(0.84, 1.24)	1.03	(0.83, 1.28)	0.99	(0.65, 1.50)
Distance to nearest HF (Reference Group: <2Km)								
2–5	0.80	(0.66, 0.98)**						
5+	1.02	(0.83, 1.25)						
Don't Know	0.79	(0.53, 1.19)						
Area of residence (Reference Group: Urban)								
Rural	0.84	(0.69, 1.03)	0.62	(0.47, 0.82)*	0.59	(0.43, 0.80)*	0.75	(0.46, 1.23)
Region (Reference Group: Central 1)								
Central 2	0.75	(0.56, 1.02)	0.74	(0.53, 1.02)	0.75	(0.51, 1.10)	0.71	(0.41, 1.21)
Kampala	0.60	(0.43, 0.84)*	0.49	(0.32, 0.75)*	0.53	(0.33, 0.86)*	0.47	(0.23, 0.93)*
East Central	0.45	(0.31, 0.66)*	0.46	(0.31, 0.66)*	0.58	(0.37, 0.90)*	0.30	(0.15, 0.60)*
Mid-Eastern	0.40	(0.25, 0.64)*	0.44	(0.27, 0.71)*	0.37	(0.23, 0.62)*	0.53	(0.24, 1.14)
North East	0.54	(0.38, 0.77)*	0.58	(0.40, 0.82)*	0.56	(0.37, 0.84)*	0.53	(0.30, 0.95)*
West Nile	0.38	(0.26, 0.56)*	0.36	(0.23, 0.56)*	0.38	(0.22, 0.64)*	0.33	(0.18, 0.63)*
Mid Northern	0.72	(0.52, 1.01)	0.80	(0.56, 1.13)	0.78	(0.53, 1.16)	0.59	(0.33, 1.11)
South western	0.82	(0.60, 1.14)	0.87	(0.62, 1.22)	0.93	(0.62, 1.38)	0.65	(0.38, 1.14)
Mid-western	0.73	(0.53, 102)	0.78	(0.55, 1.11)	0.82	(0.54, 1.24)	0.69	(0.40, 1.17)

*p-value<0.01, **P-value<0.05, cOR- unadjusted Odds Ratio, aOR-Adjusted Odds Ratio.

sexual partners in the 12 months preceding the survey (463, 43.2%) (Table 3).

Discussion

We analysed data from 2011 UAIS, to estimate the proportion of individuals ever tested for HIV; prevalence by testing venue; HIV PR for those tested in a health facility to those tested in a community setting; and assessed factors associated with HIV positivity for each subgroup.

Findings show that 59.0% of the respondents tested for HIV in a health facility in the 12 months preceding the survey. Overall HIV PR was 1.8. PR was 2 or more times higher among males; age groups 15–19, 30–39 and 40–49 years; those with no education; never or previously married; and among those who had no sexual partner in the 12 months preceding the survey.

Higher HIV prevalence among health facility compared to community-based testing found in this study is consistent with findings elsewhere (Chirawu et al., 2010; Govindasamy et al., 2015; Lugada et al., 2010; Sharma et al., 2015; Silvestri et al., 2011). For example, in a rural community in East Central Uganda, HIV prevalence was 2.7 (17.3% vs. 7.1%) times higher in the clinic-based compared to home-based arm (Lugada et al., 2010) while in Zimbabwe, positivity rates were 32.9% and 18.8% for clinic and community-based arms respectively (Chirawu et al., 2010). These findings demonstrate higher HIV prevalence (~2-fold) in health facility testing compared to prevalence based on HIV testing from a community setting.

Furthermore, HIV prevalence was 3.6% (10.9% vs 7.3%) higher and 1.9% (5.4% vs 7.3%) lower among individuals tested in health facility and those tested in a community setting respectively compared to overall prevalence in the general population (Ministry of Heath and ICF international, 2012).

Higher HIV prevalence among health facility compared to community-based testing have been attributed to ill health among individuals accessing health facilities in many studies (Chirawu et al., 2010; Sharma et al., 2016; Silvestri et al., 2011; Ssebunya et al., 2018). Other studies attributed higher HIV prevalence among health facility testers to a desire by individuals to know status after a risky sexual behaviour (Ssebunya et al., 2018).

In this study, we observed similar HIV prevalence (PR=1.1) for those tested at a health facility compared to those tested at community setting for the age group 20–29 years. This age group comprise mainly newly married individuals who have a desire to have children and

are therefore more likely to visit a health facility for antenatal or child health reasons and not a recent high-risk sexual behaviour or ill health. Similarly, HIV PR for married/cohabiting individuals and those who had one sexual partner was close to 1, (i.e., 1.3, and 1.5 respectively). These population sub groups are more likely to be in stable sexual/family relationship and therefore are more likely to visit a health facility for antenatal or family health reasons and thus less likely to test HIV positive. Additionally, in sub-Saharan Africa, caring for sick family members are often delegated to women who attend to/take care of sick family members seeking health care. This phenomenon may explain the PR (~1), observed among females who had tested at health facilities compared to those tested at a community setting.

Several studies have also found lower fertility among HIV positive women compared to their HIV negative counterparts (Gregson et al., 2002; Zaba et al., 2000). This implies that women who visit health facilities for antenatal reasons are less likely to test HIV positive, similar to findings in this study.

Additionally, this study found higher HIV PR among never-married individuals; previously married; and those with no sexual partners tested at health facilities compared those tested at a community setting. This may be attributed to ill health or a recent risky sexual behaviour in these population subgroups as noted above.

Regarding factors associated with HIV positivity, similar factors were found for those tested in a health facility and those testing in a community setting for all socio-demographic characteristics except gender. Among those tested in a health facility, HIV positivity was significantly lower among females compared to males. However, positivity was significantly higher among females tested in a community setting compared to males consistent with findings elsewhere (Amornkul et al., 2009; Kwesigabo et al., 2000; World Health Organization, 2015; Zaba et al., 2000). Lower positivity rates among females tested at health facilities compared to males may be attributed to males seeking care at health facilities due to ill health, while women may visit health facilities for antenatal or other family health reasons as noted above.

We also observed higher HIV positivity rates among individuals who reported no sexual partner compared to those who had only one partner in the 12 months preceding the survey irrespective of the venue of testing. Higher HIV positivity rates in this subgroup may be attributed to concealment of existing or earlier sexual relationships consistent with a study in Uganda, that found never and previously married individuals had more sexual partners compared to those who were married (Nalugoda et al., 2014).

Although prevalence of health facility testing was generally low at 59%, testing was lower among males, individuals who had two or more sexual partners, those with primary or lower level of education consistent with findings from other studies (Amornkul et al., 2009; Manda et al., 2015; Mekonnen et al., 2015; Mtenga et al., 2015; Mtowa et al., 2017; Muyunda et al., 2018; Nalugoda et al., 2014). The proportion of males tested in a health facility was only 34.9% compared to 72.7% of females. In Uganda, there is almost universal HCT coverage among women seeking antenatal and postnatal services (>95%) (Uganda Bureau of Statistics (UBOS) and ICF International Inc, 2012), accounting for higher testing rates among females compared males found in this study. We also observed lower testing rates in age groups 15–19years and 40–49 years consistent with other studies in Sub-Saharan Africa (Staveteig et al., 2017). The low testing rates in older age groups has been attributed to reluctance by health workers to prioritise testing for older age groups and poor health seeking behaviour among youths and older persons (Kiplagat, 2018). Other well-documented barriers for HCT access by men and youth include fear to find out ones test results; avoiding to divulge personal information to health workers (Facha et al., 2016; Mohlabane et al., 2016).

Limitations

This study used population survey data, many limitations including, recall of information such as, when and where the last HIV test was conducted may be influenced by the respondent's personal experiences during the test and whether the experience was desirable. Individuals may prefer not to report negative experiences such as positive HIV test result during a previous test. Reasons for testing for HIV at a health facility or community setting were not captured during the survey.

Conclusions

We found higher HIV prevalence among individuals who tested in a health facility compared to those tested in a community setting. HIV PR was more than a two-fold in males; age groups 15–19 and 40–49 years; never married and those who had no sexual partners in the 12 months preceding the survey while prevalence of facility testing was lower in these population subgroup groups. Higher HIV prevalence among specific population subgroups accessing health facilities and the low testing rates in those population subgroups call for continuous review of Health Management Information System (HMIS) data to inform scaling up of HIV testing

interventions. Additionally, HMIS data comprise data from public health facilities only, excluding private health facilities accessed by wealthier and more educated individuals whose HIV prevalence rates may be different from that of the general population.

Acknowledgements

We acknowledge Ministry of Health Uganda, ICF Macro International and their partners in conducting the 2011 Uganda AIDS Indicator Survey, and for the permission to use the dataset.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported through the Wellcome Trust, grant 107754/Z/15/Z-DELTAS Africa via Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA). The views expressed in this publication are those of the authors and not necessarily those of the AAS, NEPAD Agency, Wellcome Trust or the UK government.

Ethics

Ethical clearance was obtained from the University of Witwatersrand Human Research Ethics Committee (HREC) and Uganda National Council of Science and Technology (UNCST). Permission was also obtained from Ministry of Health, Uganda, and the Measure DHS (ORC MACRO).

Data availability statement

Data used for this study is publicly available from Measure DHS: https://dhsprogram.com.

ORCID

Joseph Ouma http://orcid.org/0000-0001-5249-7440
Caroline Jeffery http://orcid.org/0000-0002-8023-0708

References

Amornkul, P. N., Vandenhoudt, H., Nasokho, P., Odhiambo, F., Mwaengo, D., Hightower, A., Buvé, A., Misore, A., Vulule, J., Vitek, C., Glynn, J., Greenberg, A., Slutsker, L., De Cock, K. M., & Hernandez, A. V. (2009). HIV prevalence and associated risk factors among individuals aged 13–34 years in rural Western Kenya. *PLoS ONE*, *4*(7), https://doi.org/10.1371/journal.pone.0006470

Chirawu, P., Langhaug, L., Mavhu, W., Pascoe, S., Dirawo, J., & Cowan, F. (2010). Acceptability and challenges of implementing voluntary counselling and testing (VCT) in rural Zimbabwe: Evidence from the Regai Dzive Shiri Project. *AIDS Care*, *22*(1), 81–88. https://doi.org/10.1080/09540120903012577

Eaton, J. W., Rehle, T. M., Jooste, S., Nkambule, R., Kim, A. A., Mahy, M., & Hallett, T. B. (2014). Recent HIV prevalence trends among pregnant women and all women in sub-Saharan Africa: Implications for HIV estimates. *Wolters Kluwer Health*, *28*, 507–514. https://doi.org/10.1097/QAD.0000000000000412

Fabiani, M., Fylkesnes, K., Nattabi, B., Ayella, E. O., & Declich, S. (2003). Evaluating two adjustment methods to extrapolate HIV prevalence from pregnant women to the general female population in sub-Saharan Africa. *AIDS*, *17*(3), 399–405. https://doi.org/10.1097/01.aids.0000042976.95433.7e

Facha, W., Kassahun, W., & Workicho, A. (2016). Predictors of provider- initiated HIV testing and counseling refusal by outpatient department clients in Wolaita zone, Southern Ethiopia: A case control study. *BMC Public Health*, 1–8. https://doi.org/10.1186/s12889–016-3452-8

Gonese, E., Dzangare, J., Gregson, S., Jonga, N., Mugurungi, O., & Mishra, V. (2010). Comparison of HIV prevalence estimates for Zimbabwe from antenatal clinic surveillance (2006) and the 2005–06 Zimbabwe demographic and health survey. *PLoS ONE*, *5*(11), 1–7. https://doi.org/10.1371/journal.pone.0013819

Govindasamy, D., Ferrand, R. A., Wilmore, S. M., Ford, N., Ahmed, S., Afnan-Holmes, H., & Kranzer, K. (2015). Uptake and yield of HIV testing and counselling among children and adolescents in sub-Saharan Africa: A systematic review. *Journal of the International AIDS Society*, *18*(1), 20182. https://doi.org/10.7448/IAS.18.1.20182

Gregson, S., Dharmayat, K., Pereboom, M., Takaruza, A., Mugurungi, O., Schur, N., & Nyamukapa, C. A. (2015). Do HIV prevalence trends in antenatal clinic surveillance represent trends in the general population in the antiretroviral therapy era? The case of Manicaland, East Zimbabwe. *AIDS*, *29*(14), 1845–1853. https://doi.org/10.1097/QAD.0000000000000754

Gregson, S., Terceiria, N., Kakowa, M., Mason, P. R., Anderson, R. M., & Chandiwana SK, C. M. (2002). Study of bias in antenatal clinic HIV-1 surveillance data in a high contraceptive prevalence population in sub-Saharan Africa. *AIDS*, *16*(4), 643–652. https://doi.org/10.1097/00002030-200203080-00017

Judith, R. G., Anne, B., Michel, C., Rosemary, M. M., Maina, K., Isaac, M., & Francis T, L. Z. (2001). Factors influencing the difference in HIV prevalence between antenatal clinic and general population in sub-Saharan Africa. *AIDS*, *15*, 1717–1725. https://doi.org/10.1097/00002030-200109070-00016

Katz, D., Baptista, J., Azen, S. P., & Pike, M. C. (1978). Obtaining confidence intervals for the risk ratio in cohort studies. *International Biometric Society*, *34*(3), 469–474. https://doi.org/10.2307/2530610

Kiplagat, J. (2018). HIV testing and counselling experiences: A qualitative study of older adults living with HIV in western

Kenya. *BMC Geriatrics*, *18*(257), 1–10. https://doi.org/10.1186/s12877-018-0941-x.

Koopman, P. A. (1984). Confidence intervals for the ratio of two binomial proportions. *International Biometric Society*, *40*(2), 513–517. https://doi.org/10.2307/2531405

Kwesigabo, G., Killewo, J. Z. J., Urassa, W., Mbena, E., Mhalu, F., Lugalla, J. L. P., Godoy, C., Biberfeld, G., Emmelin, M., Wall, S., & Sandstrom, A. (2000). Monitoring of HIV-1 infection prevalence and trends in the general population using pregnant women as a sentinel population: 9 years experience from the Kagera region of Tanzania. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *23*(5), 410–417. https://doi.org/10.1097/00126334-200004150-00008

Leon, N., Mathews, C., Lewin, S., Osler, M., Boulle, A., & Lombard, C. (2014). A comparison of linkage to HIV care after provider-initiated HIV testing and counselling (PITC) versus voluntary HIV counselling and testing (VCT) for patients with sexually transmitted infections in Cape Town, South Africa. *BMC Health Services Research*, *14*(1), 350. https://doi.org/10.1186/1472-6963-14-350

Lugada, E., Levin, J., Abang, B., Mermin, J., Mugalanzi, E., Namara, G., Gupta, S., Grosskurth, H., Jaffar, S., Coutinho, A., & Bunnell, R. (2010). Comparison of home and clinic-based HIV testing among household members of persons taking antiretroviral therapy in Uganda: results from a randomized trial. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *55*(2), 245–252. https://doi.org/10.1097/QAI.0b013e3181e9e069

Lule, F., Granich, R., & Hargreaves, J. (2019). Bulletin of the World Health Organization from caution to urgency: The evolution of HIV testing and counselling in Africa. *Bulletin of the World Health Organization*, 1–11. https://doi.org/10.2471/BLT.11.100818

Manda, S., Masenyetse, L., Cai, B., & Meyer, R. (2015). Mapping HIV prevalence using population and antenatal sentinel-based HIV surveys: A multi-stage approach. *Population Health Metrics*, *13*, 22. https://doi.org/10.1186/s12963-015-0055-z

Mekonnen, Z. A., Lerebo, W. T., Gebrehiwot, T. G., & Abadura, S. A. (2015). Multilevel analysis of individual and community level factors associated with institutional delivery in Ethiopia. *BMC Research Notes*, 1–9. https://doi.org/10.1186/s13104-015-1343-1

Ministry of Health Uganda. (2016). *National HIV testing services policy and implementation guidelines*. https://health.go.ug/sites/default/files/UGANDA HTS POLICY AND IMPLEMENTATION GUIDELINES 2016 %282%29.pdf

Ministry of Heath and ICF International. (2012). *Uganda AIDS indicator survey (AIS) 2011*. https://doi.org/AIS10

Mohlabane, N., Tutshana, B., & Peltzer, K. M. A. (2016). Barriers and facilitators associated with HIV testing uptake in South Africa health facilities offering HIV Counseling and testing. *Health SA GESONDHEID*, *21*, 86–95. https://doi.org/

Montoy, J. C. C., Dow, W. H., & Kaplan, B. C. (2016). Patient choice in opt-in, active choice, and opt-out HIV screening: Randomized clinical trial. *BMJ (Online)*, *352*, https://doi.org/10.1136/bmj.h6895

Mtenga, S. M., Pfeiffer, C., Merten, S., Mamdani, M., Exavery, A., Haafkens, J., Tanner, M., & Geubbels, E. (2015). Prevalence and social drivers of HIV among married and cohabitating heterosexual adults in south-eastern Tanzania: Analysis of adult health community cohort data. *Global Health Action*, *1*, 1–10. https://doi.org/10.3402/gha.v8.28941

Mtowa, A., Gerritsen, A. A. M., Mtenga, S., Mwangome, M., & Geubbels, E. (2017). Socio-demographic inequalities in HIV testing behaviour and HIV prevalence among older adults in rural Tanzania, 2013. *AIDS Care*, *29*(9), 1162–1168. https://doi.org/10.1080/09540121.2017.1308462

Musinguzi, J., Kirungi, W., Opio, A., Montana, L., Mishra, V., Madraa, E., Biryahwaho, B., Mermin, J., Bunnell, R., Cross, A., Hladik, W., McFarland, W., & Stoneburner, R. (2009). Comparison of HIV prevalence estimates from sentinel surveillance and a national population-based survey in Uganda, 2004-2005. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *51*(1), 78–84. https://doi.org/10.1097/QAI.0b013e3181990713

Muyunda, B., Musonda, P., Mee, P., Todd, J., & Michelo, C. (2018). Educational attainment as a predictor of HIV testing uptake among women of child-bearing age: Analysis of 2014 demographic and health survey in Zambia. *Frontiers in Public Health*, *6*, 192. https://doi.org/10.3389/fpubh.2018.00192

Nalugoda, F., Guwatudde, D., Bwaninka, J. B., Makumbi, F. E., Lutalo, T., Kagaayi, J., Sewankambo, N. K., Kigozi, G., Serwadda, D. M., Kong, X., Wawer, M. J., Wabwire-Mangen, F., & Gray, R. H. (2014). Marriage and the risk of incident HIV infection in Rakai, Uganda. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *65*(1), 91–98. https://doi.org/10.1097/QAI.0b013e3182a7f08a

Roura, M., Watson-jones, D., Kahawita, T. M., Ferguson, L., & Ross, D. A. (2013). Provider-initiated testing and counselling programmes in sub-Saharan Africa: A systematic review of their operational implementation. *AIDS*, *27*, 617–626. https://doi.org/10.1097/QAD.0b013e32835b7048

Sharma, M., Ying, R., Tarr, G., & Barnabas, R. (2015). Systematic review and meta-analysis of community and facility-based HIV testing to address linkage to care gaps in sub-Saharan Africa. *Nature*, *528*(7580), S77. https://doi.org/10.1038/nature16044

Sharma, M., Ying, R., Tarr, G., Barnabas, R., Division, I. D., & Hutchinson, F. (2016). Systematic review and meta-analysis of community and facility-based HIV testing to address linkage to care gaps in sub-Saharan Africa. *Nature*, *528*(7580), 1–26. https://doi.org/10.1038/nature16044.A

Silvestri, D. M., Modjarrad, K., Blevins, M. L., Halale, E., Vermund, S. H., & Mckinzie, J. P. (2011). A comparison of HIV detection rates using routine opt-out provider-initiated HIV testing and counseling versus a standard of care approach in a rural African setting. *Journal of Acquired Immune Deficiency Syndromes*, *56*(1), 71. https://doi.org/10.1097/01.qai.0000397355.61168.f5

Ssebunya, R. N., Wanyenze, R. K., Namale, L., Lukolyo, H., Kisitu, G. P., Nahirya-Ntege, P., & Kekitiinwa, A. (2018). Prevalence and correlates of HIV testing among adolescents 10–19 years in a post-conflict pastoralist community of Karamoja region, Uganda. *BMC Public Health*, *18*(1), 612. https://doi.org/10.1186/s12889-018-5544-0

StataCorp. (2017). *Stata statistical software: Version 15* [Computer software].[Q6] https://www.stata.com/.

Staveteig, S., Croft, T. N., Kampa, K. T., & Head, S. K. (2017). Reaching the 'first 90': Gaps in coverage of HIV testing among people living with HIV in 16 African countries.

PLOS ONE, *12*(10). Article e0186316. https://doi.org/10.1371/journal.pone.0186316

Uganda AIDS Commission. (2017). *Uganda HIV/AIDS country progress report July 2016–June 2017*. https://www.unaids.org/sites/default/files/country/documents/UGA_2018_countryreport.pdf

Uganda Bureau of Statistics (UBOS) and ICF International Inc. (2012). *Uganda demographic and health survey 2011*.

UNAIDS. (September 2016). Fact sheet 2016 | Global Statistics-2015. hhttps://www.unaids.org/sites/default/files/media_asset/20150901_FactSheet_2015_en.pdf.

Wanyenze, R. K., Nawavvu, C., Namale, A. S., Mayanja, B., Bunnell, R., Abang, B., Amanyire, G., Amanyire, N., & Kamya, M. R. (2008). Acceptability of routine HIV counselling and testing, and HIV seroprevalence in Ugandan hospitals. *Bulletin of the World Health Organization*, *86*, 302–309. https://doi.org/10.2471/BLT.07.042580

Wanyenze, R. K., Nawavvu, C., Ouma, J., Namale, A., Colebunders, R., & Kamya, M. R. (2009). Provider-initiated HIV testing for paediatric inpatients and their caretakers is feasible and acceptable. *Tropical Medicine & International Health*, https://doi.org/10.1111/j.1365-3156.2009.02417.x

Wilson, K. C., Mhangara, M., Dzangare, J., Eaton, J. W., Hallett, T. B., Mugurungi, O., & Gregson, S. (2017). Does nonlocal women's attendance at antenatal clinics distort HIV prevalence surveillance estimates in pregnant women in Zimbabwe? *AIDS*, *31*(Suppl 1), S95–S102. https://doi.org/10.1097/QAD.0000000000001337

World Health Organization. (2015). *Consolidated guidelines on HIV testing services: 5Cs: Consent, confidentiality, counselling, correct results, and connection*.

World Health Organization, & Joint United Nations Programme on HIV/AIDS. (2007). *Guidance on provider-initiated HIV testing and counselling in health facilities*. World Health Organization.

Zaba, B. W., Carpenter, L. M., Boerma, J. T., Gregson, S., Nakiyingi, J., & Urassa, M. (2000). Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa. *AIDS*, *14*(17), 2741–2750. https://doi.org/10.1097/00002030-200012010-00014

Appendices

Appendix 1. Health care classification in Uganda

(1) Health Center (HC) II is the lowest level of service delivery. It is located at parish level. It forms the first contact with the formal health sector for the community. It provides Primary Health Care (PHC) and outpatient clinical services to about 5000 people. It has no inpatient and laboratory services and is headed by a nurse.
(2) Health Center (HC) III is located in the sub-Counties, provides PHC, laboratory, clinical outpatient as well as maternity services. Its catchment population is 20,000 people. It is a referral facility for the community and HC II.
(3) Health Center (HC) IV is located at the county level or sub-district level and covers a population of 100,000 people. It provides PHC, clinical, maternity, laboratory, blood transfusion and emergency surgery services. It is a referral facility for the community, HC II and III.

Appendix 2. Method for computing confidence interval

Let X and Y be independent binomial variates based on the sample sizes m and n and parameters p_1 and p_2, respectively. Let the, We computed the prevalence ratio confidence intervals using the Katz et al. (1978) method as follows.

Letting $T = (X/m)/(Y/n)$, then the random variable $\ln(T)$ is approximately normally distributed with approximate mean $\ln(\theta)$ and estimated variance $\hat{\sigma}^2 = ((1/x) - (1/m)) + ((1/y) - (1/n))$.

The approximate two sided $1 - \alpha/2$ confidence interval for θ is given by

$$\{t * \exp(-Z_{1-\alpha/2}) * \hat{\sigma},\ \ t * \exp(Z_{1-\alpha/2}) * \hat{\sigma}\},$$

where $Z_{1-\alpha/2}$ is the $1 - \alpha/2$ percentile of the standard normal distribution, and t is the observed value of T.

Appendix 3. Prevalence of ever testing for HIV by testing venue

Characteristic	Overall N	HIV test taken in: Health facility n	Health facility Weighted percentage	Community n	Community Weighted percentage
Total	11685	8978	77.1	2707	22.9
Gender					
Male	4038	2582	64.2	1456	35.8
Female	7647	6396	84.0	1251	16.0
Age					
15–19	1498	953	65.1	545	34.1
20–29	4789	3894	81.2	895	18.8
30–39	3422	2716	79.6	706	20.5
40–49	1976	1415	71.7	561	28.3
Education level					
No Education	1216	1007	82.3	209	17.7
Primary	6406	5003	78.3	1403	21.7
Secondary+	4063	2968	73.7	1095	26.3
Marital status					
Never married	2159	1301	61.4	858	14.5
Married/Cohabiting	8169	6609	80.8	1560	19.2
Previously married	1357	1068	79.8	289	20.2
Number of sexual partners in last 12 months					
0	1957	1252	65.4	705	34.6
1	8649	6980	80.6	1669	19.4
2+	1079	746	69.9	333	30.0
Currently working					
No	2894	2219	77.5	675	22.5
Yes	8791	6759	77.0	2032	23.0
Distance to nearest HF					
<2	2441	1920	78.9	521	21.1
2–5	4823	3696	77.0	1127	23.0
5+	4009	3040	76.0	969	24.0
Don't Know	412	322	77.0	90	23.0
Area of residence					
Rural	8870	6845	77.2	2025	22.8
Urban	2815	2133	76.7	682	23.4
Region					
Central 1	1151	889	77.2	262	22.8
Central 2	1177	861	72.4	316	27.6
Kampala	1464	1068	72.2	396	27.8
East Central	1013	722	72.5	291	27.5
Mid-Eastern	901	732	81.0	169	19.0
North East	1134	875	75.4	259	24.6
West Nile	1220	896	72.8	324	27.2
Mid Northern	1394	1107	79.4	287	20.6
South Western	1024	853	83.6	171	16.4
Mid-Western	1207	975	80.5	232	19.5

Note: Coverage proportions are weighted using population survey weights.

Appendix 4. Factors associated with HIV positive among those tested in the 12 months preceding the survey

Characteristic	Overall cOR	Overall (95% CI)	Overall aOR	Overall (95% CI)	Tested in health facility aOR	Tested in health facility (95% CI)	Tested in community aOR	Tested in community (95% CI)
Gender (Reference group: Male)								
Female	1.05	(0.85, 1.29)	1.16	(0.92, 1.47)	1.00	(0.77, 1.29)*	1.76	(1.00, 3.10)*
Age (Reference group: 20–29)								
15–19	0.67	(0.45, 0.99)*	0.59	(0.40, 0.88)*	0.68	(0.45, 1.03)	0.34	(0.12, 0.96)*
30–39	1.78	(1.43, 2.22)*	1.56	(1.25, 1.94)*	1.66	(1.31, 2.10)*	1.03	(0.59, 1.80)
40–49	2.45	(1.88, 3.19)*	1.96	(1.51, 2.55)*	2.09	(1.58, 2.77)*	1.49	(0.82, 2.71)
Education level (Reference group: Primary)								
No Education	1.03	(0.76, 1.40)	0.82	(0.61, 1.11)	0.88	(0.64, 1.20)	0.60	(0.28, 1.37)
Secondary	0.59	(0.48, 0.75)*	0.66	(0.52, 0.83)*	0.61	(0.47, 0.78)*	0.93	(0.64, 1.61)
Marital status (Reference group: Married/Living together)								
Never married	0.60	(0.43, 0.85)*	0.74	(0.50, 1.11)	0.85	(0.55, 1.31)	0.46	(0.19, 1.15)
Previously married	3.26	(2.53, 4.19)*	2.53	(1.94, 3.28)*	2.76	(2.08, 3.66)*	1.64	(0.88, 3.05)

(*Continued*)

Continued.

	Overall				Tested in health facility		Tested in community	
Characteristic	cOR	(95% CI)	aOR	(95% CI)	aOR	(95% CI)	aOR	(95% CI)
Number of sexual partners in 12 months preceding survey (Reference group: 1) sig 10%								
0	1.89	(1.46, 2.46)*	1.93	(1.00, 1.75)	1.39	(1.02, 1.90)*	1.28	(0.63, 2.60)
2+	1.31	(0.94, 1.82)	1.36	(0.99, 1.88)	1.15	(0.79, 1.67)	1.98	(1.05, 3.75)**
Currently working (Reference group: Not employed)								
Employed	1.21	(0.96, 1.52)	1.01	(0.81, 1.25)	1.01	(0.79, 1.29)	0.94	(0.54, 1.65)
Distance to nearest HF (Reference group: <2 Km)								
2–5	0.72	(0.55, 0.93)*						
5+	1.07	(0.82, 1.39)						
Don't Know	0.89	(0.51, 1.56)						
Area of residence (Reference group: Urban)								
Rural	0.80	(0.63, 1.02)	0.61	(0.44, 0.85)*	0.58	(0.41, 0.82)*	0.90	(0.39, 2.06)
Region (Reference group: Central 1)								
Central 2	0.69	(0.47, 1.03)**	0.68	(0.46, 1.00)**	0.68	(0.45, 1.04)	0.66	(0.25, 1.71)
Kampala	0.66	(0.43, 1.01)**	0.54	(0.33, 0.88)*	0.51	(0.30, 0.87)*	0.75	(0.24, 2.38)
East Central	0.58	(0.35, 0.97)*	0.54	(0.34, 0.86)*	0.58	(0.36, 0.97)*	0.41	(0.15, 1.13)
Mid-Eastern	0.36	(0.20, 0.64)*	0.46	(0.27, 0.78)*	0.38	(0.21, 0.68)*	0.76	(0.30, 1.94)
North East	0.54	(0.34, 0.85)*	0.61	(0.40, 0.94)*	0.54	(0.34, 0.86)*	1.02	(0.44, 2.37)
West Nile	0.38	(0.21, 0.68)*	0.38	(0.22, 0.65)*	0.34	(0.19, 0.62)*	0.51	(0.19, 1.38)
Mid Northern	0.78	(0.51, 1.19)	0.74	(0.48, 1.13)	0.78	(0.50, 1.22)	0.36	(0.13, 1.01)
South western	0.86	(0.56, 1.32)	0.94	(0.61, 1.45)	0.87	(0.55, 1.38)	1.27	(0.49, 3.30)
Mid-western	0.72	(0.46, 1.15)	0.78	(0.49, 1.22)	0.75	(0.46, 1.22)	0.83	(0.31, 2.26)

Appendix VI: Original paper two: BMC Public Health

Ouma et al. BMC Public Health (2020) 20:379
https://doi.org/10.1186/s12889-020-8436-z

BMC Public Health

RESEARCH ARTICLE Open Access

Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda

Joseph Ouma[1*], Caroline Jeffery[2†], Joseph J. Valadez[2], Rhoda K. Wanyenze[3], Jim Todd[4†] and Jonathan Levin[1†]

Abstract

Background: National or regional population-based HIV prevalence surveys have small sample sizes at district or sub-district levels; this leads to wide confidence intervals when estimating HIV prevalence at district level for programme monitoring and decision making. Health facility programme data, collected during service delivery is widely available, but since people self-select for HIV testing, HIV prevalence estimates based on it, is subject to selection bias. We present a statistical annealing technique, Hybrid Prevalence Estimation (HPE), that combines a small population-based survey sample with a facility-based sample to generate district level HIV prevalence estimates with associated confidence intervals.

Methods: We apply the HPE methodology to combine the 2011 Uganda AIDS indicator survey with the 2011 health facility HIV testing data to obtain HIV prevalence estimates for districts in Uganda. Multilevel logistic regression was used to obtain the propensity of testing for HIV in a health facility, and the propensity to test was used to combine the population survey and health facility HIV testing data to obtain the HPEs. We assessed comparability of the HPEs and survey-based estimates using Bland Altman analysis.

Results: The estimates ranged from 0.012 to 0.178 and had narrower confidence intervals compared to survey-based estimates. The average difference between HPEs and population survey estimates was 0.00 (95% CI: − 0.04, 0.04). The HPE standard errors were 28.9% (95% CI: 23.4–34.4) reduced, compared to survey-based standard errors. Overall reduction in HPE standard errors compared survey-based standard errors ranged from 5.4 to 95%.

Conclusions: Facility data can be combined with population survey data to obtain more accurate HIV prevalence estimates for geographical areas with small population survey sample sizes. We recommend use of the methodology by district level managers to obtain more accurate HIV prevalence estimates to guide decision making without incurring additional data collection costs.

Keywords: Combining, Bias, Population survey, Health Information System, Hybrid Prevalence Estimate, District Health Information System

* Correspondence: oumajosephd@gmail.com
†Caroline Jeffery, Jim Todd and Jonathan Levin contributed equally to this work.
[1]Division of Epidemiology and Biostatistics, School of Public Health, University of Witwatersrand, Johannesburg, South Africa
Full list of author information is available at the end of the article

Background

Accurate data are needed for monitoring health programmes and interventions and for appropriate allocation of resources. In most of sub-Saharan Africa (SSA), where the HIV/AIDS epidemic is generalized, national population surveys, such as AIDS Indicator Surveys (AIS), are preferred to provide reliable health indicator estimates for programme monitoring [1]. The surveys are however designed to provide estimates at national and regional levels, but small sample sizes at district or sub-district levels lead to less reliable indicator estimates, that have wider confidence intervals [1–4].

Health Information Systems such as the District Health Information System (DHIS2) provide another source of information that can be used for monitoring the HIV/AIDS epidemic. This data is collected more regularly, available at more decentralized levels, e.g. districts and costs less to collect. WHO, UNAIDS and other development partners recommend use of routine facility data in addition to other data sources to monitor programme performance, assess intervention coverage and measure levels of disease in a population [5]. Use of routine health facility data informed the adjustments in HIV prevalence estimates in many countries in Eastern and Southern Africa [6]. Several other studies highlight the utility of data from routine service delivery in informing service delivery decisions [7, 8]. Routine service delivery data, however, are collected only on individuals who attend/access health facilities and thus provide potentially biased estimates of population indicators.

In addition, development partners and ministries of health in middle and low-income countries have invested in electronic health information systems including the DHIS2, to improve the quality and timeliness of data from the systems. In Uganda, Ministry of Health (MoH) with support from development partners conduct quarterly reviews to validate data reported into DHIS2 [9]. With this investment, there is a need to find ways to utilize this source of information to inform service delivery decisions. Combining routine data with a relatively small sample of respondents from population survey data has been found to produce more accurate indicator estimates [10, 11].

Statistical models in packages such as SPECTRUM or THEMBISA attempt to use both routine and population survey data to calculate HIV/AIDS indicators. Model inputs such as ANC prevalence, mortality, number of individuals on ART and recent HIV prevalence when not available, complicate their use [12]. A simpler and more robust method may be easier to use and give good results.

Larmarange and Bendaud obtained district level estimates from population survey data from 17 countries using a kernel density approach implemented in PrevR [13]. In districts with inadequate number of observations in the survey sample, estimates were obtained based on observations from neighboring districts or administrative units and were categorized as "uncertain" estimates [13]. Using a similar approached, PrevR, UNAIDS found "uncertain" estimates in up to 86% of the districts in Mozambique and in 79% of the districts in Uganda [13, 14].

In this study, we explore use of the readily available health facility service delivery data in combination with population survey data to obtain more accurate HIV prevalence estimates at district level for monitoring interventions and disease impact in the general population. We implement the Hybrid Prevalence Estimation (HPE) methodology to obtain HIV prevalence estimate and 95% confidence interval for districts in Uganda. The estimation process accounts for sample size limitations associated with population survey data at district level and self-selection bias associated with health facility testers, a limitation that many researchers have not been able to address adequately [2–4, 15–18].

Methods

Data sources

We analyzed data from the 2011 Uganda AIDS Indicator Survey (UAIS) and health facility HIV testing data from the national DHIS2 collected during 2011. UAIS data was downloaded from the Measure DHS website www.measuredhs.com after obtaining consent from ICF/Macro international, while health facility testing data was extracted from the DHIS2 hosted at MoH after obtaining written permission from MoH. Ethical clearance to conduct this study was obtained from the University of Witwatersrand Human Research Ethics Committee (HREC) and Uganda National Council for Science and Technology (UNCST).

Uganda AIDS Indicator survey

The UAIS is a nationally representative, population-based, HIV serological survey, designed to provide HIV prevalence estimates at national and regional levels [19]. The survey used a two-stage cluster sampling design. For the 2011 survey, Uganda was divided into 10 geographical regions each consisting of 8–15 neighboring districts. Clusters were randomly selected from each region with probability proportional to number of households in the cluster. The estimated number of households per cluster were projections from the 2002 National Population and Housing Census (NPHC) [20]. Clusters were enumeration areas from the 2002 NPHC. Sample sizes were allocated equally across the 10 geographical regions. A systematic sample of 25 households were then selected from each cluster using the 2002 NPHC sampling frame. All adults present in the selected household and who consented to participate in the survey were interviewed [19]. More details about the survey are available from www.measuredhs.com.

For this study, a total of 19,475 individuals (8532 men and 10,943 women) aged 15–49 years and tested for HIV during the survey were considered. Variables included in the analysis were (i) at cluster level: area of residence (urban/rural) and region of the country and at (ii) individual level: respondents' gender, marital status, education level attained, number of sexual partners including husband/wife in the 12 months preceding the survey, employment status and distance to nearest health facility.

A multilevel logistic regression model was fitted to the UAIS data to obtain the respondents' probability of testing for HIV in a health facility. The model was fitted using a total of 470 clusters. The average number of observations per cluster were 45(min = 11 and max = 64). Unequal sample selection probabilities were accounted for by incorporating scaled sampling weights. Carle's methodology was applied to adjust/scale the sampling weights [21]. The models were fitted using maximum likelihood method in Stata statistical software, release 15 [22].

Survey respondents were considered to have tested for HIV at a health facility if they reported that they tested for HIV in health facility and received their test results in the 12 months preceding the survey. Pregnant or breastfeeding women who tested for HIV during antenatal care attendance and individuals who tested at an HIV care centre such as The AIDS Support Organization (TASO) and AIDS Information Centre (AIC) were included in the analysis. Health facilities included facilities owned and managed by government (public) and private organizations that reported HIV testing data to the national DHIS2.

Health facility data

Health facility HIV testing data comprised of data reported to the national DHIS2. The system was developed to provide accurate, timely and quality routine data for monitoring and planning for the health sector in Uganda [9, 23]. Training and technical support from development partners and MoH has led to improvement in the quality and reliability of data in the system [9]. Aggregated HIV testing data is reported by health facilities to the DHIS2 on a monthly basis. The data includes HIV testing at all inpatient and outpatient departments in health facilities. For 2011 reporting period, data was disaggregated by age (i.e. 0–14, 15–49 and 50+ years) and gender (male and female). For this study, we considered males and females aged 15–49 years.

Indicators considered for this analysis were: number of individuals who were tested and received their HIV test results (A) and number of individuals who tested HIV positive (B). For ANC data, we considered number of pregnant women counseled, tested and received their HIV test results (C) at first antenatal visit and the number who tested HIV positive (D). HIV counseling and testing algorithm in Uganda recommends HIV testing for any individual whose most recent negative HIV test result was conducted more than 3 months prior to the current visit to the health facility [24]. Some individuals may test multiple times within a year but may not disclose to health workers resulting in double counting, a key limitation for this study. Furthermore, some pregnant women may test for HIV before seeking antenatal care and test again during antenatal attendance leading to double counting in the data reported to the national DHIS2.

Variables based on DHIS2 data were defined as follows;

- Total number of individuals tested for HIV = $(A + C)$
- Number HIV positive = $(B + D)$
- Total number of males tested for HIV = *males in A*
- Number of males tested HIV positive = males in *B*
- Total number of females tested for HIV = (*females in A*) + *C*
- Number of female tested HIV positive = (*females in B*) + *D*

Addressing possible bias in health facility data

Routine facility data collected as part of service delivery consists of individuals who self-select, limiting its' use for general population health indicator monitoring. To obtain general population indicator estimates, some researchers have used census projections as denominators, however this approach often results in coverage estimates that are greater than 100% [25]. Population surveys are preferred to obtain health indicator denominators since their design takes into account population distribution in the country [25–28]. The UAIS comprise two subpopulations, namely individuals who tested for HIV in a health facility in the 12 months preceding the survey (the facility testers) and those who did not test for HIV in a health facility (the non-facility testers) for the same period. We assume that the UAIS estimates of HIV prevalence for those who tested for HIV in a health facility and for those who did not test for HIV in a health facility are accurate at regional levels, since estimates of domain proportions from a multistage survey are unbiased. We apply this assumption to adjust the denominators of the DHIS2 data so that at the regional level, DHIS2 HIV prevalence estimates are similar to UAIS prevalence estimates. The adjustment process was carried out as follows:

1. We obtained the HIVs prevalence $\hat{k}_f$ among health facility testers in each region in the UAIS data.
2. We adjusted denominators in the DHIS2 data for each region using $n^r_{ajdusted} = \frac{n_{pos}}{\hat{k}_f}$, where n_{pos} is the observed number of individuals who tested HIV positive in each region in the DHIS2 data.

3. Calculated an adjustment factor (δ_f) for each region, using $\delta_f = \frac{n^r_{ajdusted}}{n_r}$, where nr is the observed number of individuals who tested for HIV in each region from the DHIS2 data.
4. We applied the adjustment factor (δ_f), to obtain $n^d_{ajdusted}$, the adjusted number of individuals who tested for HIV in a health facility at district level using, $n^d_{ajdusted} = \delta_f * n_d$, where n_d is the observed number of individuals tested for HIV at district level.
5. HIV prevalence (P_f) based on DHIS2 adjusted data in the district was then obtained as a ratio of n_{pos}, the total observed positives and $n_{adjusted}$ the adjusted number of individuals who tested for HIV in the district, i.e. $P_f = \frac{n_{pos}}{n^d_{ajdusted}}$

Hybrid prevalence estimation methodology

We consider n individuals in the UAIS to include n_c individuals who tested for HIV at a health facility during the 12 months preceding the survey and know their test result and $n_{\underline{c}}$ individuals who did not test for HIV at a health facility and therefore do not know their HIV status. i.e. $n = n_c + n_{\underline{c}}$. Using health facility prevalence computed in step 5 above, we computed district HIV prevalence as a weighted average of prevalence from DHIS2 data, P_f and prevalence among individuals who did not test for HIV in a health facility, $\hat{P}_{\underline{s}}$ estimated from the UAIS data.

$$\hat{P} = \hat{\pi}_c P_f + (1-\hat{\pi}_c)\hat{P}_{\underline{s}} \quad (1)$$

where;

$\hat{P}$ – HPE/combined estimate, $\hat{\pi}_c$ – the estimated probability of testing for HIV in a health facility, P_f– Adjusted HIV prevalence for individuals tested at a health facility and $\hat{P}_{\underline{s}}$ – HIV prevalence for individuals tested during the survey and had not tested for HIV in a health facility in the 12 months preceding the survey. We estimated $\hat{\pi}_c$ from UAIS data using multilevel logistic regression adjusting for both individual and cluster level factors. Applying this model, we account for clustering at cluster level [25]. Although the probability of testing for HIV in a health facility was obtained at individual level, we used average district level probability of testing to combine the estimates. Since average probability of HIV testing is obtained from a survey sample containing both facility and non-facility testers, we estimate the variance and standard errors (SE) for the HPE respectively as follows;

$$Var(\hat{P}) = \frac{1}{n}\left\{\hat{P}_{\underline{s}}(1-\hat{P}_{\underline{s}})(1-\hat{\pi}_c) + (1-\hat{\pi}_c)\left(P_f - \hat{P}_{\underline{s}}\right)^2\right\}$$
$$\text{and } SE = \sqrt{var(\hat{P})} \quad (2)$$

We assess accuracy of the HPEs compared to survey-based prevalence estimates by computing the percentage change in standard errors. We further assessed agreement of the estimates obtained using the HPE methodology with those from population survey method (Direct population survey estimate) using a Bland Altman analysis [26, 27].

All analysis was carried out in Stata statistical analysis software, Release 15 [22] and R version 3.5.0 [28].

Results

Of the 19,475 individuals, 6729 (34.6%) tested for HIV in a health facility in the 12 months preceding the survey. HIV prevalence among those who tested in a health facility was 0.084 compared to 0.068 among those who did not test in a health facility (Table 1).

From health facility data, national (unadjusted) HIV prevalence was 0.058 (Male: 0.057 Female: 0.059). A total of 4,758,991 (female: 73.7%) individuals were tested for HIV in a health facility. (Table 1). DHIS2 HIV positivity by gender is presented Additional file 1: Appendix 1.

Weighting/annealing factor

Overall (national) average propensity to test in a health facility was 0.35 (male: 0.27, female:0.41). It ranged from 0.001 to 0.95 (Fig. 1). Mid Northern region had the highest average propensity to test for HIV in health facility, 0.44 (male: 0.40 and female: 0.49) while Mid-Eastern region had the lowest, 0.25 (male: 0.16, female: 0.32) (Fig. 1).

Hybrid prevalence estimates

HIV prevalence was highest in Central 1 region (0.11) and lowest in Mid-Western and Mid-Eastern regions (0.04 in each region). District level HPEs ranged from 0.01 to 0.18. Average HIV prevalence by region were; Central 1: 0.11 (min; 0.06, max; 0.18), Central 2: 0.10 (0.08, 0.17), East Central: 0.05 (0.02, 0.09), Mid-Eastern: 0.04 (0.01, 0.09), Mid Northern: 0.09 (0.05, 0.14), Mid-Western: 0.08 (0.03, 0.16), North East:0.04 (0.04, 0.10), South West: 0.08 (0.04, 0.13) and West Nile: 0.04 (0.03, 0.07) (Table 2). Table 2 also presents HPE, survey and DHIS2 based district HIV prevalence estimates by district.

Figure 2 presents HIV prevalence maps in; both sexes (map a); in males (map b); and in females (map c). HPEs had similar patterns for both sexes, males and females consistent with the regional level prevalence estimates from population survey in Table 1. Districts in Central 1 region, Mid northern region, Island, and those along lake shores had higher overall, male and female HIV prevalence estimates (Fig. 2, and Additional file 1: Appendix 2) while districts in mid-eastern and West Nile region had lower HIV prevalence estimates. HPEs were not calculated for two districts (Bukwo in mid-eastern

Table 1 Regional level HIV prevalence estimates

Region	Population Survey prevalence (in proportion) (Number HIV+)			Health Facility Prevalence (Number HIV+)[a]
	Overall	Tested in Health Facility	Not tested in health-Facility	
Central 1	0.106 (185)	0.123 (74)	0.096 (111)	0.094 (40,880)
Central 2	0.090 (166)	0.079 (57)	0.096 (109)	0.070 (36,125)
East Central	0.057 (108)	0.081 (39)	0.048 (69)	0.037 (17,207)
Kampala	0.071 (156)	0.080 (62)	0.066 (94)	0.098 (25,447)
Mid-Eastern	0.041 (88)	0.042 (24)	0.041 (64)	0.040 (13,926)
North East	0.053 (98)	0.053 (45)	0.052 (53)	0.026 (15,106)
Mid Northern	0.083 (159)	0.099 (81)	0.071 (78)	0.076 (39,559)
Mid-Western	0.083 (170)	0.089 (63)	0.080 (107)	0.062 (52,102)
West Nile	0.048 (88)	0.055 (33)	0.045 (55)	0.031 (7758)
South Western	0.080 (149)	0.103 (64)	0.069 (85)	0.053 (29,338)
National	0.073 (1367)	0.084 (542)	0.068 (825)	0.058 (277,448)

NOTE: Regional HIV Prevalence from population survey and health facility testing data
[a]HIV prevalence from the unadjusted health facility data

region and Ntoroko in mid-western region) because UAIS data points was not available for those districts.

Figure 3 compares district HIV prevalence estimates from population survey and HPE while in Fig. 4, we compare HPE and the adjusted DHIS2 data for selected districts. Prevalence comparison between HPE and survey for all districts is presented in Additional file 1: Appendix 3. The figures show that HPEs had narrower confidence intervals compared to direct survey estimates indicating an improvement in the precision of the estimates.

Of the 110 districts, 51 (46.4%) had lower HPEs (point estimates) and 59 (53.6%) had higher HPE compared to the survey-based district prevalence estimates (Fig. 3, Additional file 1: Appendix 3).

HPEs were however lower than the DHIS2 prevalence estimates in 74 (67.3%) and higher in 36 (32.7%) of the districts in Uganda (Fig. 4, Additional file 1: Appendix 4).

A joint comparison of the HP estimates with both survey-based and health facility-based prevalence estimates show that 33 (30.0%) of the districts had lower HPEs while 18 (16.4%) had higher estimates compared

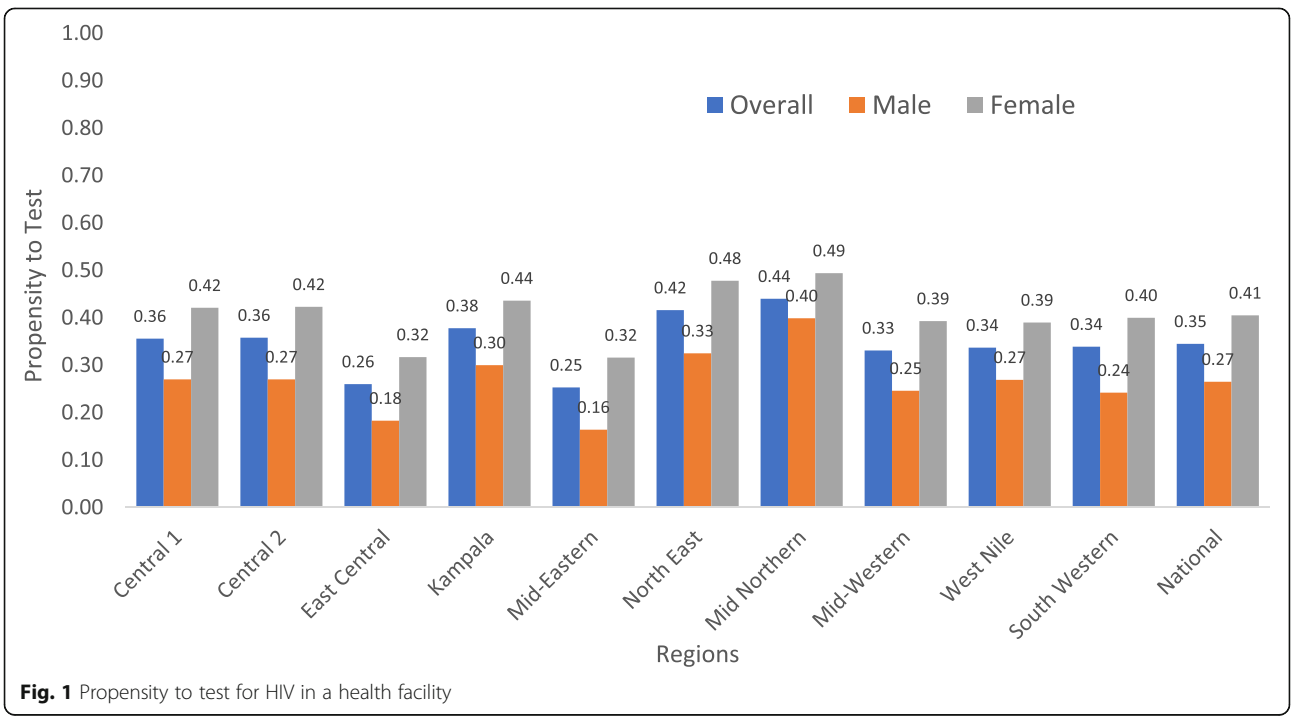


Fig. 1 Propensity to test for HIV in a health facility

Table 2 HPE HIV prevalence estimates, (HPE, Survey and unadjusted DHIS2)

Region and District	HPE		Population Survey		Facility Data (unadjusted)
	Estimate (95% CI)		Estimate (95% CI)		
Central 1					
Bukomansimbi	0.064	(0.027, 0.101)	0.074	(0.021, 0.127)	0.083
Butambala	0.103	(0.017, 0.190)	0.124	(0.000, 1.000)	0.063
Gomba	0.090	(0.036, 0.144)	0.158	(0.060, 0.257)	0.063
Kalangala	0.138	(0.064, 0.212)	0.195	(0.000, 1.000)	0.135
Kalungu	0.142	(0.065, 0.219)	0.102	(0.025, 0.179)	0.086
Lwengo	0.143	(0.090, 0.195)	0.157	(0.090, 0.223)	0.105
Lyantonde	0.121	(0.012, 0.229)	0.109	(0.000, 1.000)	0.085
Masaka	0.178	(0.129, 0.227)	0.156	(0.096, 0.216)	0.131
Mpigi	0.097	(0.065, 0.129)	0.108	(0.059, 0.158)	0.101
Rakai	0.081	(0.053, 0.109)	0.076	(0.042, 0.109)	0.067
Sembabule	0.072	(0.037, 0.106)	0.067	(0.015, 0.118)	0.078
Wakiso	0.107	(0.090, 0.124)	0.096	(0.070, 0.123)	0.094
Central 2					
Buikwe	0.083	(0.060, 0.106)	0.079	(0.042, 0.116)	0.086
Buvuma	0.170	(0.102, 0.238)	0.185	(0.096, 0.273)	0.085
Kayunga	0.069	(0.041, 0.096)	0.065	(0.030, 0.100)	0.049
Kiboga	0.090	(0.035, 0.145)	0.063	(0.008, 0.117)	0.069
Kyankwanzi	0.111	(0.049, 0.173)	0.128	(0.054, 0.203)	0.045
Luwero	0.094	(0.066, 0.123)	0.080	(0.044, 0.117)	0.075
Mityana	0.136	(0.097, 0.175)	0.104	(0.058, 0.151)	0.125
Mubende	0.077	(0.053, 0.101)	0.094	(0.056, 0.132)	0.057
Mukono	0.090	(0.063, 0.116)	0.092	(0.054, 0.129)	0.069
Nakaseke	0.073	(0.038, 0.109)	0.083	(0.025, 0.141)	0.083
Nakasongola	0.083	(0.041, 0.125)	0.076	(0.018, 0.134)	0.088
East Central					
Bugiri	0.037	(0.019, 0.055)	0.027	(0.008, 0.045)	0.029
Buyende	0.055	(0.020, 0.091)	0.044	(0.006, 0.081)	0.033
Iganga	0.062	(0.037, 0.087)	0.054	(0.025, 0.082)	0.035
Jinja	0.088	(0.064, 0.112)	0.090	(0.056, 0.124)	0.057
Kaliro	0.036	(0.001, 0.071)	0.046	(0.000, 0.098)	0.023
Kamuli	0.051	(0.034, 0.068)	0.058	(0.033, 0.082)	0.036
Luuka	0.023	(0.005, 0.042)	0.018	(0.000, 0.037)	0.018
Mayuge	0.042	(0.023, 0.062)	0.092	(0.043, 0.140)	0.036
Namayingo	0.089	(0.043, 0.135)	0.093	(0.035, 0.151)	0.079
Namutumba	0.037	(0.011, 0.062)	0.038	(0.000, 0.076)	0.030
Kampala	0.076	(0.067, 0.085)	0.071	(0.058, 0.084)	0.098
Mid Eastern					
Budaka	0.065	(0.018, 0.111)	0.050	(0.007, 0.093)	0.029
Bududa	0.025	(0.000, 0.055)	0.046	(0.006, 0.087)	0.014
Bulambuli	0.026	(0.000, 0.055)	0.023	(0.000, 0.053)	0.040
Busia	0.068	(0.038, 0.097)	0.074	(0.037, 0.112)	0.063
Butaleja	0.026	(0.003, 0.049)	0.026	(0.001, 0.052)	0.018

Table 2 HPE HIV prevalence estimates, (HPE, Survey and unadjusted DHIS2) *(Continued)*

Region and District	HPE		Population Survey		Facility Data (unadjusted)
	Estimate (95% CI)		Estimate (95% CI)		
Kapchorwa	0.092	(0.000, 0.186)	0.097	(0.000, 1.000)	0.021
Kibuku	0.036	(0.000, 0.074)	0.014	(0.000, 0.040)	0.027
Kween	0.013	(0.000, 0.032)	0.007	(0.000, 0.022)	0.022
Manafwa	0.034	(0.014, 0.055)	0.034	(0.006, 0.062)	0.035
Mbale	0.037	(0.022, 0.052)	0.035	(0.015, 0.055)	0.070
Pallisa	0.012	(0.002, 0.022)	0.007	(0.000, 0.019)	0.021
Sironko	0.051	(0.022, 0.079)	0.050	(0.020, 0.080)	0.030
Tororo	0.071	(0.044, 0.097)	0.069	(0.041, 0.098)	0.044
Mid Northern					
Agago	0.080	(0.047, 0.113)	0.046	(0.011, 0.081)	0.071
Alebtong	0.066	(0.035, 0.097)	0.058	(0.016, 0.100)	0.068
Amolatar	0.108	(0.058, 0.159)	0.158	(0.079, 0.237)	0.078
Amuru	0.059	(0.027, 0.091)	0.028	(0.000, 0.061)	0.043
Apac	0.062	(0.043, 0.080)	0.052	(0.022, 0.082)	0.081
Dokolo	0.070	(0.041, 0.099)	0.055	(0.015, 0.094)	0.061
Gulu	0.090	(0.066, 0.113)	0.100	(0.049, 0.151)	0.086
Kitgum	0.081	(0.042, 0.121)	0.083	(0.024, 0.141)	0.072
Kole	0.047	(0.023, 0.070)	0.029	(0.001, 0.057)	0.051
Lamwo	0.088	(0.041, 0.134)	0.121	(0.043, 0.198)	0.071
Lira	0.108	(0.079, 0.138)	0.111	(0.066, 0.157)	0.098
Nwoya	0.123	(0.033, 0.212)	0.163	(0.000, 1.000)	0.052
Otuke	0.136	(0.048, 0.224)	0.128	(0.000, 1.000)	0.075
Oyam	0.066	(0.043, 0.089)	0.068	(0.035, 0.100)	0.058
Pader	0.140	(0.088, 0.192)	0.165	(0.090, 0.239)	0.087
Mid Western					
Buliisa	0.065	(0.013, 0.117)	0.038	(0.000, 0.785)	0.065
Bundibugyo	0.032	(0.010, 0.054)	0.032	(0.000, 0.071)	0.028
Hoima	0.075	(0.050, 0.100)	0.086	(0.051, 0.121)	0.053
Kabarole	0.159	(0.124, 0.194)	0.137	(0.094, 0.180)	0.099
Kamwenge	0.065	(0.037, 0.093)	0.054	(0.023, 0.086)	0.050
Kasese	0.066	(0.046, 0.085)	0.050	(0.027, 0.073)	0.050
Kibaale	0.068	(0.045, 0.091)	0.079	(0.046, 0.113)	0.049
Kiryandongo	0.059	(0.019, 0.098)	0.071	(0.016, 0.126)	0.048
Kyegegwa	0.118	(0.058, 0.177)	0.127	(0.056, 0.198)	0.046
Kyenjojo	0.077	(0.047, 0.106)	0.121	(0.071, 0.172)	0.070
Masindi	0.063	(0.035, 0.091)	0.060	(0.023, 0.097)	0.062
North East					
Abim	0.066	(0.004, 0.128)	0.098	(0.000, 1.000)	0.031
Amudat	0.090	(0.000, 0.204)	0.049	(0.000, 0.899)	0.022
Amuria	0.102	(0.067, 0.138)	0.102	(0.056, 0.148)	0.039
Bukedea	0.044	(0.014, 0.074)	0.045	(0.002, 0.088)	0.022
Kaabong	0.037	(0.016, 0.057)	0.021	(0.000, 0.044)	0.022
Kaberamaido	0.081	(0.045, 0.117)	0.061	(0.020, 0.102)	0.034

Table 2 HPE HIV prevalence estimates, (HPE, Survey and unadjusted DHIS2) *(Continued)*

Region and District	HPE		Population Survey		Facility Data (unadjusted)
	Estimate (95% CI)		Estimate (95% CI)		
Katakwi	0.074	(0.041, 0.107)	0.080	(0.034, 0.126)	0.035
Kotido	0.052	(0.009, 0.095)	0.046	(0.000, 0.096)	0.017
Kumi	0.040	(0.023, 0.057)	0.025	(0.000, 0.051)	0.024
Moroto	0.041	(0.014, 0.069)	0.065	(0.000, 0.132)	0.023
Nakapiripirit	0.049	(0.002, 0.097)	0.038	(0.000, 0.792)	0.018
Napak	0.075	(0.018, 0.132)	0.073	(0.003, 0.143)	0.023
Ngora	0.046	(0.015, 0.078)	0.061	(0.014, 0.108)	0.020
Serere	0.049	(0.029, 0.069)	0.046	(0.014, 0.077)	0.028
Soroti	0.089	(0.060, 0.118)	0.058	(0.019, 0.097)	0.034
South Western					
Buhweju	0.035	(0.000, 0.077)	0.022	(0.000, 0.592)	0.025
Bushenyi	0.096	(0.060, 0.132)	0.079	(0.036, 0.123)	0.068
Ibanda	0.065	(0.027, 0.102)	0.053	(0.002, 0.104)	0.059
Isingiro	0.102	(0.061, 0.143)	0.112	(0.064, 0.160)	0.045
Kabale	0.043	(0.024, 0.063)	0.037	(0.010, 0.064)	0.035
Kanungu	0.077	(0.042, 0.112)	0.084	(0.036, 0.131)	0.047
Kiruhura	0.095	(0.046, 0.144)	0.086	(0.025, 0.148)	0.072
Kisoro	0.053	(0.015, 0.091)	0.059	(0.012, 0.105)	0.026
Mbarara	0.101	(0.068, 0.134)	0.124	(0.072, 0.177)	0.059
Mitooma	0.097	(0.042, 0.151)	0.076	(0.012, 0.140)	0.069
Ntungamo	0.063	(0.040, 0.085)	0.062	(0.030, 0.095)	0.050
Rubirizi	0.133	(0.067, 0.199)	0.140	(0.044, 0.236)	0.057
Rukungiri	0.082	(0.053, 0.111)	0.076	(0.038, 0.115)	0.064
Sheema	0.095	(0.056, 0.134)	0.124	(0.063, 0.185)	0.069
West Nile					
Adjumani	0.044	(0.020, 0.067)	0.039	(0.012, 0.065)	0.023
Arua	0.043	(0.031, 0.056)	0.049	(0.030, 0.069)	0.035
Koboko	0.048	(0.019, 0.077)	0.063	(0.025, 0.102)	0.025
Maracha	0.029	(0.001, 0.058)	0.007	(0.000, 0.019)	0.014
Moyo	0.033	(0.011, 0.055)	0.037	(0.005, 0.069)	0.019
Nebbi	0.068	(0.045, 0.091)	0.070	(0.042, 0.098)	0.036
Yumbe	0.026	(0.010, 0.042)	0.027	(0.006, 0.048)	0.014
Zombo	0.056	(0.031, 0.082)	0.040	(0.011, 0.069)	0.050

to both the survey and health facility-based prevalence estimates.

Precision of HPE and population survey estimates

Standard errors of the HPEs were generally lower compared to SEs from survey-based estimates (Fig. 5). Of the districts, 105 (95.5%) had lower HPE SEs compared to SEs from survey-based estimates. Overall, the HPE standard errors were decreased by 28.9% (95% CI: 23.4–34.4) compared to survey-based standard errors.

Similarity of HPE and survey-based estimates

On average, there is no difference between survey and HPE estimates, 0.0 (95% CI: – 0.04,0.04) (Fig. 6a). Average difference for males was – 0.01 (95% CI: – 0.05,0.03) while for females was 0.00 (95% CI: – 0.06,0.06). Although there seems to be a bias (0.01) when assessing the agreement between HP and survey-based estimates for males (Fig. 6b), the 95% confidence interval of the difference between the estimates are narrow. Additionally, there was no systematic pattern of the points as the average of the estimates increases.

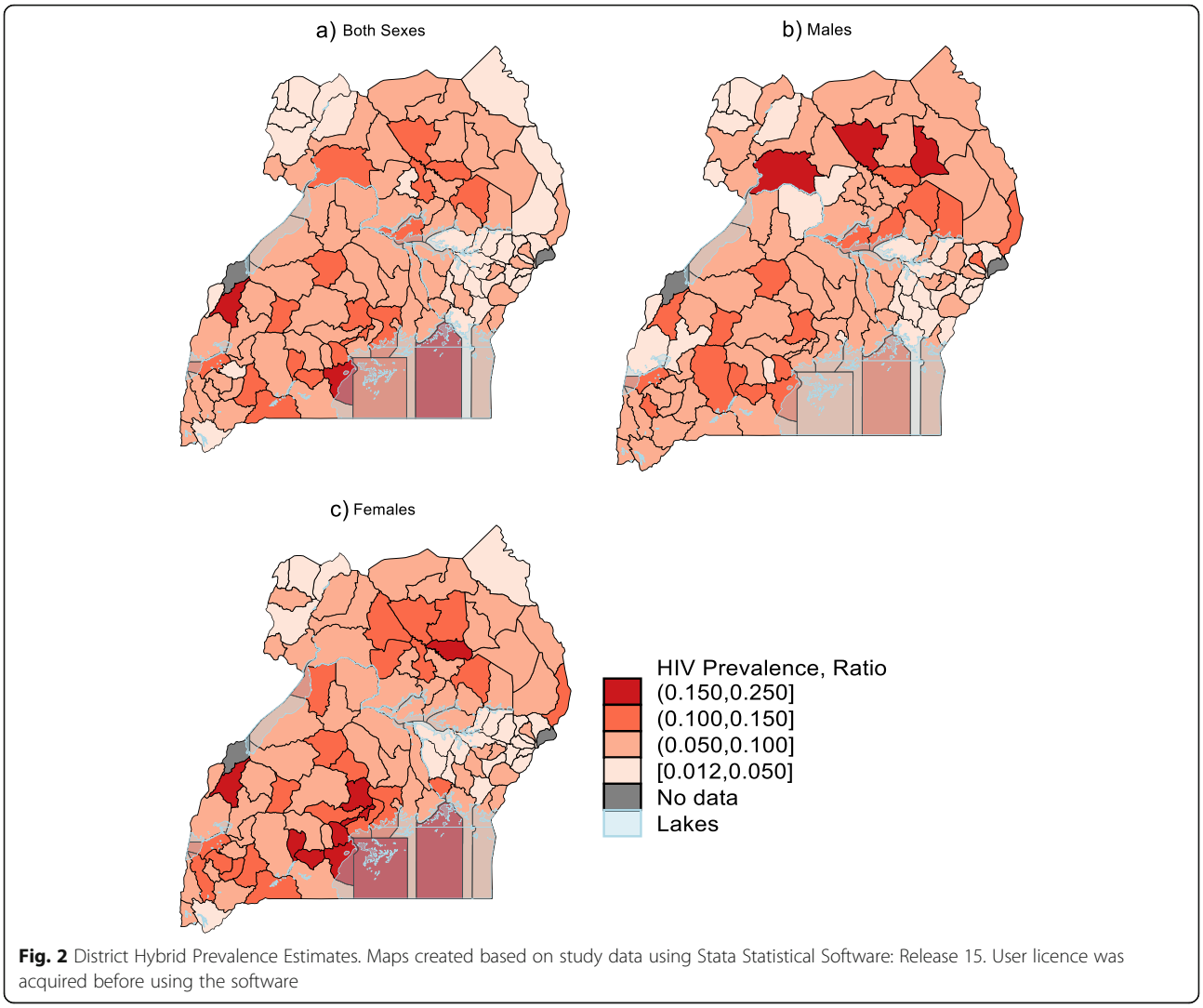


Fig. 2 District Hybrid Prevalence Estimates. Maps created based on study data using Stata Statistical Software: Release 15. User licence was acquired before using the software

The mean difference between the HPE and DHIS2 estimates was 0.01 (95% CI: − 0.05,0.06) (Fig. 7a). Average difference for males was − 0.01 (95% CI: − 0.07, 0.06) while for females was 0.02 (95% CI: − 0.05, 0.09), (Fig. 7b and c respectively). The size of the difference increased with increase in the mean of the estimates. This is seen from the wider variability of the points about the no-difference (zero) line as the values of the average of the estimates increase (Fig. 7a-c). The average difference was 0.02 and confidence intervals were wider when comparing HPEs and survey-based estimates for females (Fig. 7c).

Discussion

In this study, we implement a novel approach, the Hybrid Prevalence Estimation methodology to obtain HIV prevalence estimates for districts in Uganda. We combined DHIS2 HIV testing data with information of non-facility testers from the 2011 UAIS data to obtain district level HIV prevalence estimates.

Although national population surveys are the gold standard for calculating population level health indicators, district level estimates from these surveys are less accurate due to the reduced sample sizes at district or lower administrative levels. The demand for accurate indicator estimates at district or lower administrative levels for programme monitoring motivates use of innovative approaches to provide the estimates. We obtained district level HIV prevalence estimates by combining population survey information with DHIS2 data using a Hybrid Prevalence Estimation methodology. Our estimates had narrower confidence intervals compared to estimates from the population survey at the district level, consistent with findings elsewhere [10, 11]. The HPE was calculated from three parameters; 1) Prevalence in the health facility sample, 2) prevalence among non-facility testers from the population survey sample and 3) the propensity to test for HIV in a health facility from the population survey sample. We also observed that HIV prevalence estimate

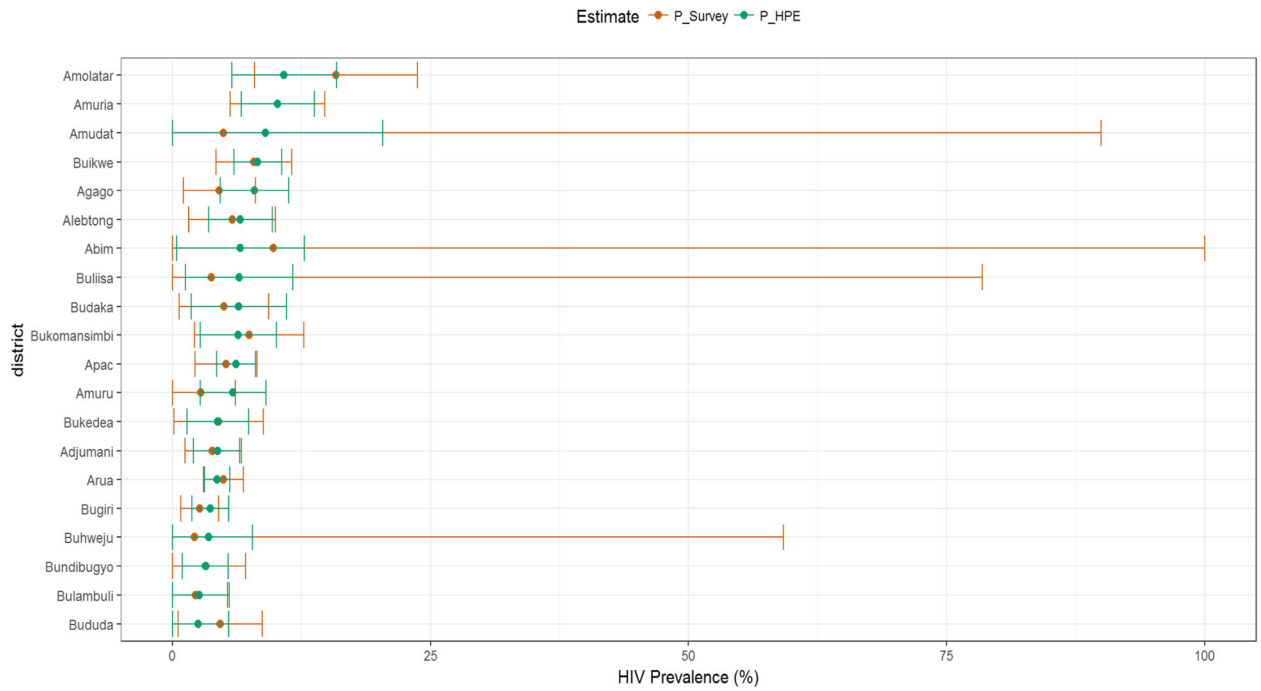


Fig. 3 District prevalence estimates from combined and population survey data. P_survey- survey based prevalence estimate while P_HPE is HIV prevalence based on the HPE methodology

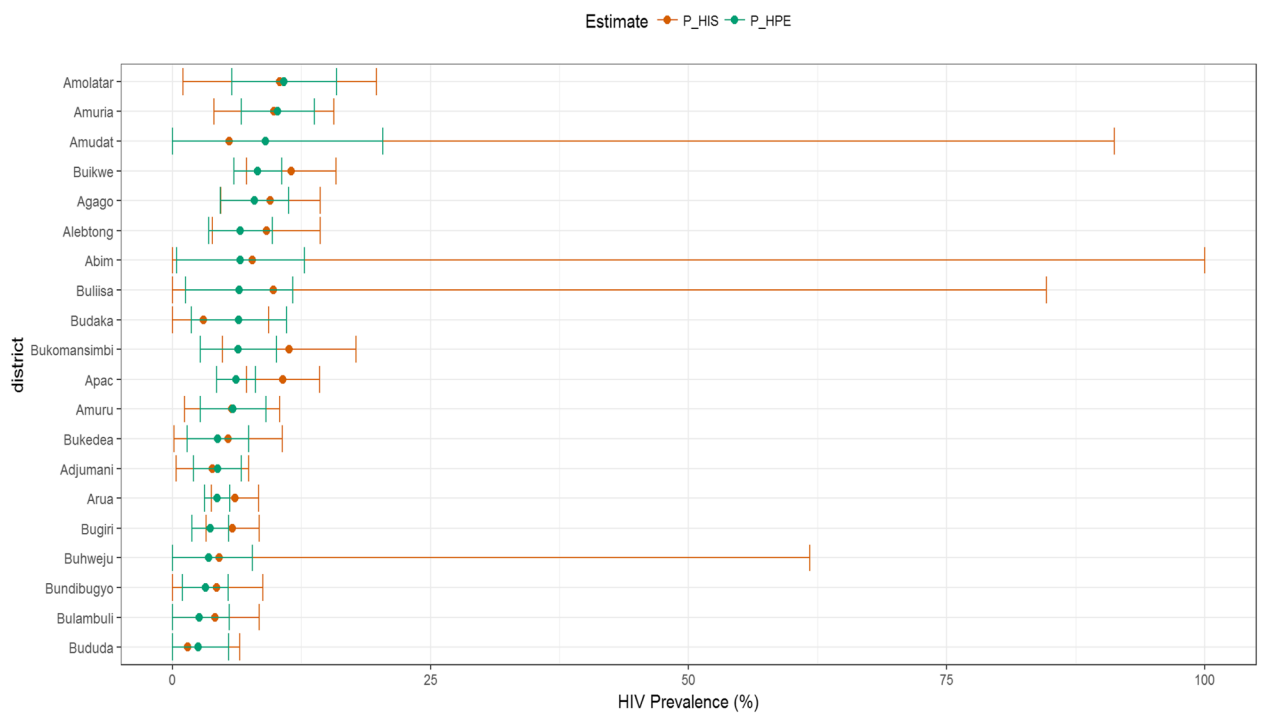


Fig. 4 District prevalence estimates from combined and DHIS2 data. P_HIS- Health facility-based prevalence estimate while P_HPE is HIV prevalence based on the HPE methodology

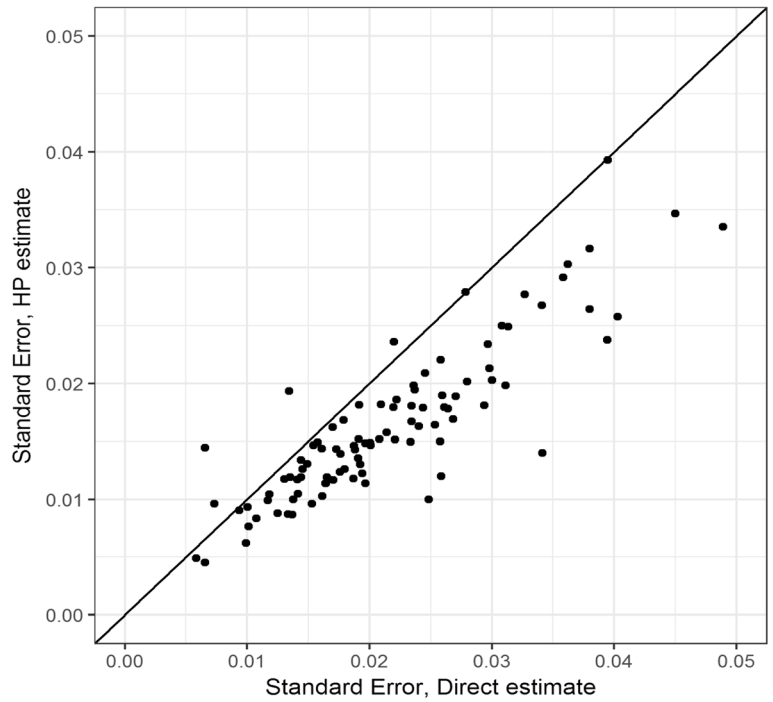


Fig. 5 Standard errors of estimates from survey and the HPE

obtained using the HPE methodology was similar to the population survey HIV prevalence estimates for male and females combined, and for males only while it was lower for females. Additionally, UAIS based prevalence estimates were generally higher while DHIS2 prevalence estimates were lower than the consistent with findings elsewhere [29].

In the UAIS, the population can be divided into two domains: 1) those that have access to health facilities, get tested for HIV, and are linked to appropriate care if found HIV positive, and 2) those that do not access health facilities and may remain unknown in the health care system. Barriers to health care access for the latter subpopulation may include factors such as low/no education and not being in stable sexual relationship that also increase the risk of HIV transmission [30, 31]. Combining survey with DHIS2 data therefore generates more precise indicator estimates that can be used to improve planning and service delivery for the general population at district levels where service delivery decision are implemented.

Facility level data has known limitations including selection bias, as it is not a random sample from the population for measuring general population level prevalence [15, 16, 18, 32]. Studies in Uganda [33, 34], Tanzania [35] and Zambia [36, 37], have also found facility-based antenatal HIV testing data has biased estimates of HIV prevalence' and therefore not appropriate for calculating HIV/AIDS indicators including HIV prevalence in the general

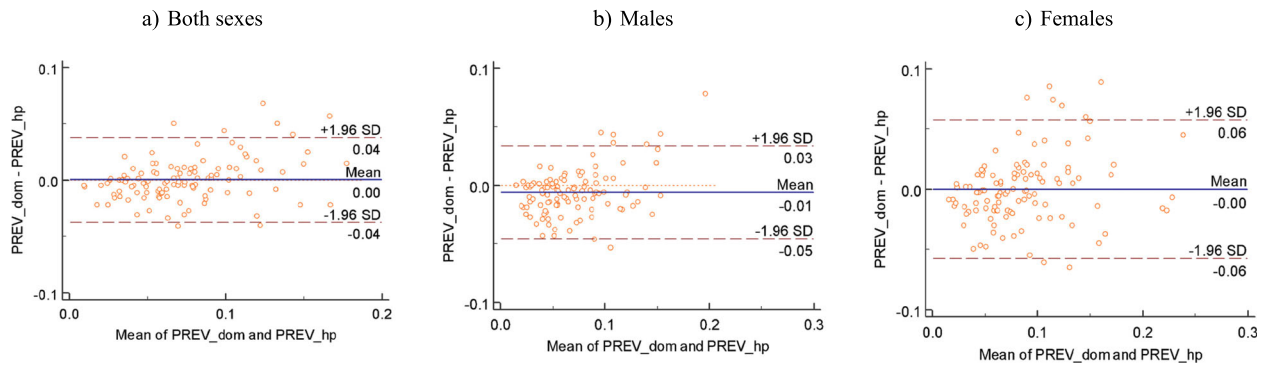


Fig. 6 Difference plot comparing HPE and Direct survey estimates. PREV_HIS- HIV Prevalence based on health facility data, PREV_hp-HIV prevalence based on the HPE methodology while PREV_dom- HIV prevalence based on survey data only

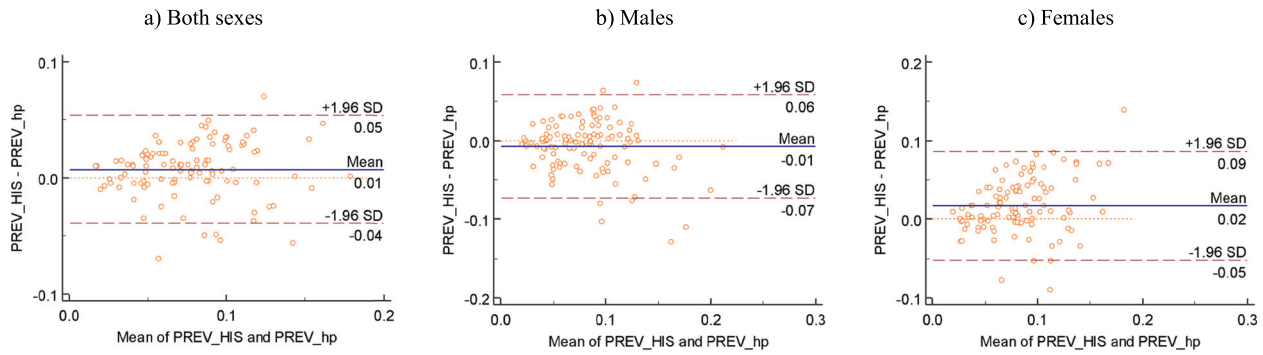


Fig. 7 Difference plot comparing HPE and DHIS2 estimates. PREV_HIS- HIV Prevalence based on health facility data, PREV_hp-HIV prevalence based on the HPE methodology while PREV_dom- HIV prevalence based on survey data only

population. The HPE methodology requires use of a small population survey sample [10, 11] to correct for bias in indicator estimates from health facility testing data. We used Uganda AIDS indicator survey data to correct for the bias in DHIS2 so as to obtain the HIV prevalence estimates for districts in Uganda. Other population surveys such as the demographic and health surveys can be similarly combined with facility-based data to obtain general population indicator estimates for planning and decision making, especially in low resourced environments where resource constraint limits collection of large sample sizes.

We applied a weighting factor, propensity to test in a health facility, calculated using multilevel logistic regression to combine the two data sources. Individuals and cluster level predictors of testing for HIV were included in the model. Predictors of access to testing or health care system may also impact HIV disease risk as noted elsewhere [10]. Multilevel logistic regression is also appropriate for the UAIS design and enables inclusion of both individual and cluster level risk factors in the modelling process. The model also accounts for clustering [21, 25, 38].

There was no difference in prevalence estimates from the HPE and Survey based approaches but confidence intervals of the HPEs were narrower, demonstrating efficiency of the HPE methodology in obtaining population level estimates as observed elsewhere [11, 18, 39].

Strengths and limitations

We applied multilevel modeling which has multiple advantages over classical models including use of HIV risk factors at individual and cluster levels. We used data from the 2011 UAIS, a more recent survey, the Population HIV Impact Assessment (PHIA), completed in 2017 was not publicly available at the time of this study. DHS data are prone to refusal to participate, this may have bias on the results of this study as those who refuse to participate may have characteristics different from those who participated in the study. Furthermore, this study was limited to complete case analysis thus reducing the effective sample size used for the analysis. DHIS2 data includes individuals who may have tested multiple times which can lead to the use of wrong or unrepresentative denominators for individuals tested at health facilities. Studies elsewhere report repeat testing ranging from 3 to 13% [40, 41]. We further note that some health facilities, especially privately owned, do not report their data to the national DHIS2 further lowering the representativeness of health facility HIV testing data.

Conclusions

The growing demand for accurate information for programme management and policy formulation will require strategies that use all the available information efficiently with little or no additional resource investments. Countries and development partners continue to build and strengthen DHIS2 through capacity building and regular data quality assessments. We applied a simple tool, HPE methodology, to support efficient use of DHIS data in combination with small survey samples to obtain more accurate indicator estimates at district or lower administrative levels. HPE obtained in this study had reduced standard errors (by 28.8%) compared to survey-based estimates demonstrating improved accuracy and reliability of the estimates. We therefore recommend use of the methodology to combine DHIS2 data with population survey data to obtain population level indicator estimates for lower administrative levels where the survey samples are small for accurate indicator estimation.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12889-020-8436-z.

Additional file 1: Appendix 1. Population survey and DHIS2 prevalence estimates. **Appendix 2.** HPE HIV prevalence estimates and associated 95% CI. **Appendix 3.** Comparison of district prevalence estimates for the HP and survey-based estimates. **Appendix 4.** Comparison of district prevalence estimates for the HPE and DHIS2-based estimates.

Abbreviations

AIC: AIDS Information Centre; AIDS: Acquired Immune Deficiency Syndrome; AIS: Aids Indicator Survey; CI: Confidence Interval; DHIS2: District Health Information System Version 2; DHS: Demographic Health Survey; HIV: Human Immunodeficiency Virus; HPE: Hybrid Prevalence Estimate; HREC: Human Research Ethics Committee; MOH: Ministry of Health; NPHC: National Population and Housing Census; PHIA: Population HIV Impact Assessment; SE: Standard error; SSA: Sub-Saharan Africa; TASO: The AIDS Support Organization; UAIS: Uganda AIDS Indicator Survey; UNAIDS: The Joint United Nations Program on HIV/AIDS; UNCST: Uganda National Council for Science and Technology; WHO: World Health Organization

Acknowledgements

We acknowledge the Ugandan Ministry of Health and its partners in conducting the Uganda AIDS Indicator Survey, and for the permission to use both DHIS2 and UAIS datasets.

Authors' contributions

JO, JJV and JL designed the research. JO compiled and analysed the data and wrote the first manuscript draft. JL, CJ, JT, RKW provided critical content analysis, helped draft and review the manuscript. All authors have read and approved the final manuscript.

Funding

This study was supported through the Wellcome Trust, grant 107754/Z/15/Z-DELTAS Africa via Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant No. 107754/Z/15/Z) and the UK government. The views expressed in this publication are those of the authors and not necessarily those of the AAS, NEPAD Agency, Wellcome Trust or the UK government.

Availability of data and materials

The 2011 AIDS indicator survey datasets analyzed during the current study are available from https://dhsprogram.com/what-we-do/survey/survey-display-373.cfm. Health facility HIV testing dataset can be accessed from Ministry of Health, Uganda following a reasonable request. Ethics clearance from Uganda National Council for Science and Technology (UNCST) is required to access the data. All data used in this study were identified by participant unique IDs, no additional identifying information was included in the data.

Ethics approval and consent to participate

Ethical clearance to conduct this study was obtained from the University of Witwatersrand Human Research Ethics Committee (HREC), clearance Certificate number M171053. Further clearance was obtained from the Uganda National Council for Science and Technology (UNCST) with registration number HS2366. Data for the study was obtained from surveys conducted in Uganda. The study was a secondary analysis of data and therefore consent to participate is not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Division of Epidemiology and Biostatistics, School of Public Health, University of Witwatersrand, Johannesburg, South Africa. [2]METRe Group, Department of International Health, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK. [3]Department of Disease Control and Environmental Health, Makerere University School of Public Health, Kampala, Uganda. [4]Department of Population Health, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK.

Received: 10 January 2020 Accepted: 28 February 2020
Published online: 23 March 2020

References

1. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Monitoring HIV impact using Population-Based Surveys. 2015.
2. McGovern ME, Marra G, Radice R, Canning D, Newell M-L, Barnighausen T. Adjusting HIV prevalence estimates for non-participation: an application to demographic surveillance. J Int AIDS Soc. 2015;18(1):19954.
3. Marston M, Harriss K, Slaymaker E. Non-response bias in estimates of HIV prevalence due to the mobility of absentees in national population-based surveys: a study of nine national surveys. Sex Transm Infect. 2008;84(Suppl 1):71–8.
4. Vinod M, Hong R, Khan S, Gu Y. Evaluating HIV Estimates from National Population-Based Surveys for Bias resulting from non-Reponse. DHS analytical studies no. 12. Calverton, Maryland: Macro International Inc.; 2008.
5. Chan M, Kazatchkine M, Lob-levyt J, Obaid T, Schweizer J, Veneman A, et al. Meeting the demand for results and accountability : a call for action on health data from eight Global Health agencies. PLoS Med. 2010;7(1):5–8.
6. UNAIDS. ENDING AIDS: Progress Towards the 90–90-90 Targets. Global Aids Update. 2017. Available from: http://www.unaids.org/sites/default/files/media_asset/Global_AIDS_update_2017_en.pdf.
7. Rice B, Boulle A, Baral S, Egger M, Mee P, Fearon E, et al. Strengthening Routine Data Systems to Track the HIV Epidemic and Guide the Response in Sub-Saharan Africa. JMIR Public Heal Surveill. 2018;4(2):e36.
8. Sheng B, Marsh K, Slavkovic AB, Gregson S, Eaton JW, Bao L. Statistical models for incorporating data from routine HIV testing of pregnant women at antenatal clinics into HIV/AIDS epidemic estimates. AIDS. 2017;31(Suppl 1):S87–94.
9. Kiberu VM, Matovu JK, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. BMC Med Inform Decis Mak. 2014;14(1):40.
10. Hedt BL, Pagano M. Health indicators: eliminating bias from convenience sampling estimators. Stat Med. 2011;30(5):560–8.
11. Jeffery C, Pagano M, Hemingway J, Valadez JJ. Hybrid prevalence estimation : method to improve intervention coverage estimations. PNAS. 2018;115(51):13063–8.
12. Avenir Health. Spectrum Manual: Spectrum System of Policy Models. Available from: https://www.avenirhealth.org/Download/Spectrum/Manuals/AIMManualEnglish.pdf.
13. Larmarange J, Bendaud V. HIV estimates at second subnational level from national population-based surveys. AIDS. 2014;28(Suppl 4):S469–2476.
14. UNAIDS. Developing Subnational Estimates of HIV Prevalence and the Number of People Living with HIV. 2014.
15. Wilson KC, Mhangara M, Dzangare J, Eaton JW, Hallett TB, Mugurungi O, et al. Does nonlocal women's attendance at antenatal clinics distort HIV prevalence surveillance estimates in pregnant women in Zimbabwe? AIDS. 2017;31(Suppl 1):S95–102.
16. Zaba BW, Carpenter LM, Boerma JT, Gregson S, Nakiyingi J, Urassa M. Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa. AIDS. 2000;14(17):2741–50.
17. Manda S, Masenyetse L, Cai B, Meyer R. Mapping HIV prevalence using population and antenatal sentinel-based HIV surveys: a multi-stage approach. Popul Health Metrics. 2015;13:22.
18. Gregson S, Terceiria N, Kakowa M, Mason PR, Anderson RM, Chandiwana SKCM. Study of bias in antenatal clinic HIV-1 surveillance data in a high contraceptive prevalence population in sub-Saharan Africa. AIDS. 2002;16(4):643–52.
19. Ministry of Heath and ICF international. Uganda AIDS Indicator Survey (AIS) 2011. Kampala Uganda and Rockville, Maryland, USA; 2012. Available from: http://health.go.ug/docs/UAIS_2011_REPORT.pdf.
20. Uganda Bureau of Statistics. 2002 Uganda Population and Housing Census Administrative Report 2007.
21. Carle AC. Fitting multilevel models in complex survey data with design weights: recommendations. BMC Med Res Methodol. 2009;9(1):1–13.
22. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC; 2017. Available from: https://www.stata.com/.
23. Ministry of Health Kampala Uganda. No Title. Electronic Health Management Information System. Available from: http://www.health.go.ug/oldsite/node/76.

24. World Health Organization. Consolidated guidelines on person-centred HIV patient monitoring and case surveillance. AIDS. 2017;31(Suppl 1):S87–94 Available from: http://www.differentiatedcare.org/Portals/0/adam/Content/Y1Jnet-yCkO4DGu8XS9VMg/File/9789241512633-eng.pdf.
25. Rabe-Hesketh S, Skrondal A. Multilevel modelling of complex survey data. J R Stat Soc Ser A Stat Soc. 2006;169(4):805–27.
26. Bland U, Giavarina D. Lessons in biostatistics Biochemia Medica. 2015;25(2):141–51.
27. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-10.
28. Core Team R. R: a language and environment for statistical computing [internet]. Vienna: The R Foundation; 2013. Available from: http://www.r-project.org.
29. Gouws E, Mishra V, Fowler TB. Comparison of adult HIV prevalence from national population-based surveys and antenatal clinic surveillance in countries with generalised epidemics: Implications for calibrating surveillance data. Sex Transm Infect. 2008;84(Suppl I):i17–i23.
30. Muyunda B, Musonda P, Mee P, Todd J, Michelo C. Educational attainment as a predictor of HIV testing uptake among women of child-bearing age: analysis of 2014 demographic and health survey in Zambia. Front Public Heal. 2018;6:192.
31. Kiros G, Workagegn F, Gebretsadik LA. Predictors of HIV-test utilization in PMTCT among antenatal care attendees in government health centers: institution-based cross-sectional study using health belief model in Addis Ababa, Ethiopia, 2013. HIV/AIDS - Res Palliat Care. 2015;215–22.
32. Gregson S, Dharmayat K, Pereboom M, Takaruza A, Mugurungi O, Schur N, Nyamukapa CA. Do HIV prevalence trends in antenatal clinic surveillance represent trends in the general population in the antiretroviral therapy era? The case of Manicaland, East Zimbabwe. AIDS. 2015;29(14):1845–53.
33. Fabiani M, Fylkesnes K, Nattabi B, Ayella EO, Declich S. Evaluating two adjustment methods to extrapolate HIV prevalence from pregnant women to the general female population in sub-Saharan Africa. AIDS. 2003;17(3):399–405.
34. Musinguzi J, Kirungi W, Opio A, Montana L, Mishra V, Madraa E, et al. Comparison of HIV prevalence estimates from sentinel surveillance and a National Population-Based Survey in Uganda, 2004-2005. JAIDS J Acquir Immune Defic Syndr. 2009;51(1):78–84.
35. Kwesigabo G, Killewo JZ, Urassa W, Mbena E, Mhalu F, Lugalla JL, et al. Monitoring of HIV-1 infection prevalence and trends in the general population using pregnant women as a sentinel population: 9 years experience from the Kagera region of Tanzania. J Acquir Immune Defic Syndr. 2000;23(5):410–7.
36. Fylkesnes K, Musonda RM, Sichone M, Ndhlovu Z, Tembo F, Monze M. Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population-based surveys. AIDS. 2001;15(7):907–16.
37. Fylkesnes K, Ndhlovu Z, Kasumba K, Musonda RM, Sichone M. Studying dynamics of the HIV epidemic. AIDS. 1998;12(10):1227–42.
38. Wong GY, Mason WM. The hierarchical logistic regression model for multilevel analysis. J Am Stat Assoc. 1985;80(391):513–24.
39. Judith RG, Anne B, Michel C, Rosemary MM, Maina K, Isaac M, Francis TLZ. Factors influencing the difference in HIV prevalence between antenatal clinic and general population in sub- Saharan Africa. AIDS. 2001;15:1717–25.
40. Kulkarni S, Tymejczyk O, Gadisa T, Lahuerta M, Remien RH, Melaku Z, et al. "Testing, testing": multiple HIV-positive tests among patients initiating antiretroviral therapy in Ethiopia. J Int Assoc Provid AIDS Care. 2017;16(6):546–54.
41. Maina I, Wanjala P, Soti D, Kipruto H, Boerma T. Using health-facility data to assess subnational coverage of maternal and child health indicators, Kenya. Bull World Health Organ. 2017;95:683–94.

Publisher's Note

Appendix VII: Original paper three: Plos One,

PLOS ONE

RESEARCH ARTICLE

Model-based small area estimation methods and precise district-level HIV prevalence estimates in Uganda

Joseph Ouma[1]*, Caroline Jeffery[2], Colletar Anna Awor[3], Allan Muruta[4], Joshua Musinguzi[4], Rhoda K. Wanyenze[5], Sam Biraro[6], Jonathan Levin[1], Joseph J. Valadez[2]

1 Division of Epidemiology and Biostatistics, School of Public Health, University of Witwatersrand, Johannesburg, South Africa, **2** METRe Group, Department of International Health, Liverpool School of Tropical Medicine, Liverpool, United Kingdom, **3** Data Science and Informatics Branch, Centers for Disease Control and Prevention, Uganda, **4** AIDS Control Program, Ministry of Health, Uganda, **5** Department of Disease Control and Environmental Health, Makerere University School of Public Health, Kampala, Uganda, **6** ICAP at Columbia University, Nakasero, Kampala, Uganda

* oumajosephd@gmail.com

Citation: Ouma J, Jeffery C, Awor CA, Muruta A, Musinguzi J, Wanyenze RK, et al. (2021) Model-based small area estimation methods and precise district-level HIV prevalence estimates in Uganda. PLoS ONE 16(8): e0253375. https://doi.org/10.1371/journal.pone.0253375

Editor: José Antonio Ortega, University of Salamanca, SPAIN

Received: August 15, 2020

Accepted: June 3, 2021

Published: August 6, 2021

Data Availability Statement: Regarding data availability, a third party collected data used for this research. The data was collected by different agencies as follows. 1. Uganda Population and HIV Impact Assessment (UPHIA) 2016/17 was collected as collaborative effort by Uganda Bureau of Statistics, ICAP at Columbia University, Ministry of Health Uganda, Centres for Disease Control and Prevention. The dataset has not been released publicly and can be accessed upon request through the Chairperson UPHIA Data Advisory Committee. Dr Musinguzi Joshua, email -

Abstract

Background

Model-based small area estimation methods can help generate parameter estimates at the district level, where planned population survey sample sizes are not large enough to support direct estimates of HIV prevalence with adequate precision. We computed district-level HIV prevalence estimates and their 95% confidence intervals for districts in Uganda.

Methods

Our analysis used direct survey and model-based estimation methods, including Fay-Herriot (area-level) and Battese-Harter-Fuller (unit-level) small area models. We used regression analysis to assess for consistency in estimating HIV prevalence. We use a ratio analysis of the mean square error and the coefficient of variation of the estimates to evaluate precision. The models were applied to Uganda Population-Based HIV Impact Assessment 2016/2017 data with auxiliary information from the 2016 Lot Quality Assurance Sampling survey and antenatal care data from district health information system datasets for unit-level and area-level models, respectively.

Results

Estimates from the model-based and the direct survey methods were similar. However, direct survey estimates were unstable compared with the model-based estimates. Area-level model estimates were more stable than unit-level model estimates. The correlation between unit-level and direct survey estimates was ($\beta_1 = 0.66$, $r^2 = 0.862$), and correlation between area-level model and direct survey estimates was ($\beta_1 = 0.44$, $r^2 = 0.698$). The error associated with the estimates decreased by 37.5% and 33.1% for the unit-level and area-level models, respectively, compared to the direct survey estimates.

jmusinguzi@infocom.co.ug 2. Data from the Uganda Population and Housing Census, conducted in 2014 is available from the census report (https://www.ubos.org/wp-content/uploads/publications/03_20182014_National_Census_Main_Report.pdf). This data is publicly available and can be extracted from the report as described in the methods section of our study. 3. Antenatal HIV testing data, extracted from DHIS2, can been accessed upon request from Ministry of health Uganda (https://www.health.go.ug/), through the permanent secretary, Ministry of Health Uganda. Dr Diana Atwine. email: diana.atwine@health.go.ug 4. Lot Quality Assurance Sampling survey data can also be accessed upon request from ministry of local government (https://molg.go.ug/) or Ministry of Health through the permanent secretary Dr Diana Atwine. Email: diana.atwine@health.go.ug. It contains individual level sensitive information We confirm that none of the authors had special privileges to access any of the datasets. We obtained ethical clearance from the Uganda National Council of Science and Technology to use the datasets. This is a key requirement for anyone accessing these datasets in Uganda.

Funding: This study was supported through the DELTAS Africa Initiative SSACAB (Grant No.107754/Z/15/Z). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant No. 107754/Z/15/Z) and the UK government. This research has been supported in part by the US President's Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention (CDC under the terms of cooperative agreement U2GGH001226. The findings and conclusions are those of the authors and do not necessarily represent the official position of the funding agencies.

Competing interests: The authors have declared that no competing interests exist.

Conclusions

Although the unit-level model estimates were less precise than the area-level model estimates, they were highly correlated with the direct survey estimates and had less standard error associated with estimates than the area-level model. Unit-level models provide more accurate and reliable data to support local decision-making when unit-level auxiliary information is available.

Introduction

Model-based small area estimation (SAE) methods can help monitor the impact of public health interventions and appropriately allocate resources in small geographical areas where the domain-specific sample is not large enough to support direct estimates of adequate precision. Other terms used to refer to small geographical areas include "small domain" or "local area" [1]. SAE methods link a study/outcome variable with auxiliary data from other sources to produce more precise indicator estimates than direct-survey estimates (i.e., design-based estimates based on survey data alone) for the target local area.

Sources of auxiliary information may include routine administrative data from Health Information Systems (HMIS), censuses or other surveys. However, external covariate data are limited for predicting local area estimates. For example, routine data only capture information from individuals who interact with health facilities [2–4]. A study combining national population survey and routine data found a 28 improvement in the precision of the estimates [5]. General population censuses, on the other hand, are conducted decennially and do not capture information in the interim or information about HIV/AIDS risk factors such as number of sexual partners or condom use during last high-risk sex [6–12], rendering these censuses unsuitable for assessing outcomes that change rapidly. Annual HIV risk factor surveys with adequate level of precision, such as in the community Lot Quality Assurance Surveys (LQAS) conducted annually in Uganda districts, help generate timely and reliable estimates of district-level HIV prevalence.

The LQAS methodology is recommended as a tool for monitoring public health interventions [13, 14]. In Uganda, the LQAS methodology is used to monitor district-level health service interventions annually [15, 16].

Model-based SAE methods are classified into two types: 1) Unit-level models such as the Battese-Harter-Fuller model [17], which links the study or outcome variable with unit or individual-level auxiliary variables (e.g. HIV status as an outcome and individuals' sex as the auxiliary variable) and 2) Area-level models such as the Fay-Herriot model [1], which links the study variable with summary or aggregate data of the auxiliary variable at the target geographical area (e.g., direct survey-based HIV prevalence as the outcome and percent of individuals who are women in a district as the auxiliary variable). Area-level models are applied if individual or unit-level covariate data are not available [18, 19] and are more popular than unit-level SAE methods because of the ease in accessing aggregate area-specific covariate information. Area-level models assume homogeneity of units within an area and ignore internal variability between and units within the area. Unit-level model parameters are estimated more accurately using sampling unit-level observations [1]. Unit-level models are efficient and are associated with small mean square errors (MSEs) compared to the area-level models [20].

SAE studies in Africa have been limited to area-based models to estimate institutional births in Ghana [21]; to identify unmet need for contraceptives in Ghana [22], and to estimate HIV

prevalence in South Africa [23]. To our knowledge, our study is the first to use unit-level models to estimate local HIV prevalence in Uganda.

Uganda's HIV prevalence distribution varies across geographical regions and among population groups. National HIV prevalence is 6.2% among persons aged 15–64 years (women, 7.6%; men, 4.7%) and varies from 3.1% to 8.0% across regions [24]. Availability of district-level prevalence information is therefore critical for resources allocation and decision-making.

We applied unit-level and area-level models to estimate HIV prevalence for districts in Uganda. We compared the precision of the model-based estimates to direct survey estimates and the precision of the unit-level model to the area-level model estimates. Our findings include several alternative district-level estimates that may be helpful for monitoring district-level HIV/AIDS intervention programs.

Methods

Data sources

Uganda Population HIV Impact Assessment. The Uganda Population HIV Impact Assessment (UPHIA) 2016–2017 used a two-stage, stratified cluster-sampling design. In the first stage, 520 enumeration areas (EAs) or clusters were randomly selected using probability proportional to the number of households in the cluster; in the second stage, a sample of 25 households were randomly selected using equal probabilities from each EA. The EAs were based on the 2014 National Population and Housing Census (NPHC) [25]. Uganda's Ministry of Health conducted the survey with technical support from ICAP at Columbia University, Centers for Disease Control and Prevention, and ICF/Macro International. For detailed survey information, see the official survey report [24]. For our study, we analyzed data from 16,828 adults aged 15–64 years from 70 districts; participants provided written informed consent were tested for HIV during the survey.

Lot quality assurance sampling surveys. In Uganda, annual LQAS surveys are used to obtain district-level indicator estimates for monitoring health interventions [26]. In LQAS, each program area is defined to be a district subdivided into 4–7 supervision areas (SA). SA comprise either a sub-county or neighboring sub-counties with similar socioeconomic characteristics. Using probability proportional to the number of households in the village, survey staff randomly select a sample of 19 villages for districts with 5–7 SAs and 24 villages for those with four SAs. A minimum sample of 19 and 24 respondents per SA for districts with 5–7 and four SAs, respectively, are selected to maintain the combined SA misclassification (Alpha + Beta) errors to $<15\%$. These sample sizes enable computation of district (program)-level coverage with $<10\%$ error margin for indicators [27]. A reference household is randomly identified using equal probability sampling within the SA and the next nearest household from the exit of the reference household is selected to start the survey. Eligible and consenting adult respondents are interviewed from selected households. Parallel sampling is applied to select eligible respondents for the following population categories: mothers of children aged 0–59 months, youth aged 15–24 years, women aged 15–49 years, and men aged 15–54 years. A sample of 19 or 24 respondents are selected for each of the population groups. The design does not permit selection of youth, men, and women from the same household because similar indicators are assessed in these population groups. For full details, see the survey reports [16]. We analyzed data from youth, men, and women.

Other sources of covariate data. Summary district level covariate data were obtained from the NPHC 2014 [28], and HIV prevalence from antenatal care attendance was obtained from the District Health Information System version 2 (DHIS2) [29]. Variables obtained from the NPHC 2014 include district population density, percentage of the population living in

urban areas, and proportion of individuals who accessed a health facility in the 12 months preceding the survey.

Statistical analysis

We computed district-level HIV prevalence estimates for 70 districts that conducted both UPHIA and LQAS surveys in 2016 using direct survey, area-level SAE models, and unit-level SAE models. We further assessed the estimates for consistency in estimating HIV prevalence and compared the precision of the estimates via the confidence intervals and the coefficient of variation of the estimates.

1. Direct survey estimates. If y_{ij} is the HIV status (positive, 1; negative, 0) of the *j-th* individual in the *i-th* district and p_i is the proportion of HIV-positive people in district *i*, then taking into account the sampling weights, the direct estimate of district HIV prevalence is obtained as follows:

$$p_i = A/B \tag{1.1}$$

Where $A = \sum_i \sum_j w_{ij} y_{ij}$ and $B = \sum_i \sum_j w_{ij}$ and $i = 1,2, \ldots, m$, $j = 1,2, \ldots, N_i$.

Where m is the number of districts and N_i is the number of individuals in district *i*.

Using the direct survey estimate (Eq 1.1), we used the UPHIA 2016–2017 dataset to compute weighted district HIV prevalence estimates and associated standard errors (SE) based on the sampled observation from each district. More details about the survey weights are available [24]. SE were computed using standard survey estimation (linearization) methods.

2. Area-level model estimation. The multivariate Fay-Herriot model [1] is the most commonly used explicit area-level model to estimate area parameters when area-level auxiliary data are available. In area-level models, the outcome obtained from direct estimation is regressed against summary/aggregate explanatory variables that are available only at the administrative/geographical area of interest [18]. The model is defined in two stages: developing a sampling model for the direct survey estimates and applying a linking model to obtain area-level parameter estimates.

2.1. Model estimation and prediction. Taking $\theta_i = g(\bar{y}_i)$, assuming g(.) is a logit link function, and relating it to a vector of p area-level covariates, $\boldsymbol{z_i}$, where $\boldsymbol{z}_i = (z_{1i}, z_{2i}, \ldots, z_{pi})$ for the m areas, the linking model for the area-level parameter θ_i is defined as

$$\boldsymbol{\theta}_i = \boldsymbol{z}_i^T \boldsymbol{\beta} + \boldsymbol{\upsilon}_i \tag{2.1}$$

Where: $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^T$ is a $p \times 1$ vector of regression coefficients and $\upsilon_i's$ are area-specific random effects assumed to be independent and identically distributed (i.i.d.) with $E(\upsilon_i) = 0$ and $var\ (\upsilon_i) = \sigma_\upsilon^2$ (i.e., $\upsilon_i \sim N\,(0, \sigma_\upsilon^2)$). The area-level random effects, υ_i, capture the unstructured heterogeneity among the areas (districts) not explained by the sampling error variance.

The unbiased direct estimator of θ_i is obtained using a sampling model in model 2.2.

$$\hat{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_i + \boldsymbol{e}_i \tag{2.2}$$

for $i = 1,2, \ldots, m$, $e_i \sim N\,(0, \psi_i)$. Where e_i is the sampling error with known sampling variance, $Var\ (e_i) = \psi_i$ and $E(e_i) = 0$ for all areas.

Model 2.2 is referred to as the sampling model because θ_i is unobservable and is estimated based on the sampled observations in the area.

Combining 2.1 and 2.2, we obtain the mixed model

$$\hat{\boldsymbol{\theta}}_i = \boldsymbol{z}_i^T\boldsymbol{\beta} + \boldsymbol{\upsilon}_i + \boldsymbol{e}_i \tag{2.3}$$

Where $\upsilon_i \sim N(0,\ \sigma_v^2)$, $e_i \sim N(0,\psi_i)$, $i = 1,2,\ldots,m$ and υ_i is independent of e_i

The Fay-Herriot, area-level model estimate is then obtained as a weighted combination of the direct ($\hat{\theta}_i$) and regression-synthetic estimators ($z_i^T\beta$).

$$\hat{\boldsymbol{\theta}}_i^{FH} = \hat{\boldsymbol{\gamma}}_i\hat{\boldsymbol{\theta}}_i + (\mathbf{1}-\hat{\boldsymbol{\gamma}}_i)\boldsymbol{z}_i^T\boldsymbol{\beta} \tag{2.4}$$

Where: $\hat{\gamma}_i = \frac{\hat{\sigma}_v^2}{\hat{\sigma}_v^2+\sigma_i^2}$

The weighting component is the ratio of the model error variance to the total variance. From Eq 2.4, the estimate $\hat{\theta}_i^{FH}$ tends to $\hat{\theta}_i$ for large values of the model variance ($\hat{\sigma}_v^2$) and tends to $z_i^T\beta$ for small values of the model variance relative to σ_i^2. Model parameters and SE are estimated using maximum likelihood methods [18, 30].

2.2. Auxiliary variables for the area-level model. Variants of the area-level model were fitted with different combinations of the predictor variables as shown in S1 File. The general population direct HIV prevalence estimate (p_i) was found to be related to HIV prevalence from ANC attendance (P_{ANCi}) $y_i = logit(p_i)$ and $\boldsymbol{z}_i = logit(p_{ANCi})$. This model had the lowest value of the Akaike Information Criterion (AIC) (Table 1 in S1 File). The models were fitted using the SAE package in R version 3.6.2 [31].

3. Unit-level model estimation. When unit-level or individual-level auxiliary data $\boldsymbol{X}_{ij} = (x_{ij1},\ x_{ij2},\ ,x_{ij3},\ldots.x_{ijq},)$ for a vector of q auxiliary variables are available, the Battese-Harter-Fuller [17] unit-level model is often applied to obtain small area parameter estimates. Under the unit-level model, data for both the outcome and the explanatory variables are available for each population element in the district, irrespective of the administrative area/domain of interest [18].

3.1. Model estimation and prediction. Letting p_{ij} be the probability of individual j from district i being HIV positive ($p_{ij} = Pr(y_{ij} = 1 \mid x_{ij},\ \mu_i)$, where y_{ij} corresponds to the HIV status of individual i in district, a logistic regression model with area-level effect is used to estimate p_i [18, 32].

$$\text{logit}(p_{ij}) = \log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \boldsymbol{X}_{\boldsymbol{ij}}^{\boldsymbol{t}}\boldsymbol{\beta} + \upsilon_i + e_{ij} \tag{3.1}$$

Where y_{ij} is assumed to be independent Bernoulli (p_{ij}) conditioned on p_{ij}'s with area random effects υ_i; β is the vector of regression model parameters. υ_i is assumed to be independent and identically distributed with $E(\upsilon_i) = 0$ and $Var\ (\upsilon_i) = \sigma_v^2$.

Assuming absence of area-level auxiliary information, the indirect estimators of district HIV prevalence, based on model 3.1 is the empirical best predictor obtained as follows:

$$\hat{p}_i = \frac{1}{N_i}\left\{\sum\nolimits_{j\in S} y_{ij} + \sum\nolimits_{j\in S'}\hat{p}_{ij}\right\} \tag{3.2}$$

Where $\hat{p}_{ij} = \frac{exp\,(\hat{\eta}_{ij})}{1+exp\,(\hat{\eta}_{ij})}$ and $\hat{\eta}_{ij} = \hat{\beta}_i X^t{}_{ij} + \hat{v}_i$

The component $\sum_{j\in S} y_{ij}$, is the sum of n_i values of HIV infection for sampled individuals from the *i-th* district while $\sum_{j\in S'}\hat{p}_{ij}$ is the sum over the estimated probability of infection for the non-sampled individuals in district i, and N_i corresponds to number of individuals in each

district. Model fitting and parameter estimation were implemented using the SAE [31] package in R software, version 3.6.2 [31]

The MSE of $\hat{p}_i$ is obtained using the parametric bootstrap estimation method for finite populations [33, 34] as described in S2 File.

3.2. Auxiliary variables for the unit-level model. Auxiliary variables from the 2016 LQAS data include age group in years (15–19, 20–24, 25–34, 35–44, and ≥45), sex (male and female), level of education (none, primary, secondary, and tertiary), marital status (single, married, and previously married: widowed/divorced/separated) and number of sexual partners including spouse in the 12 months preceding the survey (0, 1, and ≥2). We selected these variables because they were significantly associated with HIV positivity in our previous study [6].

4. Comparison of the unit-level and area-level model estimates. We used summary statistics, regression analysis, and graphical assessment for consistency to compare estimates from direct, area-level, and unit-level models. We assessed gain in precision of the model-based estimates compared to the direct survey estimates using the coefficient of variation of the estimate and ratio of the MSE. We further computed ratios of the unit-level model and the area-level model estimates to assess for improvement in precision of the unit-level model estimates.

Ethics approval and consent to participate

Ethical clearance to conduct this study was obtained from the University of Witwatersrand Human Research Ethics Committee (HREC), clearance Certificate number M171053. Further clearance was obtained from the Uganda National Council for Science and Technology (UNCST) with registration number HS2366. Data for the study were obtained from surveys conducted in Uganda. The study was a secondary analysis of data, so consent to participate is not applicable.

Results

Selected characteristics of survey participants

For both surveys, most respondents were women, had incomplete primary education, were married/cohabiting, and had one sexual partner in the 12 months preceding the surveys (Table 1).

HIV prevalence estimates

HIV prevalence estimates for the districts included in the analysis are summarized in Table 2. District-specific estimates with 95% confidence intervals are presented in S1 Table. On average, direct survey estimates were higher (mean = 0.064, SD = 0.034) than area-level (mean = 0.056, SD = 0.018) and unit-level model (mean = 0.058, SD = 0.026) estimates. Direct survey estimates also had a higher variation than the model-based estimates, with minimum and maximum values of 0.010 and 0.148, respectively. The estimates from the area-level model had the least variation (Table 2).

Unit-level model HIV prevalence estimates Fig 1C had a similar pattern compared to the direct survey prevalence estimates (Fig 1A). HIV prevalence estimates based on the area-level model had the least recognizable pattern between districts (Fig 1B) consistent with the summary statistics in Table 2. HIV prevalence was generally higher in districts in Central, South Western, and Northern regions of Uganda and in districts bordering lakes.

Model diagnostics

We regressed model-based estimates against direct survey estimates to assess bias and reliability of the model-based estimates. The bias diagnostics plot also presents a scatter plot of the estimates and shows the effect of extreme values in the estimates. Unit-level model estimates

Table 1. Characteristics of respondents.

	UPHIA 2016	LQAS 2016
Characteristic	**Weighted % (n = 16,862)**	**%(n = 34,109)**
Sex		
Male	47.9 (7,302)	48.5 (16,545)
Female	52.1 (9,560)	51.5 (17,562)
Age, years		
15–19	23.9 (3,649)	23.5 (7,997)
20–24	18.6 (2,859)	24.6 (8,403)
25–34	25.2 (4,190)	23.3 (7,946)
35–44	16.3 (2,885)	18.8 (6,407)
45–64	14.4 (3,279)	9.8 (3,356)
Education level#		
No education	7.8 (1,582)	7.4 (2,530)
Incomplete primary	43.7 (7,663)	42.6 (14,538)
Primary	16.2 (2,604)	23.2 (7,905)
Incomplete secondary	24.0 (3,695)	17.1 (5,845)
Complete secondary+	8.3 (1,218)	9.7 (3,291)
Marital status##		
Never married	31.1 (4,620)	33.5 (11,435)
Married/cohabiting	55.7 (9,775)	61.6 (21,022)
Widowed	3.5 (705)	1.5 (524)
Separated	9.8 (1,728)	3.3 (1,127)
Number of sexual partners in last 12 months###		
0	14.1 (2,131)	31.5 (9,723)
1	70.2 (10,261)	57.5 (17,760)
≥2	15.6 (2,125)	11.0 (3,410)

Notes: UPHIA 2016 survey data weighted using the population survey weight. Abbreviations: UPHIA, Uganda Population-Based HIV Impact Assessment; LQAS, Lot Quality Assurance Surveys. Data missing for #-101, ##-39 and ###-2345 respondents.

https://doi.org/10.1371/journal.pone.0253375.t001

were strongly correlated with the direct survey estimates ($\beta_1 = 0.66$; $r^2 = 0.862$; Fig 2B) compared to area-level model estimates ($\beta_1 = 0.44$; $r^2 = 0.698$; Fig 2A).

Precision and consistency of HIV prevalence estimates

The model-based estimates were similar to the direct survey estimates, but the point estimates had less variation compared to the direct survey estimates (Fig 3A and 3C). We also note that direct survey estimates were significantly different (i.e., they were either higher or lower) compared with model-based estimates for small survey sample sizes Fig 3A and 3C), demonstrating the shrinkage of the model-based estimates toward the point estimates. However, direct survey and model-based estimates tended to be similar with

Table 2. Summary of district-level HIV prevalence estimates in Uganda.

	Mean (SD)	Minimum	Maximum	First Quartile (Q1)	Second Quartile (Q2)	Third Quartile (Q3)
Direct survey	0.064 (0.034)	0.010	0.148	0.038	0.055	0.089
Area-level model	0.056 (0.018)	0.0157	0.096	0.042	0.058	0.067
Unit-level model	0.058 (0.026)	0.004	0.142	0.042	0.042	0.074

https://doi.org/10.1371/journal.pone.0253375.t002

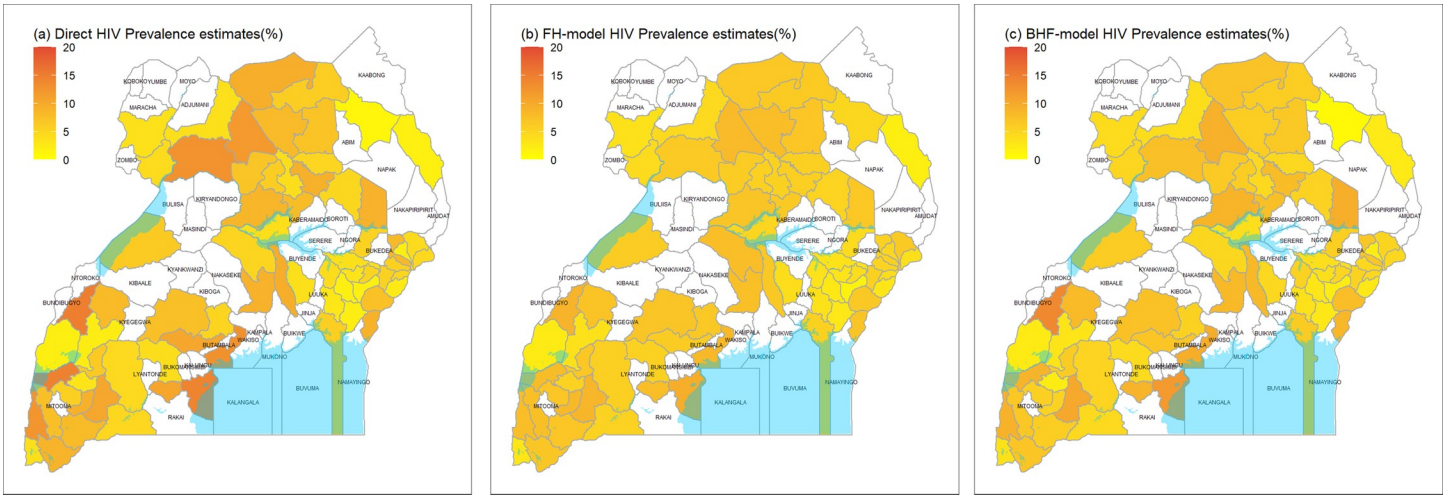


Fig 1. Comparison of HIV prevalence estimates in Uganda. Fig 1 presents district-level prevalence estimates: (a) Direct survey estimates, (b) area-level model estimates, and (c) unit-level model estimates. Scales for each of the maps were maintained to show the extent of extreme values. Unshaded areas represent district that did not complete LQAS surveys in 2016.

https://doi.org/10.1371/journal.pone.0253375.g001

increasing survey sample sizes in the districts. For example, Nwoya district with survey sample of 44 individuals, the HIV prevalence estimate was 0.127 (95% CI: 0.000–1.000), 0.062 (95% CI: 0.034–0.091), and 0.075 (95% CI: 0.039–0.111), whereas for Mbale district with a survey sample of 874 individuals, the estimates were 0.054 (95% CI, 0.028), 0.051 (95% CI: 0.029, 0.072), and 0.046 (95% CI: 0.029, 0.063) for direct survey, area-level and unit-level models respectively (S1 Table). The average improvement in the precision of the estimates was 37.5% and 33.1% for the unit-level and area-level model, respectively (data not shown).

The coefficient of variation of the direct survey estimates were generally larger and had a higher variation compared to the coefficient of variation of the model-based estimates irrespective of the survey sample size in the districts (Fig 3B and 3D). Table 3 summarizes the coefficients of variation and shows that the mean coefficient of variation for direct survey estimates was 138.0% (346.9) compared to 22.4% (7.5) for the area-level and 23.7% (18.5) unit-level models.

Assuming estimates are considered as reliable for decision making if the coefficient of variation is <20% [31], then <50% of the districts have reliable data for decision making based on the direct survey estimates, whereas >50% of the districts would have reliable data based on the SAE methods (Table 3). Specifically, only 14 (20%), 36 (51.4%), and 36 (51.4%) of the districts would have reliable information for decision making based on direct survey estimates, area-level model estimates, and unit-level model estimates, respectively (data not shown).

Consistency of model-based estimates

Unit-level model estimates varied more than the area-level model estimates (Fig 4A). Additionally, the coefficients of variation of the unit-level model estimates were consistently larger for districts with small survey sample sizes and consistently smaller for districts with large survey sample sizes (Fig 4B). Implying that the unit-level model estimates converge more rapidly to the point estimate compared to the area-level model estimates as the survey sample sizes in the districts increased.

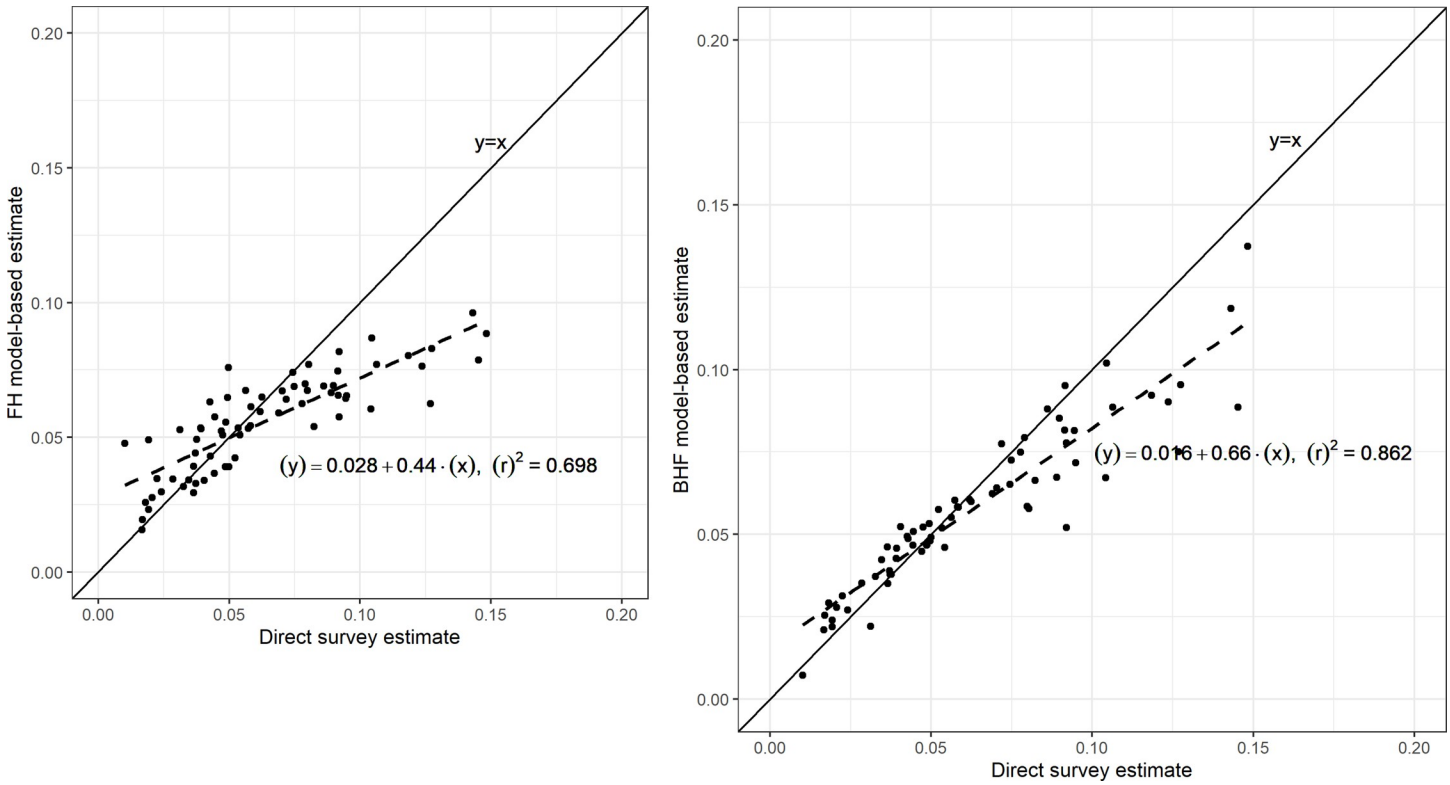


Fig 2. Correlation of model and direct survey estimates of HIV prevalence in Uganda. Regression and scatter plot of (a) area-level model estimates compared to direct survey estimates and (b) unit-level model estimates compared to direct survey estimates.

https://doi.org/10.1371/journal.pone.0253375.g002

Discussion

Study findings show the feasibility of applying SAE models to population survey data with auxiliary information from community LQAS surveys and DHIS2 data to obtain more precise HIV prevalence estimates for districts in Uganda. Both the unit-level and area-level model estimates were similar to the direct survey estimates, although the unit-level model estimates were more correlated with the direct survey estimates compared to area-level model estimates. A graphical assessment shows that the model-based estimates were close to the direct survey estimates, were less polarized, and had no extreme values/outliers. Mapping model-based HIV prevalence estimates shows a similar pattern between the unit-level estimates and the direct survey estimates. Estimates for all approaches were generally similar to regional survey prevalence estimates [24].

The coefficient of variation of both the unit-level and area-level model estimates were lower than the coefficient of variation of the direct survey estimates. However, the coefficient of variation of the unit-level model estimates were larger than the coefficient of variation of the area-level model estimates, which suggests that area-level model estimates were more precise. Coefficient of variation is computed as a ratio of the SE to the mean and expressed as a percentage. It therefore expresses the sampling variability of the estimates from the point estimate, implying that estimates with large coefficients of variation would be considered over-dispersed and therefore unreliable for decision-making.

Although population surveys provide more accurate national and regional information, they yield only partial information for district-level planning and allocating resources. Uganda, like many other low and middle-income countries, lacks the resources to collect representative data for monitoring social services at the district-level; however, in Uganda's de-centralized

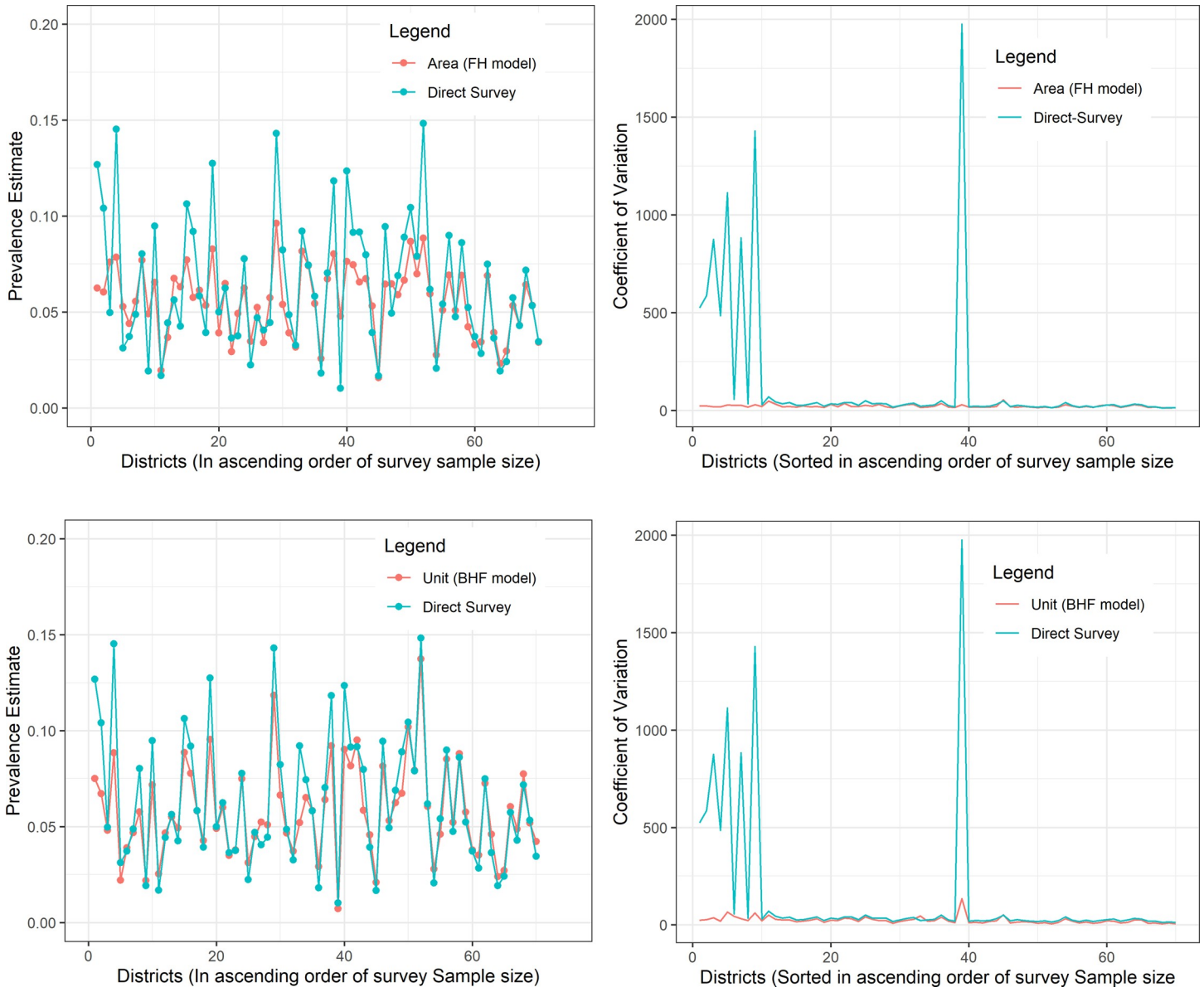


Fig 3. HIV prevalence estimates and the coefficient of variation of direct survey and model-based estimates in Uganda. HIV prevalence estimates (on the left) and the coefficient of variation (on the right) for direct survey and model-based estimates with the districts sorted in ascending order of the survey sample sizes. (a) direct survey compared to area-level model estimates using the Fay-Herriot (FH) model, (b) coefficient of variation for direct survey and area-level model estimates, (c) direct survey compared to unit-level model estimates using the Battese-Harter-Fuller (BHF) model, and (d) coefficient of variation for direct survey and unit-level model estimates.

https://doi.org/10.1371/journal.pone.0253375.g003

Table 3. Summary of coefficient of variation of HIV prevalence estimates in Uganda*.

	Mean (SD)	Minimum	Maximum	First Quartile (Q1)	Second Quartile (Q2)	Third Quartile (Q3)
Direct survey	138.0 (346.9)	13.3	1974.8	21.9	28.0	40.2
Area-level model	22.4(7.5)	13.3	54.6	17.5	19.9	19.7
Unit-level model	23.7 (18.5)	5.6	133.8	12.4	19.7	27.7

*Omitting the outlier district (Kalangala) from the analysis did not affect the model-based or direct survey estimates (results not presented). Abbreviations: SD, Standard Deviation.

https://doi.org/10.1371/journal.pone.0253375.t003

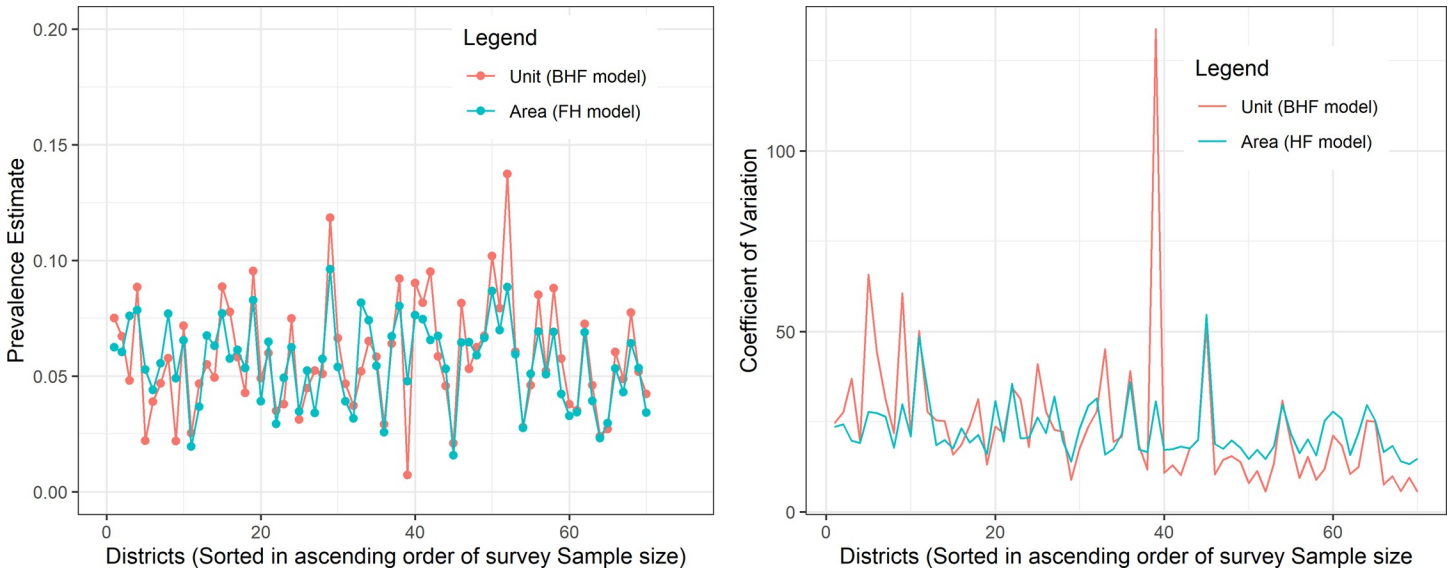


Fig 4. Area-level and unit-level model-based estimates of district HIV prevalence in Uganda. (a) HIV prevalence estimates and (b) the coefficient of variation for area-level and unit-level model-based estimates with the districts sorted in ascending order of the survey sample sizes.

https://doi.org/10.1371/journal.pone.0253375.g004

governance model, districts plan, implement, and monitor interventions on behalf of the central government. Using simple methods to generate more precise HIV prevalence estimates for districts will therefore augment district-level decision making. Application of SAE methods elsewhere in Africa have shown more precise district-level estimates [21–23] as observed in our study.

We note that the unit-level model estimates were highly correlated with direct survey estimates, although they were less precise compared to the area-level model estimates, which is contrary to SAE's methodological theory [18]. Lower correlation of the area-level model estimates with direct survey estimates compared to the unit-level model estimates versus direct survey estimates may be attributed to the aggregated area-level covariates, which mask any internal variations or heterogeneity of units. The heterogeneous spread of HIV is well documented in literature [6, 23, 24]. For example, HIV prevalence is higher in urban areas compared to rural areas [6, 23, 24]. Similarly, females have a higher HIV positivity compared to males although the differences between males and females varies by age group [24]. The difference is wider for the younger age group 15–24 years and converges with increasing age [24]. It is therefore important to take into consideration, this internal differences in obtaining the estimates as demonstrated by the BHF model applied in this study.

Additionally, use of antenatal data for HIV prevalence monitoring in the general population has limitations and biases including selection bias of routine data. Antenatal HIV surveillance data excludes non-pregnant women and men and includes information from health facilities in urban and easily accessible areas [35]; public health facilities that report to national HMIS system [20]; younger and more educated women who have higher antenatal care attendance rates [36–38]. These limitations imply that the antenatal survey data may not accurately reflect the general population HIV prevalence distribution as observed in our study.

Overall improvement in the precision of the estimates was 37.5% and 33.1% for the unit-level and area-level model, respectively. A study combining population survey with routine data found 28% improvement in the precision of estimates but lower correlation with

estimates based on routine data [5]. Use of risk factor data, representative of the general population, therefore improves the precision of the estimates compared to combining routine and survey data or use of more aggregate information.

Study findings were consistent with regional HIV prevalence estimates based on survey data. Districts with higher HIV prevalence were from regions with overall higher HIV prevalence (e.g., Kaborale district in Western region; Masaka and Mpigi in central 1 region; and Mbarara in South Western region) as observed in the national level survey [24]. These districts are urban and are major transport corridors for truck drivers. Masaka district also is inhabited by fishing communities, which typically have higher HIV prevalence [8, 9, 39, 40].

Although the model-based methods assume independence of area random effects, independence is unlikely to hold between neighboring districts. Spatial estimation approaches attempt to solve this, although these methods have their own limitations [23]. Borders and boundaries are arbitrary, and individuals tend to seek healthcare services in neighboring districts or even other regions that are convenient or offer better quality services. For example, some districts in Uganda do not have Health Centre IV or hospitals, which are known to provide better quality healthcare and a broad range of health services [23]. Furthermore, 40 out of 112 districts in the country did not conduct LQAS surveys in 2016; this implies that BHF model estimates could not be obtained for these districts limiting the breath of model evaluation.

Data-driven district-level decision making for HIV programs requires precise and reliable estimates, but survey sample sizes from population surveys are significantly smaller for districts than for regions. Our study shows that with external auxiliary data, estimates that are more precise can be obtained, but precision depends on whether the information is available at the sampling unit level or area level. Unit-level model estimates, although less precise than area-level estimates, were more consistent with direct survey estimates. This application also promotes use of annual community LQAS data, which is readily available to district service managers. SAE models also can be developed in freely available software, such as R and STATA. Models that use both area-level and unit-level covariates to obtain SAE could help provide more precise and accurate HIV prevalence estimates in other settings.

Supporting information

S1 File. Area-level model of HIV prevalence in Uganda.
(DOCX)

S2 File. Parametric bootstrap for mean square error estimation for Battese-Harter-Fuller (BHF) model.
(DOCX)

S1 Table. District HIV prevalence estimates in Uganda using direct estimates, Fay-Herriot model estimates, and Battese-Harter-Fuller model estimates. Abbreviations: CI, confidence interval.
(DOCX)

Acknowledgments

We acknowledge the Ugandan Ministry of Health and its partners for conducting the Uganda Population-Based HIV Impact Assessment and the District Health Teams for conducting community LQAS surveys; we are grateful for the permission to use DHIS2, UPHIA, and LQAS datasets.

Author Contributions

Conceptualization: Joseph Ouma, Jonathan Levin, Joseph J. Valadez.

Data curation: Joseph Ouma, Colletar Anna Awor.

Formal analysis: Joseph Ouma.

Funding acquisition: Jonathan Levin.

Methodology: Joseph Ouma, Caroline Jeffery, Colletar Anna Awor, Jonathan Levin.

Software: Joseph Ouma.

Supervision: Caroline Jeffery, Jonathan Levin, Joseph J. Valadez.

Validation: Joseph Ouma, Caroline Jeffery, Colletar Anna Awor, Rhoda K. Wanyenze.

Writing – original draft: Joseph Ouma.

Writing – review & editing: Joseph Ouma, Caroline Jeffery, Colletar Anna Awor, Allan Muruta, Joshua Musinguzi, Rhoda K. Wanyenze, Sam Biraro, Jonathan Levin, Joseph J. Valadez.

References

1. Rao JNK. Inferential issues in model-based small area estimation: Some new developments. Stat Transit. 2015; 16(4):491–510.
2. Weitoft GR, Gullberg A, Hjern A, Rosén M. Mortality statistics in immigrant research: Method for adjusting underestimation of mortality. Int J Epidemiol. 1999; 28(4):756–63. https://doi.org/10.1093/ije/28.4.756 PMID: 10480707
3. Hashimoto R, Brodt E, Skelly A, Dettori J. Administrative Database Studies: Goldmine or Goose Chase? Evid Based Spine Care J. 2014; 05(02):074–6. https://doi.org/10.1055/s-0034-1390027 PMID: 25278880
4. Van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. J Clin Epidemiol [Internet]. 2012; 65(2):126–31. Available from: https://doi.org/10.1016/j.jclinepi.2011.08.002 PMID: 22075111
5. Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Todd J, Levin J. Combining national survey with facility-based HIV testing data to obtain more accurate estimate of HIV prevalence in districts in Uganda. BMC Public Health. 2020; 20(1):1–14. https://doi.org/10.1186/s12889-019-7969-5 PMID: 31898494
6. Ouma J, Jeffery C, Valadez JJ, Wanyenze RK, Levin J. Difference in HIV prevalence by testing venue: results from population level survey in Uganda survey in Uganda. AIDS Care [Internet]. 2020;0(0):1–12. Available from: https://doi.org/10.1080/09540121.2020.1734179 PMID: 32131605
7. Jeffery C, Beckworth C, Hadden WC, Ouma J, Lwanga K, Valadez JJ, et al. Associations with HIV testing in Uganda: an analysis of the Lot Quality Assurance Sampling database 2003–2012. AIDS Care [Internet]. 2016;0(0):1–5. Available from: https://doi.org/10.1080/09540121.2015.1112350 PMID: 26586024
8. Mafigiri R, Matovu JKB, Makumbi FE, Ndyanabo A, Nabukalu D, Sakor M, et al. HIV prevalence and uptake of HIV/AIDS services among youths (15–24 Years) in fishing and neighboring communities of Kasensero, Rakai District, South Western Uganda. BMC Public Health [Internet]. 2017; 17(1):251. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=28288604 https://doi.org/10.1186/s12889-017-4166-2 PMID: 28288604
9. Lindan CP, Anglemyer A, Hladik W, Barker J, Lubwama G, Rutherford G, et al. High-risk motorcycle taxi drivers in the HIV/AIDS era: a respondent-driven sampling survey in Kampala, Uganda. Int J STD AIDS. 2015; 26(5):336–45. https://doi.org/10.1177/0956462414538006 PMID: 24970473
10. Amornkul PN, Vandenhoudt H, Nasokho P, Odhiambo F, Mwaengo D, Hightower A, et al. HIV prevalence and associated risk factors among individuals aged 13–34 years in rural Western Kenya. PLoS One. 2009; 4(7). https://doi.org/10.1371/journal.pone.0006470 PMID: 19649242
11. Asiedu C, Asiedu E, Owusu F. The Socio-Economic Determinants of HIV/AIDS Infection Rates in Lesotho, Malawi, Swaziland and Zimbabwe. Dev Policy Rev. 2012; 30(3):305–26.

12. Lakew Y, Benedict S, Haile D. Social determinants of HIV infection, hotspot areas and subpopulation groups in Ethiopia: evidence from the National Demographic and Health Survey in 2011. BMJ Open. 2015; 5(11):e008669. https://doi.org/10.1136/bmjopen-2015-008669 PMID: 26589427

13. Lanata CF, Black RE. Lot quality assurance sampling techniques in health surveys in developing countries: advantages and current constraints. World Health Stat Q. 1991; 44:133–9. PMID: 1949880

14. Lemeshow S, Taber S. Lot quality assurance sampling: single- and double-sampling plans. World Health Stat Q [Internet]. 1991; 44(3):115–32. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=1949879&retmode=ref&cmd=prlinks%5Cnpapers2://publication/uuid/77D2EBA8-8179-4C96-9AB3-5E40C9808407 PMID: 1949879

15. Businge D, Kironde S, Odong T. Utilizing lot quality assurance sampling (LQAS) surveys to guide district- and sub-district-level work-planning and decision-making: experiences from five years of implementation in East Central Uganda. 20th Int AIDS Conf July 20–25, 2014, Melbourne, Aust. 2014;4927.

16. Assessment HF. STAR-E LQAS Key results of the 2013 LQAS community surveys & Health Facility Assessment. 2013;1–27.

17. Battese GE, Harter RM, Fuller WA. An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data. J Am Stat Assoc [Internet]. 1988; 83(401):28–36. Available from: http://www.jstor.org/stable/2288915?origin=crossref%5Cnhttp://www.jstor.org/stable/2288915

18. Rao JNK, Molina I. Small Area Estimation [Internet]. Wiley; 2015. (In Wiley Online Library: Books). Available from: https://books.google.co.za/books?id=i1B_BwAAQBAJ

19. Tzavidis N, Zhang LC, Luna A, Schmid T, Rojas-Perilla N. From start to finish: a framework for the production of small area official statistics. J R Stat Soc Ser A Stat Soc. 2018; 181(4):927–79.

20. Giorgi E, Sesay SSS, Terlouw DJ, Diggle PJ. Combining data from multiple spatially referenced prevalence surveys using generalized linear geostatistical models. J R Stat Soc Ser A Stat Soc. 2015; 178 (2):445–64.

21. Johnson FA, Chandra H, Brown JJ. District-level Estimates of Institutional Births in Ghana: Application of Small Area Estimation Technique Using Census and DHS Data. 2010; 26(2):341–59.

22. Johnson FA, Padmadas SS, Chandra H, Matthews Z, Madise NJ, Amoako F, et al. Estimating unmet need for contraception by district within Ghana: An application of small-area estimation techniques. 2012; 4728(May).

23. Gutreuter S, Igumbor E, Wabiri N, Desai M, Durand L. Improving estimates of district HIV prevalence and burden in South Africa using small area estimation techniques. PLoS One [Internet]. 2019; 14(2):1–14. Available from: https://doi.org/10.1371/journal.pone.0212445 PMID: 30794619

24. Ministry of Health Uganda. Uganda Population-based HIV Impact Assessment (UPHIA) 2016–2017: Final Report. [Internet]. Kampala, Uganda; 2019. Available from: https://phia.icap.columbia.edu/wp-content/uploads/2019/07/UPHIA_Final_Report_Revise_07.11.2019_Final_for-web.pdf

25. Uganda Bureau of Statistics 2016. The National Population and Housing Census 2014 –Main Report. Kampala, Uganda; 2016.

26. Hage J, Valadez JJ. Institutionalizing and sustaining social change in health systems: The case of Uganda. Health Policy Plan. 2017; 32(9):1248–55. https://doi.org/10.1093/heapol/czx066 PMID: 28981663

27. Robertson SE, Valadez JJ. Global review of health care surveys using lot quality assurance sampling (LQAS), 1984–2004. Soc Sci Med. 2006; 63(6):1648–60. https://doi.org/10.1016/j.socscimed.2006.04.011 PMID: 16764978

28. UBOS. National housing and population census 2014-Area Specific Profiles, Wakiso District. 2017.

29. Kiberu VM, Matovu JK, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. BMC Med Inform Decis Mak. 2014; 14(1):40. https://doi.org/10.1186/1472-6947-14-40 PMID: 24886567

30. Jiang J, Lahiri P. Mixed model prediction and small area estimation. Test. 2006; 15(1):1–96.

31. Molina I, Marhuenda Y. sae: An R Package for Small Area Estimation. 2015; 7(June):81–98.

32. Chandra H, Kumar S, Aditya K. Small area estimation of proportions with different levels of auxiliary data. Biometrical J. 2018; 60(2):395–415. https://doi.org/10.1002/bimj.201600128 PMID: 29349798

33. Moretti A, Shlomo N, Sakshaug JW. Parametric bootstrap mean squared error of a small area multivariate EBLUP. Commun Stat—Simul Comput [Internet]. 2018 Dec 9;1–14. Available from: https://doi.org/10.1080/03610918.2018.1498889

34. González-Manteiga W, Lombardía MJ, Molina I, Morales D, Santamaría L. Bootstrap mean squared error of a small-area EBLUP. J Stat Comput Simul. 2008; 78(5):443–62.

35. UNAIDS/WHO. Guidelines for Conducting HIV Sentinel Serosurveys among Pregnant Women and Other Groups. 2003.

36. Gregson S, Terceiria N, Kakowa M, Mason PR, Anderson RM, Chandiwana SK CM. Study of bias in antenatal clinic HIV-1 surveillance data in a high contraceptive prevalence population in sub-Saharan Africa. AIDS. 2002; 16(4):643–52. https://doi.org/10.1097/00002030-200203080-00017 PMID: 11873009

37. Saphonn V, Hor LB, Ly SP, Chhuon S, Saidel T, Detels R. How well do antenatal clinic (ANC) attendees represent the general population? A comparison of HIV prevalence from ANC sentinel surveillance sites with a population-based survey of women aged 15–49 in Cambodia. Int J Epidemiol [Internet]. 2002; 31(2):449–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11980815 PMID: 11980815

38. Zaba BW, Carpenter LM, Boerma JT, Gregson S, Nakiyingi J, Urassa M. Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa. AIDS. 2000; 14 (17):2741–50. https://doi.org/10.1097/00002030-200012010-00014 PMID: 11125893

39. Opio A, Muyonga M, Mulumba N. HIV Infection in Fishing Communities of Lake Victoria Basin of Uganda–A Cross-Sectional Sero-Behavioral Survey. PLoS One. 2013; 8(8).

40. Hegdahl HK, Fylkesnes KM, Sandøy IF. Sex differences in HIV prevalence persist over time: Evidence from 18 countries in Sub-Saharan Africa. PLoS One [Internet]. 2016; 11(2):1–17. Available from: https://doi.org/10.1371/journal.pone.0148502 PMID: 26841112

Appendix VIII: Manuscript in draft, to be submitted to Plos one, Public health

Factors associated with decline in HIV prevalence among adults aged 15-49 years in Uganda: A decomposition analysis of 2011 and 2016 population survey data.

Joseph Ouma1§, Caroline Jeffery2, Awor Anna Colletar3, Allan Muruta4, Joshua Musinguzi4, Rhoda K. Wanyenze5, Sam Biraro6, Jonathan Levin1, Joseph J. Valadez2

1. Division of Epidemiology and Biostatistics, School of Public Health, University of Witwatersrand, Johannesburg, South Africa.

2. METRe Group, Department of International Health, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK

3. Data Science and Informatics Branch, Centers for Disease Control and Prevention, Uganda

4. AIDS Control Program, Ministry of Health, Uganda

5. Department of Disease Control and Environmental Health, Makerere University School of Public Health, Kampala, Uganda

6. ICAP at Columbia University, Uganda. 5th Floor, Lourdel Tower, Plot 1 Lourdel Road, Nakasero, Kampala, Uganda

§ Corresponding author: Joseph Ouma

Phone: +256 759 800 054
Email: oumajosephd@gmail.com

The authors declare that they have no competing interests. This study was supported through the DELTAS Africa Initiative SSACAB (Grant No. 107754/Z/15/Z). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant No. 107754/Z/15/Z) and the UK government. This research has been supported in part by the US President's Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention (CDC under the terms of cooperative agreement U2GGH001226. The findings and conclusions are those of the authors and do not necessarily represent the official position of the funding agencies.

Abstract

Introduction

Key to the strategies to curb the further spread of the HIV/AIDS virus is knowledge of the areas where efforts and resources can be re-directed. Interventions to avert the risk of infections among the high-risk population groups and enable them adopt risk reduction strategies are well developed. However to achieve the greatest impact, decomposition analysis can be applied to identify and quantify the contribution of variables and/or their coefficients to change in HIV prevalence over time. We applied the methodology to assess the change in HIV prevalence among adults aged 15-49 years tested for HIV during the 2011/12 and 2016/17 population surveys in Uganda.

Methods

We carried descriptive analysis to assess change in HIV prevalence, pooled logistic regression to assess factors associated with HIV positivity and logistic decomposition analysis to determine the contribution of the factors and their coefficients to change in HIV prevalence. All analyses were weighted using survey design weights.

Results

More individuals aged 15-19 and 20-24 years; males; individuals with secondary or tertiary level of education; previously married; individuals with one and those with 2+ sexual partners participated in the 2016/17 survey compared to 2011/12. Overall, HIV prevalence dropped by 18%. Both endowments and the coefficients of the characteristics had a significant contribution to reduction in HIV prevalence. The coefficients of age group 40-49 years contributed 59.2% of the reduction in HIV prevalence (p value<0.001).

Conclusion

Contribution of the characteristics of surveyed participants as well as the coefficients of the characteristics on change in HIV prevalence bring to the fore salient factors that underscore HIV spread. Targeting these population subgroups based on their specific and possible contribution to HIV prevalence will fasten the pace at which zero new infections occur in the general population. Other population categories to target include individuals with higher level of education, those who were previously married and those who have two or more sexual partners.

Introduction

Global efforts to halt further spread of the HIV Virus and end the pandemic has led to a general decline in new infections. By 2019, the annual number of new infections globally reduced by 40 percent from 2.8 million in 1998 to an estimated 1.7million in 2019 [1]. In Uganda the annual number of new infections reduced by 52 percent from 110,000 in 1990 to 53,000 in 2019 [1].

Although prevalence, the number of people living with HIV, has not realized a matched reduction over the years mainly due to the effective Anti-Retroviral Treatment (ART) that has enabled people to live longer with HIV, this reversal (reduction) in the number of new infections implies that it's possible to bring the epidemic to an end. Indeed the UNAIDS set ambitious targets to end the epidemic by postulating that if 90 percent of the people living with HIV knew their status; 90 percent of those who knew their status were on treatment; and 90 percent of those on treatment attained viral suppression by 2020, then the epidemic would be brought to an end by 2030[2]. While these targets may have not been achieved in many settings, there is extensive progress by all countries that adopted it. In Uganda for example achieved 84%, 87% and 88% of the targets respectively by 2019[1]. UNAIDS is currently leading the process of updating these targets to inform the global response and national strategies to be undertaken between 2021 and 2030[3].

Key to the strategies to curb the further spread of the HIV/AIDS virus is knowledge of the variables or intervention areas where efforts and resources can be re-directed. Many studies have described factors associated with decline in HIV prevalence to include increase in condom use and decline in casual sex[4–8]. Being male, young people especially young females aged 15-24years [9], residence in urban areas, being previously married are also documented as key contributors to higher prevalence[9]. Interventions to avert the risk of new infections and individual enable them to embrace risk reduction strategies have been developed. For example, the DREAMS (Determined, Resilient, Empowered, AIDS-free, Mentored and Safe) partnership was implemented in 10 high burden countries in Sub-Saharan Africa. It was able to achieve up to 40 percent decline in new HIV infections among young women aged 15-24 years by promoting HIV counseling and testing; condom promotion and provision; and improving access to youth friendly sexual and reproductive health services[10].

Knowledge of the contribution of these factors to reduction in prevalence is also important as this will enable identification of the areas were the most impact can be realized. This will further enable prioritization of intervention efforts and strategies to increase the impact of the interventions.

Decomposition analysis can be used to quantify the extent by which these variables contribute to changes in prevalence over time. The methodology stratifies the sources of effect of a factor/independent variable on an outcome into composition effects and coefficient effects[11]. Composition effect (also known as "endowments") refers to the effect due to the distribution of the survey participants by the variable of interest for example gender or age, whereas coefficient effect refers to the effect attributable to the coefficients of the variable/characteristic. This methodology has been applied in other settings including to assess change in contraceptive use[12], change in fertility preferences overtime[13] change in mortality[14], and uptake of HIV counseling and testing overtime[15].

We apply the methodology to assess the change in HIV prevalence between 2011/12 and 2016/17 among adults aged 15-49 years tested for HIV during the population surveys in Uganda. The objective of this study was therefore to identify factors associated with change in HIV prevalence between 2011 and 2016 and to assess the contribution of the endowments (characteristics) and the contribution of the coefficients of the characteristics to change in HIV prevalence.

Uganda conducts nationally representative population surveys with HIV testing to obtain HIV/AIDS indicators including HIV prevalence and uptake of HIV services. The surveys are conducted five yearly at household level, providing sufficient time to measure changes in HIV/AIDS knowledge, attitudes, behavior and potential changes in HIV prevalence at the population level[16]. The first survey was carried out in 2004/05[17] and the subsequent surveys were in 2011/12[18] and 2016/17[9]. Data from the population surveys also help to broaden understanding of the magnitude, distribution, and trends of HIV infection, as well as coverage of HIV treatment and prevention interventions. This study used data from the 2011/12 and 2016/17 surveys; data from the 2004/05 was not readily accessible.

Methods

Datasets for the study

During the 2011/12 & 2016/17 survey rounds, the country was divided into 10 geographical regions comprising 8-15 neighboring districts that share similar language and social demographic characteristics. The 10 geographical regions include Central 1, Central 2, Kampala, East-Central, Mid-Eastern, North-East, West Nile, Mid-North, Mid-West, and South-West. The surveys used a two stage, stratified cluster sampling design where Enumeration Areas (EA) were selected at the first stage using probability proportional to the number of households in the area. At the second stage, a sample of households were randomly selected from each EA using equal probabilities. All adults present in the selected households who consent to participate in the survey were interviewed and blood drawn for HIV testing. The 2011/12 survey EAs were based on the 2001 population and Housing census[19] while the 2016/17 survey EAs were based on the 2014 national housing and population census[20]. More details about the survey design and sample selection are available in the survey reports[9,18]. This study was restricted to the analysis of data for individuals aged 15-49 years who were tested for HIV during the survey.

Measures/Study variables

(1) *HIV status.* The outcome of interest for the study is HIV status. Home-based HIV testing was conducted for all consenting individuals and results provided to the respondent at home. Only final test results (either 1=Positive or 0=Negative) were provided to the survey participants.

(2) *Model explanatory/independent variables*: area of residence gender (1=male, 2=female); age (15-19, 20-24, 25-29, 30-34, 35-39 and 40-49-years); marital status (never, married/cohabiting and previously married-widowed/separated/divorced); highest level of education (non, primary, secondary and Tertiary); number of sexual partners including husband/wife in the 12 months preceding the survey (0, 1 and 2+); area of residence (1=urban, 2=rural); and region of the country (Central 1, Central 2, Kampala, East-Central, Mid-Eastern, North-East, West Nile, Mid-North, Mid-West, and South-West)

Statistical analysis

Analysis to evaluate participants' distribution by the socio-demographic characteristics was done separately for the UAIS 2011/12 and UPHIA 2016/17 surveys. Descriptive analysis was also carried out to assess change in HIV prevalence between the two survey time points.

Multivariable-pooled logistic regression was used to assess factors associated with HIV positivity in the pooled dataset. Multivariable logistic decomposition models were used to determine the changes in HIV prevalence and the associated factors across the survey years. The Blinder-Oaxaca decomposition method[11,21] was used to display the contribution of the characteristics or their coefficients to the change in HIV prevalence. The method partitions the difference in HIV prevalence between the two periods into changes in effects and changes in composition and the interaction between them. All analyses were weighted to account for complex nature of the survey design. Analyses were carried out using STATA V.15.1[22].

Results

Socio-demographic characteristics

Table 1 presents the social demographic composition of the survey participants for 2011/12 and 2016/17 surveys. It shows that, for both surveys, individuals aged 15-19 years, females, those with primary level of education, marrieds and individuals who had one sexual partner in the 12 months preceding the survey outnumbered the other population subgroups. It also shows that between 2011/12 and 2016/17, the proportion of individuals aged 15-19 and 20-24 years participating in the surveys increased while the proportion of the other age categories (25-29years, 30-34years, 35-39years and 40-49 years) decreased. The proportion of other population categories participating in the survey that increased between the 2011/12 and 2016/17 survey include males; individuals with secondary and those with tertiary+ level of education; never married and those who were divorced or separated; individuals with one and those with two or more sexual partners; and those resident in urban areas (Table 1).

Of note, 10.2% had no formal and 6.1% had tertiary level of education in 2011/12 while only 6.0% had no formal education and 11.4% had tertiary level of education in 2016/17 (Table 1). Similarly, 25% had no partner and 10.0% had two or more sexual partners in 2011/12, while in 2016/17, 12.6% and 16.2% of the survey participants had zero and two or more sexual partners respectively (Table 1).

Table 1: Background characteristics (weighted) of study participants

Characteristic	Survey period (Weighted percentages)	
Age (Years)	**2011/12 (n=19,475)**	**2016/17 (n=25,420)**
15-19	22.8	25.9
20-24	17.9	21.0
25-29	16.6	16.3
30-34	13.2	12.8
35-39	12.5	9.9
40-49	17.0	14.0
Gender		
Male	44.4	47.5
Female	55.6	52.5
Education level		
None	10.2	6.0
Primary	58.2	55.4
Secondary	25.5	27.2
Tertiary+	6.1	11.4
Marital Status		
Never Married	29.6	35.7
Married	60.7	52.4
Widowed	2.3	1.7
Divorced/Separated	7.4	10.3
Number of sexual partners		
0	25.6	12.6
1	64.4	71.2
2+	10.0	16.2
Residence		
Rural	79.5	70.3
Urban	20.5	29.7
Region		
Central 1	11.2	14.0
Central 2	10.3	12.0
Kampala	7.8	6.6
East Central	10.4	9.8
Eastern	10.5	9.5
North East	8.1	7.4
West Nile	6.3	7.2
North Central	10.3	9.9
South West	12.0	11.0
Western	13.2	12.6

Trends in HIV prevalence between the two time points (2011/12 and 2016/17)

Figure 3 presents the trends in HIV prevalence for selected socio-demographic characteristics. Trends for all socio-demographic characteristics are presented in Appendix 1. Overall, HIV prevalence dropped from 7.3% in 2011/12 to 6.0% in 2016/17 among individuals aged 15-49yeras. Population subgroups with higher (more than average) reduction in HIV prevalence between 2011/12 and 2016/17 include males (6.1% to 4.3%); age group 20-24 years (5.3% to 3.3%); age group 30-34yeras(10.2% to 8.6%); primary level of education (8.0% to 6.6%); having two or more sexual partners (9.2% to 6.5%); and individuals who were divorced/separated (16.9% to 13.4%) (Figure 1 & Appendix 1). More than average reduction in HIV prevalence was also observed in Central 1 region (10.6% to 7.6%); Central 2 (9.0% to 7.4%); West Nile (4.8% to 2.8%) and in Western region (8.3% to 5.3%) (Appendix 1).

Between 2011/12 and 2016/17 surveys, HIV prevalence increased among individuals aged 40-49years (10.7% to 12.3%); those with no education (9.2% to 9.5%); those with no sexual partners (6.3% to 6.6%) and in Eastern region of the country (4.1% to 4.8%) (Figure 1 and Appendix 1).

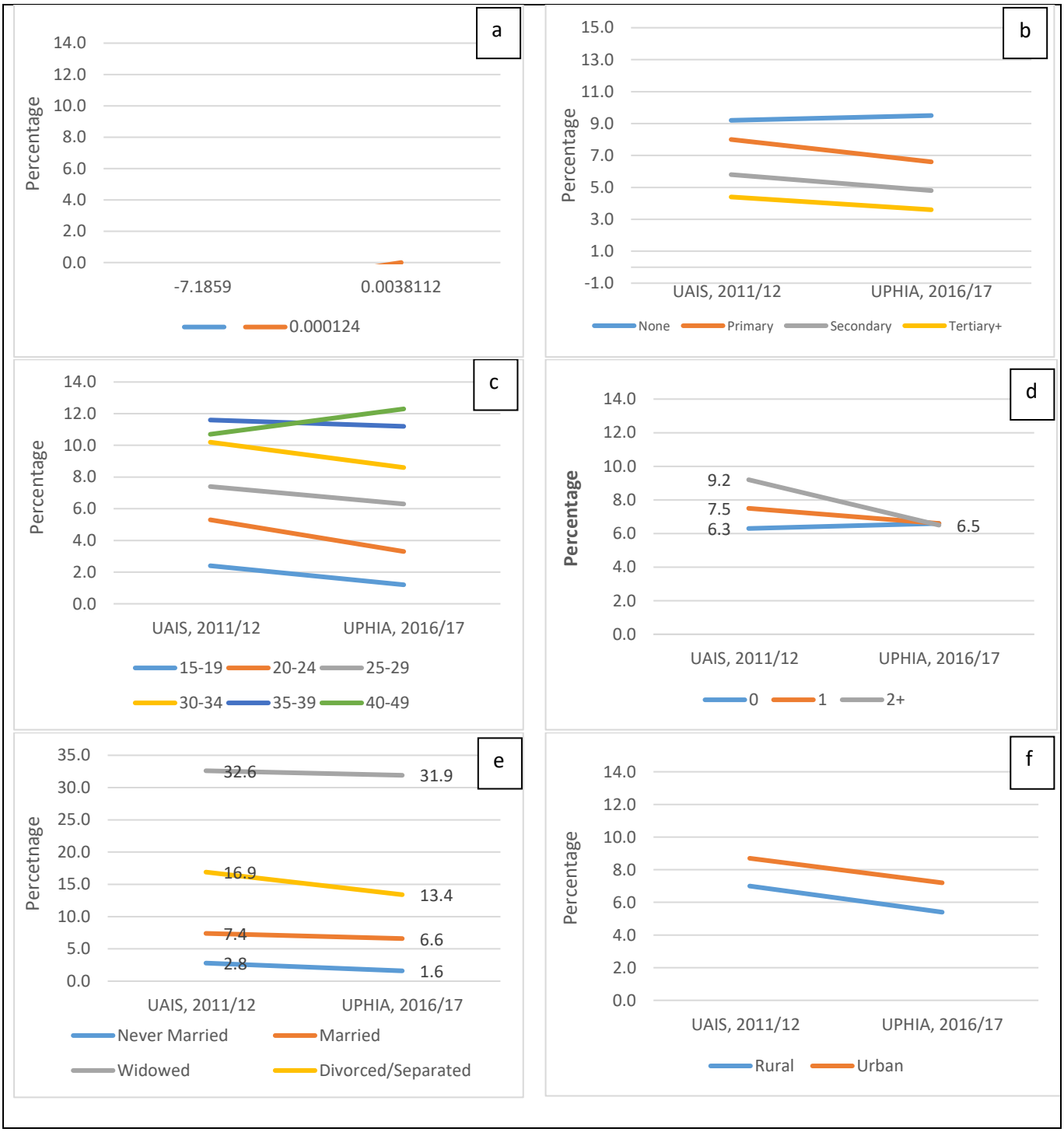


Figure 2: Trends in HIV prevalence between 2011/12 and 2016/17

Notes: Trends in HIV prevalence by: a) gender; b) highest level of education attained; c) age group; d) number of sexual partners in the 12 months preceding the survey; e) Marital status and; f) Area of residence.

Factors associated with HIV prevalence among study participants

We applied a pooled multivariable logistic regression to ascertain the factors associated with HIV positivity across the two survey periods; the results are presented in Table 2. It shows that there was a significant 16% drop in HIV prevalence in 2016/17 compared to 2011/12 and (p value< 0.001). Other factors that were significantly associated with change in HIV positivity include age, gender, marital status, education level, area of residence and number of sexual partners in the 12 months preceding the survey. We present the coefficients and standard errors of the factors associated with HIV positivity in appendix II.

HIV positivity was significantly lower among individuals aged 15-19 years compared those aged 20-24years (aOR: 0.54, 95% CI: 0.44, 0.68); among those with secondary (aOR: 0.85, 95% CI: 0.74, 0.99) and tertiary (aOR: 0.51, 95% CI: 0.41, 0.62) level of education compared to those with primary level of education and; among those never married compared to those who were marred at the time of the survey (aOR 0.82, 95% CI: 0.69, 0.97) (Table 2).

HIV positivity was significantly higher among females compared to males (aOR= 1.37; 95% CI: 1.20, 1.55); among older age groups compared to those aged 20-24 years (p value<0.001); among individuals who were widowed (aOR: 4.62, 95% CI: 3.84, 5.56) or separated/divorced (aOR: 2.27, 95% CI: 1.99, 2.59) compared to those who were married and; among those who had two or more sexual partners compared to those who had only one (aOR: 1.29, 95% CI: 1.14, 1.46) as shown in Table 2. Individuals from urban areas were also more likely to test HIV positive compared to those from rural areas (aOR: 1.58, 95% CI: 1.28, 1.94).

Table 2: Pooled logistic regression for factors associated with HIV positivity

Characteristic	HIV prevalence (weighted)	
	aOR (95%CI)	
Survey year		
2011/12	1.00	
2016/17	0.84	(0.76, 0.92)***
Age (Years)		
15-19	0.54	(0.44, 0.68)***
20-24	1.00	
25-29	1.54	(1.30, 1.83)***
30-34	2.06	(1.72, 2.46)***
35-39	2.41	(2.01, 2.88)***
40-49	2.34	(1.98, 2.75)***
Gender		
Male	1.00	
Female	1.37	(1.20, 1.55)***
Education level		
None	0.85	(0.72, 0.99)
Primary	1.00	
Secondary	0.85	(0.74, 0.99)*
Tertiary+	0.51	(0.41, 0.62)***
Marital Status		
Never Married	0.82	(0.69, 0.97)*
Married	1.00	
Widowed	4.62	(3.84, 5.56)***
Divorced/Separated	2.27	(1.99, 2.59)***
Number of sexual partners		
0	0.96	(0.85, 1.08)
1	1.00	
2+	1.29	(1.14, 1.46)***
Residence		
Rural	1.00	
Urban	1.58	(1.28, 1.94)***
Region		
Central 1	1.00	
Central 2	0.89	(0.66, 1.22)
Kampala	0.68	(0.46, 1.00)
East Central	0.56	(0.36, 0.86)**
Eastern	0.53	(0.37, 0.75)***
North East	0.54	(0.37, 0.79)**
West Nile	0.39	(0.28, 0.55)***
North Central	0.92	(0.67, 1.26)
South West	0.90	(0.64, 1.28)
Western	0.79	(0.48, 1.28)

* = significant at 0.05; ** = significant at 0.01; *** = Significant at <0.001

Assessing components of change in HIV prevalence between 2011/12 and 2016/17

We applied logistic decomposition analysis to generate the contribution of the factors and their coefficients to change in HIV positivity between 2011/12 and 2016/17 (Table 3). The results show that 59.2% of the overall percentage change in HIV positivity was attributable to changes in the effects of the age group 40-49years (p value<0.001). The analysis results further show that, change in both the characteristics (endowments) and the effects of the characteristics of the survey participant's between 2011/12 and 2016/17 resulted in a reduction in HIV prevalence. The contribution of both participants' characteristics (endowments) and the coefficients were statistically significantly (p value< 0.001), (that is the logistic multivariable decomposition model was statistically significant). Change in the effects of the characteristics contributed more to the reduction in HIV prevalence between the two time points compared to the change in the composition of the survey participants.

We also note that although the changes in the effects/coefficients contributed to a larger reduction in HIV prevalence, the contribution of the effects of most of the variables was not statistically significant. The study results also show that more than half (59.2%) of the decrease in HIV prevalence was attributable to the coefficient of age group 40-49 years (P value: <0.001). In terms of coefficients, aged groups 35-39 years (29.4%) and 40-49 years (59.2%) were the variables that were significantly associated with change in HIV positivity between 2011/12 and 2016/17.

.

Table 3. Decomposition of changes in HIV prevalence for the period 2011/12 to 2016/17

Characteristic	Due to differences in Characteristics (E)		Due to differences in Coefficients (C)		Total amount attributable
	Coefficient	Percent	Coefficient	Percent	
Age (Years)					
15-19	0.00216***	-34.296	-0.000792	12.572	-21.724
20-24	0.00		0.00		
25-29	0.000459***	-7.2861	0.0011549	-18.315	-25.601
30-34	0.000655***	-10.393	0.0010974	-17.402	-27.795
35-39	-0.000379***	6.0247	0.0018557*	-29.428	-23.403
40-49	0.000145***	-2.3027	0.0037359**	-59.245	-61.548
Gender					
Male	0.00		0.00		
Female	0.000453***	-7.1859	0.0038112	-60.438	-67.624
Education level					
None	0.000124	-1.9685	0.00038862	-6.1628	-8.131
Primary	0.00		0.00		
Secondary	0.0000563	-0.89271	0.0011594	-18.386	-19.279
Tertiary+	-0.0014889***	23.611	0.00022552	-3.5763	20.035
Marital Status					
Never Married	0.000403	-6.3837	0.0014417	-22.862	-29.246
Married	0.00		0.00		
Widowed	-0.000132***	2.1033	-0.0000793	1.2586	3.362
Divorced/Separated	0.001613***	-25.592	-0.00013664	2.1669	-23.425
# of sexual partners					
0	0.000532	-8.4421	-0.0012578	19.947	11.505
1	0.00		0.00		
2+	0.000465*	-7.3845	-.00058452	9.2694	1.885
Residence					
Rural	0.00		0.00		
Urban	0.001482***	-23.507	-.0010425	16.532	-6.975
Region					
Central 1	0.00		0.00		
Central 2	-1.1422e-06	0.01811	0.0008337	-13.222	-13.204
Kampala	0.0003067*	-4.8554	0.0012513	-19.844	-24.699
East Central	0.0000431	-0.6836	0.0008742	-13.864	-14.548
Eastern	-0.0009428***	14.952	0.0016567	-26.271	-11.319
North East	-0.0011876***	18.834	-0.0002359	3.7419	22.576
West Nile	-0.0015796***	25.049	0.0000610	-.96834	24.081
North Central	-0.0000336	0.53415	0.000918	-14.558	-14.024
South West	0.0000111	-.1762	0.00069642	-11.044	-11.220
Western	0.0002785*	-4.4175	-0.00021104	3.3467	-1.071
Subtotal	0.0034456***	-54.64	-0.009751***	-266.751	-321.392
Constant			-0.026573	421.39	
Total		-54.64	-0.009751***	154.64	100.000

* = significant at 0.05; ** = significant at 0.01; *** = significant at <0.001

Discussion

This study applied logistic decomposition analysis to assess the contribution of survey participants' characteristics (endowments) as well as the coefficients of the characteristics to change in HIV prevalence between 2011/12 and 2016/17 surveys. The study further applied pooled multivariable logistic regression to evaluate the factors associated with HIV prevalence.

Study findings show that overall HIV prevalence reduced from 7.3% in 2011/12 to 6.0% in 2016/17 among individuals aged 15-49 years tested during the survey representing a 18% decline in prevalence. Prevalence decreased in all population subgroups except among individuals with no formal education; individuals with no sexual partners in the 12 months preceding the survey and among individuals aged 40-49 years. The largest decrease was among individuals aged 20-24 years, individuals who were divorced/separated and among individuals who had two or more sexual partners in the 12 preceding the survey.

Generally, both the characteristics (endowments) of individuals surveyed and the effects of the characteristics had a significant contribution on the reduction in HIV prevalence between the two survey period. Changes in the population endowments (composition) that led to significantly larger change in HIV prevalence include young age (15-19 years), Tertiary level of education, being previously married and urban residence. Change in the characteristics ("endowments") of the survey participants is the change in the composition of survey participants in 2011/12 compared to the 2016/17. While effects of the differing coefficients of the characteristics refer to the differences of the coefficients of the characteristics in the two surveys on the overall change in HIV prevalence.

For effects of the coefficients, only differences in coefficients for older age group were likely to lead to a statistically significant change in HIV prevalence. Other characteristics with large coefficients effects (large differentials leading to large contribution to change in HIV prevalence) although may not lead to statistically significant change in HIV prevalence include female gender, secondary level of education, not being married, having no sexual partner in the 12 months preceding the survey and residence in urban areas.

A reduction in HIV prevalence have been observed in other studies conducted in Uganda over the years irrespective of the source of data. HIV prevalence decline was observed among women accessing ANC services[5,23] and among rural communities surveyed in western Uganda[6,24]. Key factors highlighted in these studies for the decline include change in sexual behavior such as

reduction in the number of sexual partners consistent with findings in our study [25]. In this study, prevalence decreased by 18% between 2011 and 2016, although the proportion of individuals with two or more sexual partners increased from 10% to 16%. Studies elsewhere have noted that individuals tend to adopt acceptable and desirable behaviors as they adjust to the realities of the epidemic[7,26]. We therefore believe that the decrease in HIV prevalence more than compensates for the increase in the proportion of individuals with two or more sexual partners and is attributable to adoption of safe sexual practices. Additionally, effective prevention programs such as male circumcision were rolled out in Uganda following studies that showed its effectiveness against HIV incidence[27] as well its acceptability as an HIV prevention tool[28,29]. The policy on male circumcision was rolled out in 2010, just before the 2011 survey in Uganda[30]. Another key intervention was the test and treat policy rolled out in 2013 that recommended initiation of ART for some population categories irrespective of their CD4 count once they tested HIV positive. This was based on the evidence of minimized chances of passing on the virus to un-infected sexual partner if the HIV positive individuals was taking ART optimally[31–37].

Increase in HIV prevalence on the other hand among individuals aged 40-49 years is attributable to cumulative effect, given the effectiveness of HIV treatment interventions that has enabled HIV infected individuals to leave longer with HIV infection[32,34,38]. The effect of this age group is thus likely to contribute more to change in prevalence, as both new incident cases as well as individuals infected earlier in their lives (cohort effect) are able to live longer with HIV. Our findings showed that change in coefficients for older age group (40-49 years) were likely to lead to a statistically significant change in HIV prevalence. The likelihood of testing HIV positive was also significantly associated with older age group in this is study.

HIV prevalence among young people is argued to provide a better reflection of the nature of HIV/AIDS epidemic in any population. It's therefore often used as an indication of the level of new infections as this population sub groups are newly exposed to risky behaviors such as initiation of sex and other vulnerabilities[15,39–43]. In our study, although the proportion of young people (15-25 years) participating in the survey increased from 40% in 2011 to 48%, HIV prevalence decreased significantly from 2.4% to 1.2% and 5.3% to 3.3% among individuals aged 15-19 years and 20-24years respectively. Additionally, the change in composition of young people was significantly associated with change in HIV prevalence. Individuals aged 15-19 years contributed to 37% reduction in HIV prevalence between 2011 and 2016.

The large reduction in HIV prevalence among young people and the significant contribution to overall change in HIV prevalence is attributable to both increased sensitivity to new information and the higher likelihood to adopt or change behavior in effect, as highlighted in other studies[4,44,45]. HIV programs such as DREAMs were rolled out in many countries in sub Saharan Africa including Uganda. The program focused on providing information to young people especially young girls on HIV prevention strategies such as delaying sexual initiation, youth friendly sexual and reproductive health care as well as condom promotion[10]. The participating countries realized between 20-40% reduction in new HIV infections[10].

Although results from our study shows that having two or more sexual partners and being previously married were significantly associated with higher positivity rates, there were statistically significant reduction in HIV prevalence among these population categories. These population categories also had statistically signification compositional effect on the reduction in HIV prevalence between 2011 and 2016. Multiple concurrent sexual relationships is a key driver of HIV distribution in settings where the epidemic is described as generalized[44,46,47]. Additionally, individuals who were previously married were found to exhibit higher prevalence due to a possible infection in an earlier relationship[44,47]. Previously married individuals have also been reported to have had multiple sexual relationship compared to those who were married in a qualitative study in western Uganda[48].

Other factors that had a large and significant composition effect on HIV prevalence include female gender, tertiary level of education and residence in urban areas. Most studies have shown HIV prevalence to be higher among females compared to males[9,18,47,49,50]. Our study show that the proportion of females participating in the surveys in 2016 reduced. Due to high levels of vulnerabilities including poverty and lack of independence among women in Uganda, interventions that focus on this differential are more likely to realize an overall impact in reducing HIV prevalence levels.

Individuals with higher education levels have easy access to information and are thus more likely to adjust their sexual behaviors and other HIV risky behavior [4,51]. Our study findings show that individuals with tertiary or higher levels of education have signification large contribution to the reduction in HIV prevalence. Additionally, the proportion of individuals with tertiary or higher levels of education increased between the two survey time points while HIV prevalence reduced significantly.

Strengths and limitations

We applied multilevel logistic regression analysis which has multiple advantages over classical models including adjusting for cluster level correlation through random effects component. Multilevel models also allow for inclusion of predictors at both individual and cluster levels. We used data from two time points only (i.e 2011 UAIS and the Population HIV Impact Assessment (PHIA), conducted in 2017. More data points would enable validation of the results obtained. Data from the population survey conducted in 2004/05 was not readily accessible while the 2019-2020 survey was interrupted by the COVID19 pandemic. Population survey data is also characterised by non-response especially for sensitive questions like condom use and multiple sexual partners.

Conclusion

Composition of the population surveyed as well as the effects of their characteristics/coefficients on HIV prevalence bring to the fore salient factors that underscore HIV spread. Despite extensive efforts to curb the spread of the virus, some population groups continue to experience higher incidences. Targeting these population subgroup based on their specific and possible contribution to HIV prevalence reduction and hence incidence will heighten the pace at which zero new infections occur in the general population. This study maybe among the first to stratify effects into endowments and those arising from the coefficients of the endowments. This is critical for identifying intervention areas to focus on to end the HIV epidemic sooner. In particular, focusing on young age groups, individuals with higher level of education, those who were previously married and those who have two or more sexual partners will have large contribution to change in HIV prevalence. Additionally, the change in coefficient of older age group has a large contribution on HIV prevalence reduction.

Ethics

Ethical clearance for this study was obtained from the University of Witwatersrand Human Research Ethics Committee (HREC) and Uganda National Council of Science and Technology (UNCST). Permission was also obtained from Ministry of Health, Uganda, and the Measure DHS (ORC MACRO). All individual-level data was identified using unique person identifiers only.

Competing interests

There are no conflicts of interest

Authors' contributions

All authors have read and approved the final manuscript. JO, JJV and JL designed the research. JO compiled and analysed the data and wrote the first manuscript draft. All authors provided critical content analysis, reviewed and approved the final manuscript.

Acknowledgements

We acknowledge the Ugandan Ministry of Health and its partners in conducting the Uganda AIDS Indicator Survey, and for the permission to use the dataset

Funding

This study was supported through the DELTAS Africa Initiative SSACAB (Grant No. 107754/Z/15/Z). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant No. 107754/Z/15/Z) and the UK government. The views expressed in this publication are those of the authors and not necessarily those of the AAS, NEPAD Agency, Wellcome Trust or the UK government.

References

1. UNAIDS. Global HIV & AIDS statistics — 2020 fact sheet _ UNAIDS. 2020.
2. (No Title) [Internet]. [cited 2021 Feb 9]. Available from: https://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf
3. 2025 AIDS Targets: 2025 Target setting and 2020-2030 resource needs and impact estimation | UNAIDS [Internet]. [cited 2021 Feb 9]. Available from: https://www.unaids.org/en/topics/2025_target_setting
4. Kilian AHD, Gregson S, Ndyanabangi B, Walusaga K, Kipp W, Sahlmüller G, et al. Reductions in risk behaviour provide the most consistent explanation for declining HIV-1 prevalence in Uganda. Aids. 1999;13(3):391–8.
5. Asiimwe-Okiror G, Opio AA, Musinguzi J, Madraa E, Tembo G, Caraël M. Change in sexual behaviour and decline in HIV infection among young pregnant women in urban Uganda. Aids. 1997;11(14):1757–63.
6. Kamali A, Carpenter LM, Whitworth JAG, Pool R, Ruberantwari A, Ojwiya A. Seven-year trends in HIV-1 infection rates, and changes in sexual behaviour, among adults in rural Uganda. Aids. 2000;14(4):427–34.
7. Fylkesnes K, Musonda RM, Sichone M, Ndhlovu Z, Tembo F, Monze M. Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population-based surveys. AIDS. 2001;15(7):907–16.
8. Hallett TB, Aberle-Grasse J, Bello G, Boulos LM, Cayemittes MPA, Cheluget B, et al. Declines in HIV prevalence can be associated with changing sexual behaviour in Uganda, urban Kenya, Zimbabwe, and urban Haiti. Sex Transm Infect. 2006;82(SUPPL. 1):1–8.
9. Ministry of Health Uganda. Uganda Population-based HIV Impact Assessment (UPHIA) 2016-2017: Final Report. [Internet]. Kampala, Uganda; 2019. Available from: https://phia.icap.columbia.edu/wp-content/uploads/2019/07/UPHIA_Final_Report_Revise_07.11.2019_Final_for-web.pdf
10. DREAMS: Partnership to Reduce HIV/AIDS in Adolescent Girls and Young Women | U.S. Agency for International Development [Internet]. [cited 2021 Feb 13]. Available from: https://www.usaid.gov/global-health/health-areas/hiv-and-aids/technical-areas/dreams
11. Blinder AS. Wage Discrimination : Reduced Form and Structural Estimates. J Hum Resour.

1973;8(4):436–55.

12. Yussuf MH, Elewonibi BR, Rwabilimbo MM, Mboya IB, Mahande1 MJ. Trends and predictors of changes in modern contraceptive use among women aged 15-49 years in Tanzania from 2004-2016: Evidence from Tanzania Demographic and Health Surveys. PLoS One [Internet]. 2020;15(6 June):1–14. Available from: http://dx.doi.org/10.1371/journal.pone.0234980

13. Ariho P, Nzabona A. Determinants of Change in Fertility among Women in Rural Areas of Uganda. 2019; Available from: https://doi.org/10.1155/2019/6429171

14. Hong R, Hor D. Factors associated with the decline of under-five mortality in Cambodia, 2000-2010: Further analysis of the Cambodia Demographic and Health Surveys . DHS Furth Anal Reports No 84 [Internet]. 2013;2000–10. Available from: http://dhsprogram.com/pubs/pdf/FA84/FA84.pdf

15. Phimemon RN, Mahande MJ, Ramadhani HO. Factors Associated with Changes in Uptake of HIV Testing among Young Women (age 15-24) in Tanzania from 2003 and 2012. DHS Working Papers No. 119. Rockville, Maryland, USA: ICF International. 2015.

16. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Monitoring HIV impact using Population-Based Surveys. 2015.

17. Uganda Ministry of Health. Uganda HIV/AIDS Sero-Behavioural Survey 2004-05. 2004;36; 107. Available from: http://pdf.usaid.gov/pdf_docs/Pnadg508.pdf

18. Ministry of Heath and ICF international. Uganda AIDS Indicator Survey (AIS) 2011 [Internet]. Kampala Uganda and Rockville, Maryland, USA; 2012. Available from: http://health.go.ug/docs/UAIS_2011_REPORT.pdf

19. Uganda Bureau of Statistics. 2002 Uganda Population and Housing Census Administrative Report. 2007.

20. UBOS. National housing and population census 2014-Area Specific Profiles, Wakiso District. 2017.

21. Powers DA, Yoshioka H, Yun M-S. Mvdcmp: Multivariate Decomposition for Nonlinear Response Models. Stata J Promot Commun Stat Stata. 2011;11(4):556–76.

22. StataCorp. Stata Statistical Software: Release 15 [Internet]. College Station, TX: StataCorp LLC; 2017. Available from: https://www.stata.com/

23. Kirungi WL, Musinguzi J, Madraa E, Mulumba N, Callejja T, Ghys P. Trends in antenatal HIV prevalence in urban Uganda associated with uptake of preventive sexual behaviour. Sex Transm Infect [Internet]. 2006;82(suppl 1 (i36-i41)). Available from: www.stijournal.com

24. Shafer LA, Biraro S, Nakiyingi-Miiro J, Kamali A, Ssematimba D, Ouma J, et al. HIV prevalence and incidence are no longer falling in southwest Uganda: Evidence from a rural population cohort 1989-2005. Aids. 2008;22(13):1641–9.

25. Mafigiri R, Matovu JKBB, Makumbi FE, Ndyanabo A, Nabukalu D, Sakor M, et al. HIV prevalence and uptake of HIV/AIDS services among youths (15-24 Years) in fishing and neighboring communities of Kasensero, Rakai District, South Western Uganda. BMC Public Health [Internet]. 2017 Dec 14 [cited 2019 Apr 2];17(1):251. Available from: http://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-017-4166-2

26. Schaefer R, Thomas R, Maswera R, Kadzura N, Nyamukapa C, Gregson S. Relationships between changes in HIV risk perception and condom use in East Zimbabwe 2003-2013: Population-based longitudinal analyses. BMC Public Health. 2020;20(1).

27. Gray R, Kigozi G, Kong X, Ssempiija V, Makumbi F, Wattya S, et al. The effectiveness of male circumcision for HIV prevention and effects on risk behaviors in a posttrial follow-up study. Aids. 2012;26(5):609–15.

28. Odoch WD, Kabali K, Ankunda R, Zulu JM, Tetui M. Introduction of male circumcision for HIV prevention in Uganda: analysis of the policy process. Heal Res Policy Syst [Internet]. 2015 Jun 20 [cited 2021 Feb 13];13(1):31. Available from: https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-015-0020-0

29. Albert LM, Akol A, L'Engle K, Tolley EE, Ramirez CB, Opio A, et al. Acceptability of male circumcision for prevention of HIV infection among men and women in Uganda. AIDS Care - Psychol Socio-Medical Asp AIDS/HIV. 2011 Dec;23(12):1578–85.

30. Safe Male Circumcision Policy | Ministry of Health Knowledge Management Portal [Internet]. [cited 2021 Feb 13]. Available from: http://library.health.go.ug/publications/policy-documents/safe-male-circumcision-policy

31. Keith A. SEARCH study exceeds 90-90-90 targets after 2 years of "test and treat" for HIV in rural East Africa | aidsmap [Internet]. 2013 [cited 2021 Jan 23]. Available from: https://www.aidsmap.com/news/jul-2016/search-study-exceeds-90-90-90-targets-after-2-years-test-and-treat-hiv-rural-east

32. Johnson LF, May MT, Dorrington RE, Cornell M, Boulle A, Egger M, et al. Estimating the impact of antiretroviral treatment on adult mortality trends in South Africa: A mathematical modelling study. PLoS Med. 2017;14(12):1–17.

33. Livia, Montana; Melissa, Neuman; Vnod M. Spatial modelling of HIV prevalence in Kenya. Vol. 0. 2007.

34. Frank TD, Carter A, Jahagirdar D, Biehl MH, Douwes-Schultz D, Larson SL, et al. Global, regional, and national incidence, prevalence, and mortality of HIV, 1980-2017, and forecasts to 2030, for 195 countries and territories: A systematic analysis for the Global Burden of Diseases, Injuries, and Risk Factors Study 2017. Lancet HIV. 2019;6(12):e831–59.

35. Das M, Chu PL, Santos G-M, Scheer S, Vittinghoff E, Mcfarland W, et al. Decreases in Community Viral Load Are Accompanied by Reductions in New HIV Infections in San Francisco. [cited 2021 Feb 13]; Available from: www.plosone.org

36. Abongomera G, Kiwuwa-Muyingo S, Revill P, Chiwaula L, Mabugu T, Phillips AN, et al. Impact of decentralisation of antiretroviral therapy services on HIV testing and care at a population level in agago district in rural Northern Uganda: Results from the Lablite population surveys. Int Health. 2017;9(2):91–9.

37. Johnson LF, Davies MA, Moultrie H, Sherman GG, Bland RM, Rehle TM, et al. The effect of early initiation of antiretroviral treatment in infants on pediatric AIDS mortality in South Africa: A model-based analysis. Pediatr Infect Dis J. 2012 May;31(5):474–80.

38. Yotebieng M, Brazier E, Addison D, Kimmel AD, Cornell M, Keiser O, et al. Research priorities to inform “Treat All” policy implementation for people living with HIV in sub-Saharan Africa: a consensus statement from the International epidemiology Databases to Evaluate AIDS (Ie DEA). J Int AIDS Soc. 2019;22(1):e25218.

39. Rodríguez MP, Hayes PR. Reducing HIV prevalence among young people : A review of the UNGASS prevalence goal and how it should be monitored Discussion paper commissioned by WHO October 2002 TABLE OF CONTENTS. Trop Med. 2002;(October).

40. Marsh KA, Nyamukapa CA, Donnelly CA, Garcia-Calleja JM, Mushati P, Garnett GP, et al. Monitoring trends in HIV prevalence among young people, aged 15 to 24 years, in Manicaland, Zimbabwe. J Int AIDS Soc [Internet]. 2011;14(1):27. Available from: http://www.jiasociety.org/content/14/1/27

41. Ortblad KF, Musoke DK, Ngabirano T, Salomon JA, Haberer JE, McConnell M, et al. Is knowledge of HIV status associated with sexual behaviours? A fixed effects analysis of a female sex worker cohort in urban Uganda. J Int AIDS Soc. 2019;22(7):1–11.

42. Igulot P, A Magadi M. Social and structural vulnerability to HIV infection in Uganda: A multilevel modelling of AIDS indicators survey data, 2004-2005 and 2011. J Community Med. 2018;1(2):2004–5.

43. Marsh K, Mahy M, Salomon JA, Hogan DR. Assessing and adjusting for differences between HIV prevalence estimates derived from national population-based surveys and antenatal care surveillance, with applications for Spectrum 2013. Aids [Internet]. 2014;28(August):S497–505. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201411004-00011

44. Zaba BW, Carpenter LM, Boerma JT, Gregson S, Nakiyingi J, Urassa M. Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa. AIDS. 2000;14(17):2741–50.

45. Gregson S, Terceiria N, Kakowa M, Mason PR, Anderson RM, Chandiwana SK CM. Study of bias in antenatal clinic HIV-1 surveillance data in a high contraceptive prevalence population in sub-Saharan Africa. AIDS. 2002;16(4):643–52.

46. Amornkul PN, Vandenhoudt H, Nasokho P, Odhiambo F, Mwaengo D, Hightower A, et al. HIV prevalence and associated risk factors among individuals aged 13-34 years in rural Western Kenya. PLoS One. 2009;4(7).

47. Kwesigabo G, Killewo JZ, Urassa W, Mbena E, Mhalu F, Lugalla JL, et al. Monitoring of HIV-1 infection prevalence and trends in the general population using pregnant women as a sentinel population: 9 years experience from the Kagera region of Tanzania. J Acquir Immune Defic Syndr. 2000;23(5):410–7.

48. Nalugoda F, Guwatudde D, Bwaninka JB, Makumbi FE, Lutalo T, Kagaayi J, et al. Marriage and the risk of incident HIV infection in Rakai, Uganda. J Acquir Immune Defic Syndr [Internet]. 2014 Jan 1;65(1):91–8. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24419066

49. Kapesa A, Basinda N, Nyanza EC, Mushi MF, Jahanpour O, Ngallaba SE. Prevalence of HIV infection and uptake of HIV/AIDS services among fisherfolk in landing Islands of Lake Victoria, north western Tanzania. BMC Health Serv Res [Internet]. 2018 Dec 18 [cited 2019 Apr

2];18(1):980. Available from: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-018-3784-4

50. Hegdahl HK, Fylkesnes KM, Sandøy IF. Sex differences in HIV prevalence persist over time: Evidence from 18 countries in Sub-Saharan Africa. PLoS One [Internet]. 2016;11(2):1–17. Available from: http://dx.doi.org/10.1371/journal.pone.0148502

51. Nangendo J, Katahoire AR, Armstrong-Hough M, Kabami J, Obeng-Amoako GO, Muwema M, et al. Prevalence, associated factors and perspectives of HIV testing among men in Uganda. PLoS One [Internet]. 2020;15(8 August):1–20. Available from: http://dx.doi.org/10.1371/journal.pone.0237402

Appendix 1: Trends in Weighted HIV prevalence between 2011 and 2016

Characteristic	HIV prevalence (weighted)		Percentage point difference in HIV prevalence
	UAIS, 2011/12	UPHIA, 2016/17	
Over all	**7.3**	**6.0**	**-1.3**
Age (Years)			
15-19	2.4	1.2	-1.2
20-24	5.3	3.3	-2.0
25-29	7.4	6.3	-1.1
30-34	10.2	8.6	-1.6
35-39	11.6	11.2	-0.4
40-49	10.7	12.3	1.6
Gender			
Male	6.1	4.3	-1.8
Female	8.3	7.5	-0.8
Education level			
None	9.2	9.5	0.3
Primary	8.0	6.6	-1.4
Secondary	5.8	4.8	-1.0
Tertiary+	4.4	3.6	-0.8
Marital Status			
Never Married	2.8	1.6	-1.2
Married	7.4	6.6	-0.8
Widowed	32.6	31.9	-0.7
Divorced/Separated	16.9	13.4	-3.5
Number of sexual partners			
0	6.3	6.6	0.3
1	7.5	6.6	-0.9
2+	9.2	6.5	-2.7
Residence			
Rural	7.0	5.4	-1.6
Urban	8.7	7.2	-1.5
Region			
Central 1	10.6	7.6	-3.0
Central 2	9.0	7.4	-1.6
Kampala	7.1	6.6	-0.5
East Central	5.7	4.4	-1.3
Eastern	4.1	4.8	0.7
North East	5.3	3.4	-1.9
West Nile	4.8	2.8	-2.0
North Central	8.3	7.0	-1.3
South West	8.0	7.7	-0.3
Western	8.3	5.5	-2.8

Appendix II: Factors associated with HIV positivity

Characteristic	Coefficient	SE	Coefficient	SE
Age (Years)				
15-19	-0.598	0.141	-0.688	0.164
20-24	0.00		0.00	
25-29	0.338	0.100	0.518	-0.115
30-34	0.605	0.110	0.819	0.117
35-39	0.677	0.107	1.058	0.120
40-49	0.550	0.103	1.107	0.117
Gender				
Male	0.00		0.00	
Female	0.220	0.0725	0.394	0.082
Education level				
None	-0.197	0.107	-0.101	0.116
Primary	0.00		0.00	
Secondary	-0.224	0.081	-0.106	0.087
Tertiary+	-0.731	0.146	-0.637	0.141
Marital Status				
Never Married	-0.278	0.157	-0.153	0.130
Married	0.00		0.00	
Widowed	1.580	0.138	1.496	0.121
Divorced/Separated	0.845	0.110	0.799	0.076
# of sexual partners				
0	0.015	0.085	-0.110	0.085
1	0.00		0.00	
2+	0.348	0.096	0.192	0.083
Residence				
Rural	0.00		0.00	
Urban	0.532	0.145	0.402	0.110
Region				
Central 1	0.00		0.00	
Central 2	-0.216	0.110	0.001	0.239
Kampala	-0.553	0.164	-0.250	0.254
East Central	-0.699	0.132	-0.487	0.339
Eastern	-0.873	0.167	-0.481	0.247
North East	-0.573	0.161	-0.644	0.290
West Nile	-0.939	0.147	-0.923	0.236
North Central	-0.199	0.146	-0.034	0.250
South West	-0.202	0.156	-0.101	0.251
Western	-0.226	0.197	-0.275	0.329
Constant	-2.757	0.110	-3.439	0.244
R2	0.0849			
s.e				