

## Collecting data on HIV-related mortality during household surveys: a randomized validation study in Malawi

MEIRU – LSHTM – JHU Partnership for improving adult mortality data<sup>1,2</sup>

---

<sup>1</sup> Members of the group include (in alphabetical order):

- Abena Amoah, London School of Hygiene and Tropical Medicine / Malawi Epidemiological Intervention Research Unit, Chilumba (Malawi)
- Sarah Brumfield, Johns Hopkins University, Baltimore (USA)
- Amelia C. Crampin, London School of Hygiene and Tropical Medicine / Malawi Epidemiological Intervention Research Unit, Lilongwe (Malawi)
- Albert N. Dube, Malawi Epidemiological Intervention Research Unit, Chilumba (Malawi)
- Stéphane Helleringer, Johns Hopkins University, Baltimore (USA)
- Estelle McLean, London School of Hygiene and Tropical Medicine, London (UK)
- Georges Reniers, London School of Hygiene and Tropical Medicine, London (UK)

<sup>2</sup> Correspondence to Stéphane Helleringer,  
Bloomberg School of Public Health  
Johns Hopkins University  
615 n. Wolfe Street  
Baltimore, MD 21205

USA  
NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.  
Email: [sheller7@jhu.edu](mailto:sheller7@jhu.edu)

## ABSTRACT

**Background:** In many African countries, there are limited representative data on HIV/AIDS mortality. We tested whether such data could be collected during household surveys periodically conducted in most African countries.

**Methods:** We added HIV questions to the module on adult and maternal mortality used in Demographic and Health Surveys. We conducted a validation study of the data generated by these questions in northern Malawi. We randomly assigned men and women aged 15–59 years old to a face-to-face interview or audio computer-assisted self-interviewing (ACASI). We compared survey reports of adult deaths to prospective data on mortality and HIV collected by the Karonga Health and Demographic Surveillance Site. We calculated the sensitivity and specificity of survey data in recording the HIV status of deceased siblings of respondents. We adjusted for partial verification bias using multiple imputations.

**Results:** We interviewed 535 participants, who reported 885 deaths at ages 15 and older. The added HIV questions yielded largely complete data on the HIV status of respondents' deceased siblings, particularly those who died recently. The adjusted sensitivity of survey data on HIV status of the deceased was high in both study groups (0.78–0.82). There were few false positive reports of the HIV status of deceased siblings (specificity = 0.96–0.98). ACASI did not improve the accuracy of survey data, but it required more time to collect mortality reports. Asking the HIV questions only took 0.4 minute ( $\approx 25$  seconds) per deceased sibling in face-to-face interviews.

**Conclusions:** Adding HIV questions to mortality questionnaires used in household surveys yields accurate data on the HIV status of deceased adults. These new data could aid in tracking progress towards global HIV elimination targets.

## INTRODUCTION

Reaching “zero AIDS-related deaths” by 2030 is a key goal of HIV prevention and treatment programs worldwide.<sup>1</sup> In most African countries, the data required to track progress towards this objective are limited.<sup>2</sup> Reliable data on HIV/AIDS mortality are only available for select populations. Clinical cohorts<sup>3,4</sup> measure mortality among patients on antiretroviral therapy (ART), but they rarely document the survival of persons with HIV (PHIV) who are not on ART and they exclude HIV-negative persons. Health and demographic surveillance systems (HDSS) collect population-based data on causes of death,<sup>5,6</sup> but they often cover small, rural populations.

In the absence of representative mortality data, national estimates of HIV/AIDS deaths in most African countries are inferred primarily from data on HIV prevalence and ART coverage.<sup>7,8</sup> Several statistical and epidemiological models exist for this purpose.<sup>8,9</sup> They can yield discrepant estimates of the number of lives saved by HIV programs.<sup>10</sup> To better calibrate the models tracking progress towards zero AIDS-related deaths in African countries, we tested whether data on HIV/AIDS mortality could be collected during household surveys.

Household surveys such as the demographic and health surveys (DHS) or the multiple indicator cluster surveys (MICS) are conducted periodically in most African countries,<sup>11,12</sup> where they constitute the main source of representative data on all-cause mortality.<sup>13</sup> They ask respondents to report the survival of close relatives.<sup>14</sup> They include questions about pregnancy-related deaths<sup>14</sup> and injuries<sup>15</sup>, but they do not inquire about HIV/AIDS mortality because of concerns that respondents may not know whether their relatives died of HIV-related causes. These concerns might now have receded, due to a) the scale-up of HIV testing,<sup>16</sup> b) increased disclosure of HIV status to family members<sup>17,18</sup> and/or c) conversations in social networks to determine if a

death can be attributed to HIV/AIDS.<sup>19</sup> Survey respondents may also be more willing to disclose such information following declines in HIV stigma.<sup>20,21</sup>

We inserted questions about the HIV status of a respondent's deceased relatives into the standard DHS module on adult/maternal mortality. These HIV questions do not allow classifying causes of deaths using codes from the International Classification of Diseases. However, by comparing HIV prevalence among survivors and deceased relatives, they allow measuring excess mortality associated with HIV (appendix A1). We conducted a validation study of the data generated by these questions in Northern Malawi.

## **Data and Methods**

### *Data sources*

To ascertain the HIV status of deceased relatives, we adapted existing questions about a respondent's own HIV status.<sup>22</sup> We asked respondents if their relative had ever been tested, and if so what were the results of their most recent test. To limit missing data, we asked respondents to assess the likelihood that their relative was infected with HIV. Response categories ranged from "highly likely" to "highly unlikely". This question was not applicable if the respondent's relative had received positive test results.

We included the HIV questions in siblings' survival histories, i.e., a questionnaire wherein respondents list all their maternal siblings, state whether they are alive, and report the current age of live siblings, or the age at death and time since death of deceased siblings.<sup>23</sup> We asked HIV questions for all deaths reported at ages 15 years and older. We classified deceased siblings as

PHIV when respondents stated that the deceased had received positive test results or was “likely” or “highly likely” to have been infected with HIV. The HIV status of the deceased was coded as “missing” if the respondent did not know the likelihood that their sibling was infected with HIV. All other deaths were deaths of HIV-negative persons.

We used data collected prospectively in Karonga district as reference against which we evaluated survey data on HIV-related mortality. This an area of Malawi where the local economy is dominated by subsistence farming, fishing and small-scale trading. Since 2002, a HDSS covers a predominantly rural population of  $\approx 40,000$  people in the southern part of Karonga district.<sup>24,25</sup> After a baseline census, key informants from HDSS villages continuously record births, deaths and migrations among the HDSS population. Each HDSS resident is attributed a unique identification (ID) number. Parental IDs are added to individual HDSS records. This allows listing the known siblings of most HDSS residents, by looking up records that share parental IDs.

Extensive HIV data are available in Karonga. Prior to the HDSS, studies of leprosy and tuberculosis have included HIV testing.<sup>26,27</sup> HDSS residents were then offered HIV testing repeatedly between 2005 and 2011. These serosurveys achieved participation rates  $>70\%$ , and many participants learned their HIV status.<sup>28</sup> HIV-related data (including HIV test results) from registers at health facilities serving HDSS communities have been linked to HDSS records.<sup>29–31</sup> Ancillary studies have asked selected HDSS residents to self-report the result and date of recent HIV tests.<sup>32,33</sup> The HDSS also elicits the HIV status of deceased residents during post-mortem verbal autopsies, but we did not consider these data due to limited reliability.<sup>34</sup>

We constructed a reference classification of adult deaths by HIV status of the deceased, based solely on pre-mortem data from prospective datasets. Deaths of PHIV were deaths for which a

clinical record or self-report of an HIV-positive test was available. Deaths of persons who were HIV-negative were deaths for which there was no evidence of a positive test, but a clinical record or self-report of an HIV-negative test was available within five years of the death.<sup>35,36</sup> Deaths without evidence of a positive or negative HIV test had missing reference data on HIV status.

### *Data collection and linkages*

We selected study participants at random among HDSS residents aged 15 to 59 years old. We oversampled residents with one or more adult death(s) among known siblings. We oversampled at a higher rate the residents whose deceased sibling(s) was(ere) classified as PHIV by reference data. We selected at most one participant per family. We recruited all participants during household visits.

We tested two modes of data collection in a parallel randomized trial: face-to-face (FTF) interviewing and audio computer assisted interviewing (ACASI). FTF is the standard approach in household surveys: an interviewer asks questions and records respondents' answers. In ACASI, respondents are equipped with headsets through which they hear pre-recorded questions; then they enter answers on their own, using a keypad.<sup>37</sup> We hypothesized that ACASI might improve reporting of the HIV status of deceased siblings because it is more confidential than FTF.<sup>38</sup>

Participants were randomized 1:1 to FTF or ACASI. We stratified the randomization by participant gender and mortality of their known siblings (i.e., whether they had a deceased sibling who was a PHIV). All interviews were conducted in the local language (Chitumbuka).

Interviewers did not have access to reference data on family composition, HIV status and mortality. For the ACASI version of the questionnaire, a team member pre-recorded questions

and instructions in Chitumbuka. We designed images indicating where respondents should push on the screen if they wanted to enter a specific answer. We used color schemes with high contrasts and symbols to ensure that color-blind participants could use ACASI. ACASI included a training section, wherein we asked respondents simple questions so they could become familiar with self-interviewing. Respondents assigned to ACASI switched to FTF if they reported vision problems, e.g., not being able to see the screen in order to answer questions. To prevent selective manipulation of study group assignments, interviewers learned the allocation of respondents to ACASI or FTF only after an individual had consented to participate and his/her vision had been assessed. All data were collected using Open Data Kit (ODK).

We used FTF to ask questions about the respondent's background (e.g., age, educational level) and to collect the standard components of siblings' survival histories (i.e., names, sex and age of siblings). When respondents assigned to ACASI reported adult deaths, interviewers handed them the tablet and headset to answer questions about that death (i.e., HIV questions, accidents/injuries, and pregnancy/childbirth). Respondents assigned to FTF were asked these questions directly by the interviewer. We collected metadata on the interview from ODK (e.g., time spent by question).

Following data collection, 2 reviewers used data on names and sex to independently link the lists of maternal siblings reported by survey respondents to HDSS records. For each match between lists, they attributed his/her HDSS ID to the reported sibling. When reviewers disagreed in their matching outcomes, a third reviewer adjudicated. To avoid confirmation bias, none of the reviewers had access to survey or reference data on siblings' vital status, ages, or HIV status.

## *Data analysis*

We described the background of respondents by assigned study group, including ownership of a mobile phone and prior use of the Internet. We included these latter variables to assess respondents' familiarity with digital tools similar to ACASI. We described the reported characteristics of siblings (sex, age). We assessed the proportion of reported adult deaths for which we could not ascertain HIV status from survey data and explored whether this proportion varied by reported time since death. We created a binary variable taking value 1 if the death had occurred within 8 years of the survey ("recent deaths"), and 0 otherwise ("earlier deaths").

The pre-specified primary outcomes were the sensitivity and specificity of survey data on HIV status of deceased siblings. We defined sensitivity as the proportion of deaths to PHIV according to reference data that were also classified as such from survey data. Specificity was the proportion of deaths to persons who were HIV-negative according to reference data that were also classified as such from survey data. The amount of time spent answering HIV questions – measured from ODK metadata– was a secondary outcome of the trial. All study group comparisons were conducted with participants included in their assigned study group (intent-to-treat). All analyses of study outcomes used survey weights to account for differences in selection probabilities and participation rates between HDSS residents in different sampling strata. We assessed differences in proportions (e.g., sensitivity) between study groups using chi-square tests, and differences in interviewing times using non-parametric rank sum tests.

Partial verification bias might affect analyses of primary outcomes,<sup>39</sup> because reference data on the HIV status of deceased siblings are only available for a subset of the deaths reported during



survey interviews. This usually leads to sensitivity estimates that are too high and specificity estimates that are too low.<sup>40</sup> We imputed missing reference data on HIV status as recommended.<sup>40,41</sup> We assumed that reference data were missing at random (MAR), i.e., their availability depended on reported characteristics of the deceased. We included sex, age at death, time since death and survey-assessed HIV status of the deceased in imputation models. We created 20 imputed datasets through chained equations, and we combined results using standard techniques.<sup>42,43</sup> We did not pre-specify these adjusted analyses.

## Results

The study was conducted between October 12<sup>th</sup>, 2018 and January 14<sup>th</sup>, 2019. We allocated 307 HDSS residents to FTF and 306 to ACASI. Participation rates were 88.6% in the FTF group (272/307) and 87.3% in the ACASI group (267/306).

[FIGURE 1 ABOUT HERE]

The educational level of participants was low (table 1), with more than a third of respondents not having completed primary education. Few respondents had prior experience of digital tools, e.g., slightly more than 1 in 10 had ever used the internet. The number of known siblings based on HDSS data varied across respondents: approximately 5% of respondents had no known sibling, whereas 48.5% of respondents in the ACASI group and 44.3% in the FTF group had 4 or more known siblings. Respondents listed 3,414 maternal siblings during survey interviews. There were 885 reported adult deaths: 427 in the ACASI group and 458 in the FTF group. In the FTF group, 38.2% of reported deaths were recent vs 44.6% in the ACASI group. The complete reported distributions of ages at death, and dates of death, are available in appendix A2.

[TABLE 1 ABOUT HERE]

Based on data generated by the HIV questions (appendix A3), the HIV status of the deceased sibling could not be determined for 6.8% of adult deaths reported by ACASI respondents (95% CI = 4.4 to 9.2) and 13.5% of deaths reported by FTF respondents (95% CI = 10.4 to 16.7, figure 2). Among recent deaths, missing survey data on HIV status was as common in ACASI (5.3%, 95% CI = 2.1 to 8.4) and FTF (6.3%, 95% CI = 2.7 to 9.9). Among earlier deaths, the proportion of siblings whose HIV status could not be ascertained from survey data was lower in ACASI data (8.0%, 95% CI = 4.6 to 11.5) than in FTF data (18.0%, 95% CI = 13.5 to 22.5).

[FIGURE 2 ABOUT HERE]

Reference data on HIV status of the deceased were available for 262 of 885 reported adult deaths (29.6%, appendix A4). Among those, 180 deceased siblings were PHIV, and 82 were HIV-negative individuals, according to reference data. The availability of reference data on HIV status did not vary between study groups. However, reference data were more frequently available for deaths that were classified as deaths of PHIV on the basis of survey data.

Among reported deaths with reference data on HIV status, sensitivity was 0.89 (95% CI = 0.85 to 0.92, table 2) in the ACASI group and 0.91 (95% CI = 0.88 to 0.93) in the FTF group. After accounting for partial verification bias, estimates of sensitivity declined to 0.82 (95% CI = 0.46 to 0.96) in the ACASI group and 0.78 (95% CI = 0.44 to 0.94) in the FTF group.

[TABLE 2 ABOUT HERE]

In the subset of deaths for which reference data were available, the specificity of survey data on HIV status of the deceased was 0.94 (95% CI = 0.90 to 0.97) in ACASI and 0.91 (95% CI = 0.86 to 0.95) in FTF. Estimates of specificity increased when we accounted for partial verification bias (0.98 in ACASI vs. 0.96 in FTF).

ACASI training lasted a median time of 7.4 minutes (appendix A5). Asking the HIV questions required a median time of 0.4 minutes per deceased adult sibling in the FTF group and 0.7 minutes in the ACASI group (figure 3). In FTF interviews, this is comparable to the median time required to collect information on pregnancy-related deaths (0.4 minutes in FTF) and on accidental/external deaths (0.2 minute in FTF).

[FIGURE 3 ABOUT HERE]

## **Discussion**

We tested whether accurate data on HIV-related mortality could be collected during household surveys. Such surveys already constitute the main source of representative data on all-cause and pregnancy-related mortality in most African countries. We thus added HIV questions to the standard adult and maternal mortality of the DHS questionnaire and conducted a validation study of these data in Karonga district in northern Malawi.

Adding HIV questions resulted in limited amounts of missing data on the HIV status of respondents' deceased siblings, particularly among deaths that had occurred within 8 years of the survey (5–6%). This is important because adult mortality rates are often estimated solely from

survey data on recent deaths and person-years.<sup>45</sup>

The added HIV questions generated a classification of deceased siblings according to their HIV status that was often accurate. In analyses accounting for partial verification bias, approximately 8 out of 10 deceased siblings who were PHIV were correctly reported as such by survey respondents. Our estimates of the specificity of survey data on HIV status were 0.96–0.98 in adjusted analyses. These levels are comparable to estimates of the sensitivity/specificity of survey data in recording pregnancy-related and injury-related deaths.<sup>46–51</sup> Whereas questions about these latter circumstances are included in most recent DHS and MICS, HIV questions have not yet been included. This is a missed opportunity because collecting additional HIV-related data requires limited time: in face-to-face interviews, asking HIV questions took 0.4 minute (i.e., 25 seconds) per deceased adult sibling, on average.

We hypothesized that ACASI might yield more accurate data than FTF, but we found limited support for this hypothesis. ACASI generated less missing data than FTF, but only for deaths that occurred more than 8 years prior to the study. ACASI data did not have higher sensitivity than FTF data, but it required more time to collect mortality data. Due to ACASI's limited effectiveness in improving accuracy, and the high burden it places on respondents, our study does not justify recommending the use of ACASI to collect mortality data in household surveys.

There are several limitations. First, our reference data do not constitute a gold standard in ascertaining the HIV status of deceased siblings. They rely partly on linkage with clinical registers from health facilities to establish the HIV status of the deceased. These linkages might be made erroneously, e.g., in the case of two persons with similar names. This might lead to estimates of the sensitivity/specificity of survey data that are too low. Reference data on HIV

status were occasionally based on self-reports (appendix A4). Since self-reported data on HIV test results might be affected by social desirability bias, our reference classification might include PHIVs who were incorrectly classified as being HIV-negative.<sup>22</sup> This would lead to estimates of the specificity of survey data that are too low. We also classified deceased siblings as HIV-negative if they had received negative test results within 5 years of their death. This definition might miss individuals who have seroconverted within a few months/years prior to death. Given low (recent) levels of HIV incidence in Malawi<sup>52</sup> and patterns of disease progression however, such misclassifications are likely rare in our reference data.

Second, reference data on HIV status were only available for a subset of the deaths reported during the survey interview. We addressed this partial verification bias through multiple imputations of the reference data. Multiple imputations assumed that reference data were missing at random (MAR) among reported deaths. If the availability of reference data was also determined by other unobserved factors, then reference data were missing not at random and multiple imputations might not have adequately adjusted our estimates of sensitivity/specificity.

Third, the results of our study might not reflect the levels of sensitivity/specificity that could be achieved if HIV questions were included in household surveys such as DHS. This is so because in Karonga, the scale-up of HIV-related services has occurred earlier than in similar settings. Between 2005 and 2011, the HDSS population was targeted by multiple HIV serosurveys, during which adult residents were offered home-based HIV testing. A large fraction of PHIV in the area became aware of their HIV status and might subsequently have disclosed this information to their siblings.<sup>53</sup> In other settings where HIV testing was not scaled-up in a similar manner, the accuracy of survey data in recording the HIV status of deceased siblings might be lower.

Fourth, study participants were residents of the HDSS area, whose siblings had died when they also resided in this area. Due to this geographic proximity, our study might have included respondents who had increased knowledge of the HIV status of their deceased siblings. The accuracy of survey data on HIV status of deceased siblings might be lower among family members who resided further away from the deceased. However, this might not necessarily be the case since migrants are often central in providing the financial support required to care for PHIV. They might also return home to visit sick relatives and might otherwise remain in touch with their siblings through mobile communications.

Finally, due to limited sample size, we did not conduct sub-group analyses. We thus did not test whether the sensitivity/specificity of survey data on HIV status of deceased siblings might vary with characteristics of a) the respondents (e.g., gender, educational levels), b) the deceased siblings (e.g., gender, age at death, time since death), or c) the interviewer. We also did not investigate whether ACASI might help improve the accuracy of survey data for some population groups, e.g., those with secondary schooling and/or prior experience with digital tools.

In conclusion, we suggest a simple strategy for improving the measurement HIV/AIDS mortality in African countries with limited data. Adding a few questions to questionnaires widely used by household surveys might generate data on the excess mortality that persons with HIV experience in a population. Such questions require limited additional time to collect, but they might help calibrate the statistical models used to track progress towards global HIV elimination objectives.

## REFERENCES CITED

1. Report of the UN Secretary-General on the Implementation of the Declaration of Commitment on HIV/AIDS and the political declarations on HIV/AIDS (Seventy-third session of the General Assembly) 2 April 2019 - A/73/824. [https://www.unaids.org/en/resources/documents/2019/A\\_73\\_824](https://www.unaids.org/en/resources/documents/2019/A_73_824). Accessed December 23, 2019.
2. Mikkelsen L, Phillips DE, AbouZahr C, et al. A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet*. 2015;386(10001):1395-1406. doi:10.1016/S0140-6736(15)60171-4
3. Johnson LF, Anderegg N, Zaniewski E, et al. Global variations in mortality in adults after initiating antiretroviral treatment: an updated analysis of the International epidemiology Databases to Evaluate AIDS cohort collaboration. *AIDS*. 2019;33 Suppl 3:S283-S294. doi:10.1097/QAD.0000000000002358
4. Haas AD, Zaniewski E, Anderegg N, et al. Retention and mortality on antiretroviral therapy in sub-Saharan Africa: collaborative analyses of HIV treatment programmes. *J Int AIDS Soc*. 2018;21(2). doi:10.1002/jia2.25084
5. Reniers G, Wamukoya M, Urassa M, et al. Data Resource Profile: Network for Analysing Longitudinal Population-based HIV/AIDS data on Africa (ALPHA Network). *Int J Epidemiol*. 2016;45(1):83-93. doi:10.1093/ije/dyv343
6. Bor J, Herbst AJ, Newell M-L, Barnighausen T. Increases in adult life expectancy in rural South Africa: valuing the scale-up of HIV treatment. *Science*. 2013;339(6122):961-965. doi:10.1126/science.1230413
7. Case KK, Johnson LF, Mahy M, Marsh K, Supervie V, Eaton JW. Summarizing the results and methods of the 2019 Joint United Nations Programme on HIV/AIDS HIV estimates. *AIDS*. 2019;33 Suppl 3:S197-S201. doi:10.1097/QAD.0000000000002440
8. GBD 2017 HIV collaborators. 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-e859. doi:10.1016/S2352-3018(19)30196-1
9. Stover J, Glaubius R, Mofenson L, et al. Updates to the Spectrum/AIM model for estimating key HIV indicators at national and subnational levels. *AIDS*. 2019;33 Suppl 3:S227-S234. doi:10.1097/QAD.0000000000002357
10. Hallett TB, Zaba B, Stover J, et al. Embracing different approaches to estimating HIV incidence, prevalence and mortality. *AIDS*. 2014;28 Suppl 4:S523-532. doi:10.1097/QAD.0000000000000488
11. Corsi DJ, Neuman M, Finlay JE, Subramanian SV. Demographic and health surveys: a profile. *Int J Epidemiol*. 2012;41(6):1602-1613. doi:10.1093/ije/dys184

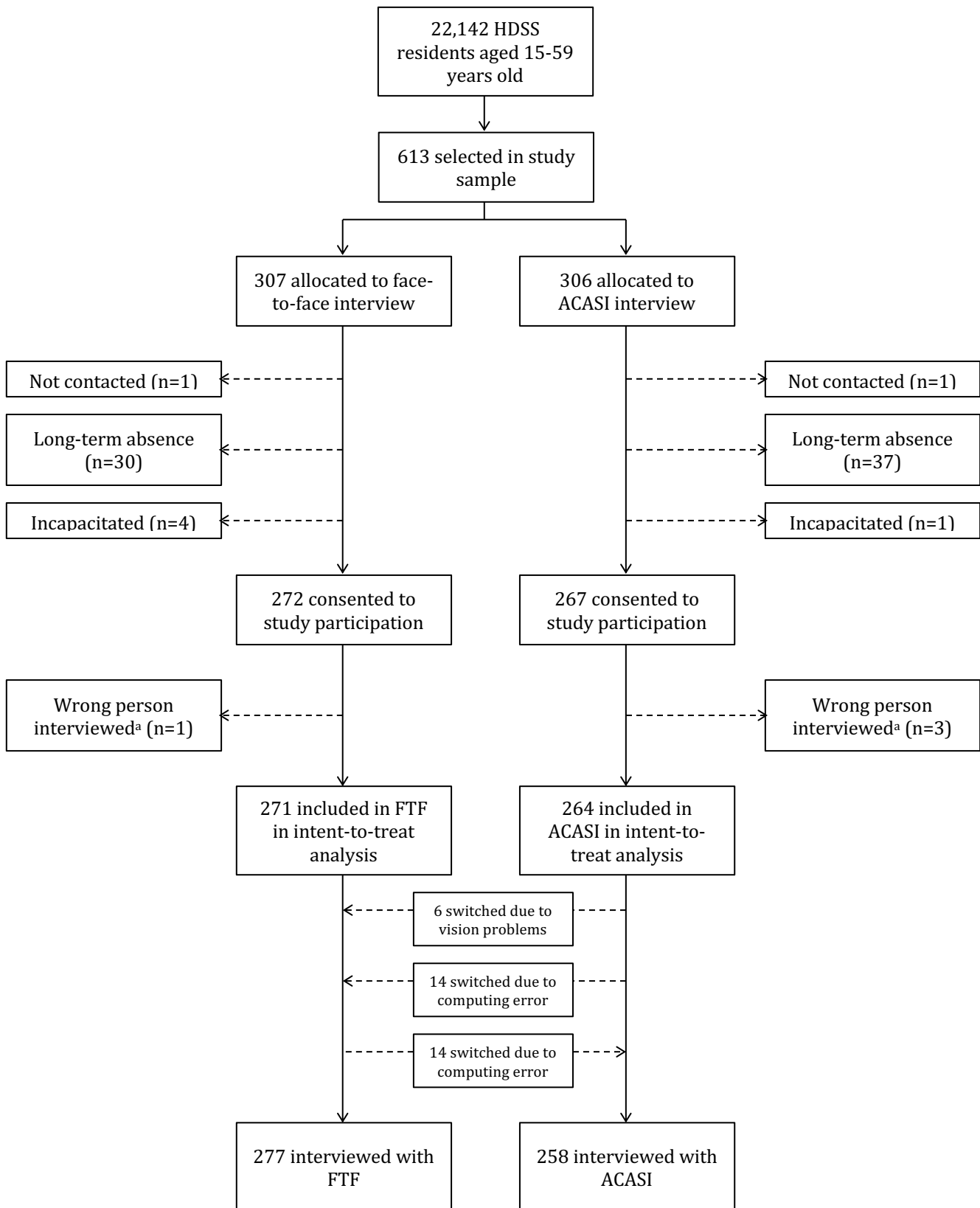
12. Khan S, Hancioglu A. Multiple Indicator Cluster Surveys: Delivering Robust Data on Children and Women across the Globe. *Stud Fam Plann.* 2019;50(3):279-286. doi:10.1111/sifp.12103
13. Reniers G, Masquelier B, Gerland P. Adult Mortality in Africa. In: Rogers RG, Crimmins EM, eds. *International Handbook of Adult Mortality*. International Handbooks of Population. Dordrecht: Springer Netherlands; 2011:151-170. doi:10.1007/978-90-481-9996-9\_7
14. Hill K, Choi Y, Timaeus I. Unconventional approaches to mortality estimation. *Demographic Research.* 2005;S4(12):281-300. doi:10.4054/DemRes.2005.13.12
15. Obermeyer Z, Murray CJL, Gakidou E. Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from the world health survey programme. *BMJ.* 2008;336(7659):1482-1486. doi:10.1136/bmj.a137
16. Maman D, Chilima B, Masiku C, et al. Closer to 90-90-90. The cascade of care after 10 years of ART scale-up in rural Malawi: a population study. *J Int AIDS Soc.* 2016;19(1):20673. doi:10.7448/IAS.19.1.20673
17. Maman S, van Rooyen H, Groves AK. HIV status disclosure to families for social support in South Africa (NIMH Project Accept/HPTN 043). *AIDS Care.* 2014;26(2):226-232. doi:10.1080/09540121.2013.819400
18. Tam M, Amzel A, Phelps BR. Disclosure of HIV serostatus among pregnant and postpartum women in sub-Saharan Africa: a systematic review. *AIDS Care.* 2015;27(4):436-450. doi:10.1080/09540121.2014.997662
19. Watkins SC. Navigating the AIDS Epidemic in Rural Malawi. *Population and Development Review.* 2004;30(4):673-705. doi:10.1111/j.1728-4457.2004.00037.x
20. Tsai AC, Venkataramani AS. The causal effect of education on HIV stigma in Uganda: Evidence from a natural experiment. *Soc Sci Med.* 2015;142:37-46. doi:10.1016/j.socscimed.2015.08.009
21. Chan BT, Tsai AC, Siedner MJ. HIV Treatment Scale-Up and HIV-Related Stigma in Sub-Saharan Africa: A Longitudinal Cross-Country Analysis. *Am J Public Health.* 2015;105(8):1581-1587. doi:10.2105/AJPH.2015.302716
22. Johnston LG, McFarland W, Sabin ML, et al. Measuring self-reported HIV status in bio-behavioural surveys. *Bull World Health Organ.* 2015;93(5):287-287A. doi:10.2471/BLT.15.153064
23. Hill K, El Arifeen S, Koenig M, Al-Sabir A, Jamil K, Raggars H. How should we measure maternal mortality in the developing world? A comparison of household deaths and sibling history approaches. *Bull World Health Organ.* 2006;84(3):173-180. doi:10.2471/blt.05.027714
24. Jahn A, Branson K, Crampin A, et al. Evaluation of a village-informant driven demographic surveillance system in Karonga, Northern Malawi. *Demographic Research.* 2007;16(8):219-248. doi:10.4054/DemRes.2007.16.8



25. Crampin AC, Dube A, Mboma S, et al. Profile: the Karonga Health and Demographic Surveillance System. *Int J Epidemiol*. 2012;41(3):676-685. doi:10.1093/ije/dys088
26. Crampin AC, Glynn JR, Ngwira BM, et al. Trends and measurement of HIV prevalence in northern Malawi. *AIDS*. 2003;17(12):1817-1825. doi:10.1097/00002030-200308150-00011
27. Shaw MA, Turner AC, Blackwell JM, Fine PE, Ponnighaus JM. Setting up HIV serology for the Karonga leprosy vaccine trial in Malaŵi. *Lepr Rev*. 1991;62(1):87-104. doi:10.5935/0305-7518.19910012
28. Molesworth AM, Ndhlovu R, Banda E, et al. High accuracy of home-based community rapid HIV testing in rural Malawi. *J Acquir Immune Defic Syndr*. 2010;55(5):625-630. doi:10.1097/QAI.0b013e3181f98628
29. Houben RMGJ, Van Boeckel TP, Mwinuka V, et al. Monitoring the impact of decentralised chronic care services on patient travel time in rural Africa--methods and results in Northern Malawi. *Int J Health Geogr*. 2012;11:49. doi:10.1186/1476-072X-11-49
30. Koole O, Houben RM, Mzembe T, et al. Improved retention of patients starting antiretroviral treatment in Karonga District, northern Malawi, 2005-2012. *J Acquir Immune Defic Syndr*. 2014;67(1):e27-e33. doi:10.1097/QAI.0000000000000252
31. Price AJ, Kayange M, Zaba B, et al. Uptake of prevention of mother-to-child-transmission using Option B+ in northern rural Malawi: a retrospective cohort study. *Sex Transm Infect*. 2014;90(4):309-314. doi:10.1136/sextrans-2013-051336
32. Chilunga FP, Musicha C, Tafatatha T, et al. Investigating associations between rural-to-urban migration and cardiometabolic disease in Malawi: a population-level study. *Int J Epidemiol*. October 2019. doi:10.1093/ije/dyz198
33. Nakanga WP, Prynn JE, Banda L, et al. Prevalence of impaired renal function among rural and urban populations: findings of a cross-sectional study in Malawi. *Wellcome Open Res*. 2019;4:92. doi:10.12688/wellcomeopenres.15255.1
34. Mclean EM, Chihana M, Mzembe T, et al. Reliability of reporting of HIV status and antiretroviral therapy usage during verbal autopsies: a large prospective study in rural Malawi. *Glob Health Action*. 2016;9. doi:10.3402/gha.v9.31084
35. Byass P, Calvert C, Miiro-Nakiyingi J, et al. InterVA-4 as a public health tool for measuring HIV/AIDS mortality: a validation study from five African countries. *Global Health Action*. 2013;6(s3):22448. doi:10.3402/gha.v6i0.22448
36. Reniers G, Blom S, Calvert C, et al. Trends in the burden of HIV mortality after roll-out of antiretroviral therapy in KwaZulu-Natal, South Africa: an observational community cohort study. *The Lancet HIV*. 2017;4(3):e113-e121. doi:10.1016/S2352-3018(16)30225-9
37. Mensch BS, Hewett PC, Erulkar AS. The reporting of sensitive behavior by adolescents: a methodological experiment in Kenya. *Demography*. 2003;40(2):247-268. doi:10.1353/dem.2003.0017

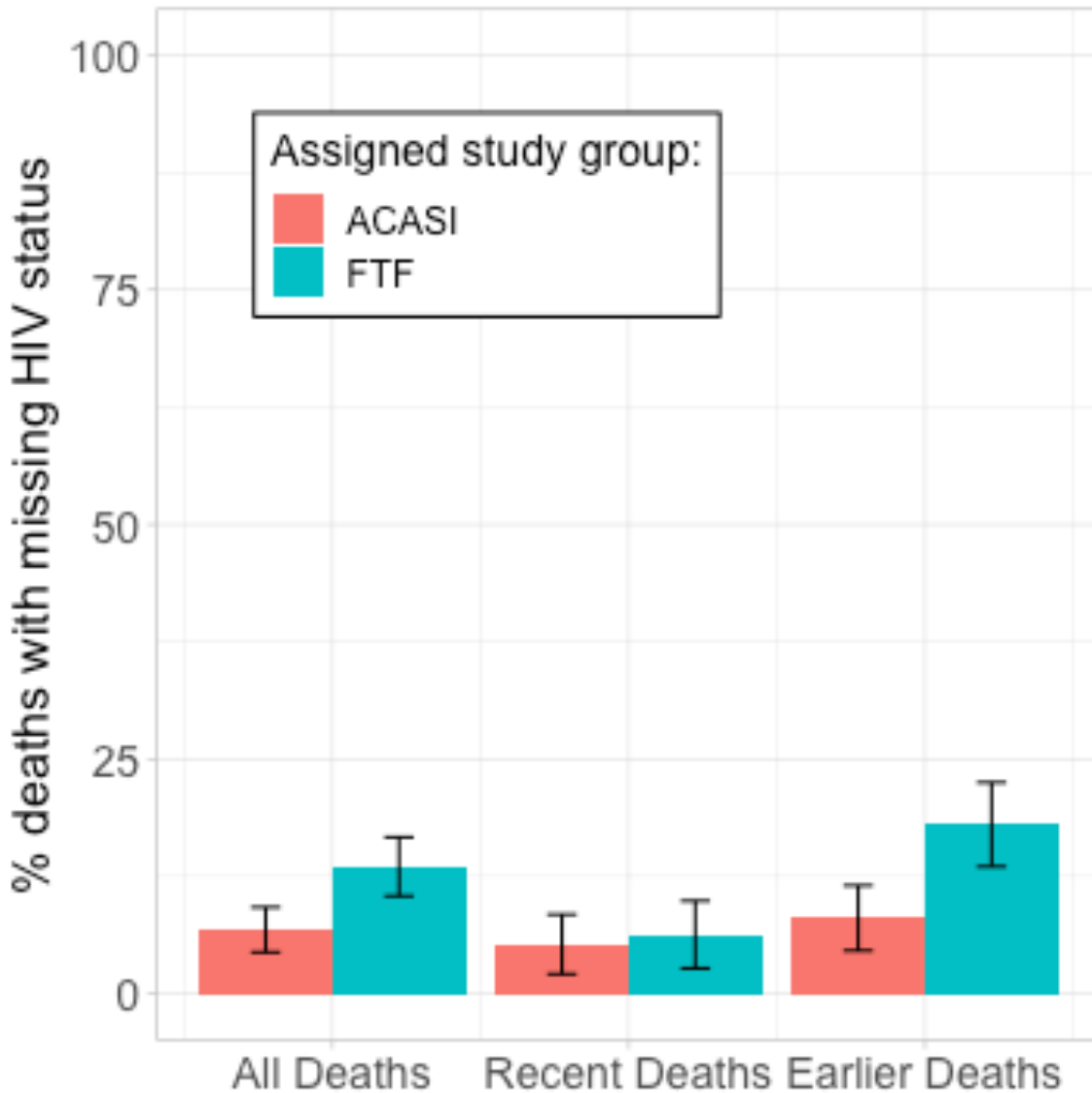
38. Phillips AE, Gomez GB, Boily M-C, Garnett GP. A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries. *Int J Epidemiol*. 2010;39(6):1541-1555. doi:10.1093/ije/dyq114
39. O'Sullivan JW, Banerjee A, Heneghan C, Pluddemann A. Verification bias. *BMJ Evid Based Med*. 2018;23(2):54-55. doi:10.1136/bmjebm-2018-110919
40. de Groot JAH, Bossuyt PMM, Reitsma JB, et al. Verification problems in diagnostic accuracy studies: consequences and solutions. *BMJ*. 2011;343:d4770. doi:10.1136/bmj.d4770
41. de Groot J a. H, Janssen KJM, Zwinderman AH, Moons KGM, Reitsma JB. Multiple imputation to correct for partial verification bias revisited. *Stat Med*. 2008;27(28):5880-5889. doi:10.1002/sim.3410
42. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581-592. doi:10.1093/biomet/63.3.581
43. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. John Wiley & Sons; 2019.
44. Walker E, Nowacki AS. Understanding Equivalence and Noninferiority Testing. *J GEN INTERN MED*. 2011;26(2):192-196. doi:10.1007/s11606-010-1513-8
45. Masquelier B, Eaton J, Gerland P, Pelletier F, Mutai K. Age patterns and sex ratios of adult mortality in countries with high HIV prevalence. *Aids*. 2017;31. doi:10.1097/QAD.0000000000001332
46. Helleringer S, Pison G, Masquelier B, et al. Improving survey data on pregnancy-related deaths in low-and middle-income countries: a validation study in Senegal. *Trop Med Int Health*. 2015;20(11):1415-1423. doi:10.1111/tmi.12583
47. Helleringer S, Pison G, Masquelier B, et al. Improving the quality of adult mortality data collected in demographic surveys: validation study of a new siblings' survival questionnaire in Niakhar, Senegal. *PLoS Med*. 2014;11(5):e1001652. doi:10.1371/journal.pmed.1001652
48. Helleringer S, Duthé G, Kanté AM, et al. Misclassification of pregnancy-related deaths in adult mortality surveys: case study in Senegal. *Trop Med Int Health*. 2013;18(1):27-34. doi:10.1111/tmi.12012
49. Shahidullah M. The sisterhood method of estimating maternal mortality: the Matlab experience. *Stud Fam Plann*. 1995;26(2):101-106.
50. Shahidullah M. A comparison of sisterhood information on causes of maternal death with the registration causes of maternal death in Matlab, Bangladesh. *Int J Epidemiol*. 1995;24(5):937-942. doi:10.1093/ije/24.5.937
51. Pison G, Masquelier B, Kante AM, et al. Estimating mortality from external causes using data from retrospective surveys: A validation study in Niakhar (Senegal). *DemRes*. 2018;38:879-896. doi:10.4054/DemRes.2018.38.32

52. Ministry of Health, Malawi. *Malawi Population-Based HIV Impact Assessment (MPHIA) 2015-2016: Final Report*. Lilongwe, Malawi: Ministry of Health; 2018.  
[https://phia.icap.columbia.edu/wp-content/uploads/2019/08/MPHIA-Final-Report\\_web.pdf](https://phia.icap.columbia.edu/wp-content/uploads/2019/08/MPHIA-Final-Report_web.pdf). Accessed December 23, 2019.
53. Price AJ, Glynn J, Chihana M, et al. Sustained 10-year gain in adult life expectancy following antiretroviral therapy roll-out in rural Malawi: July 2005 to June 2014. *Int J Epidemiol*. 2017;46(2):479-491. doi:10.1093/ije/dyw208



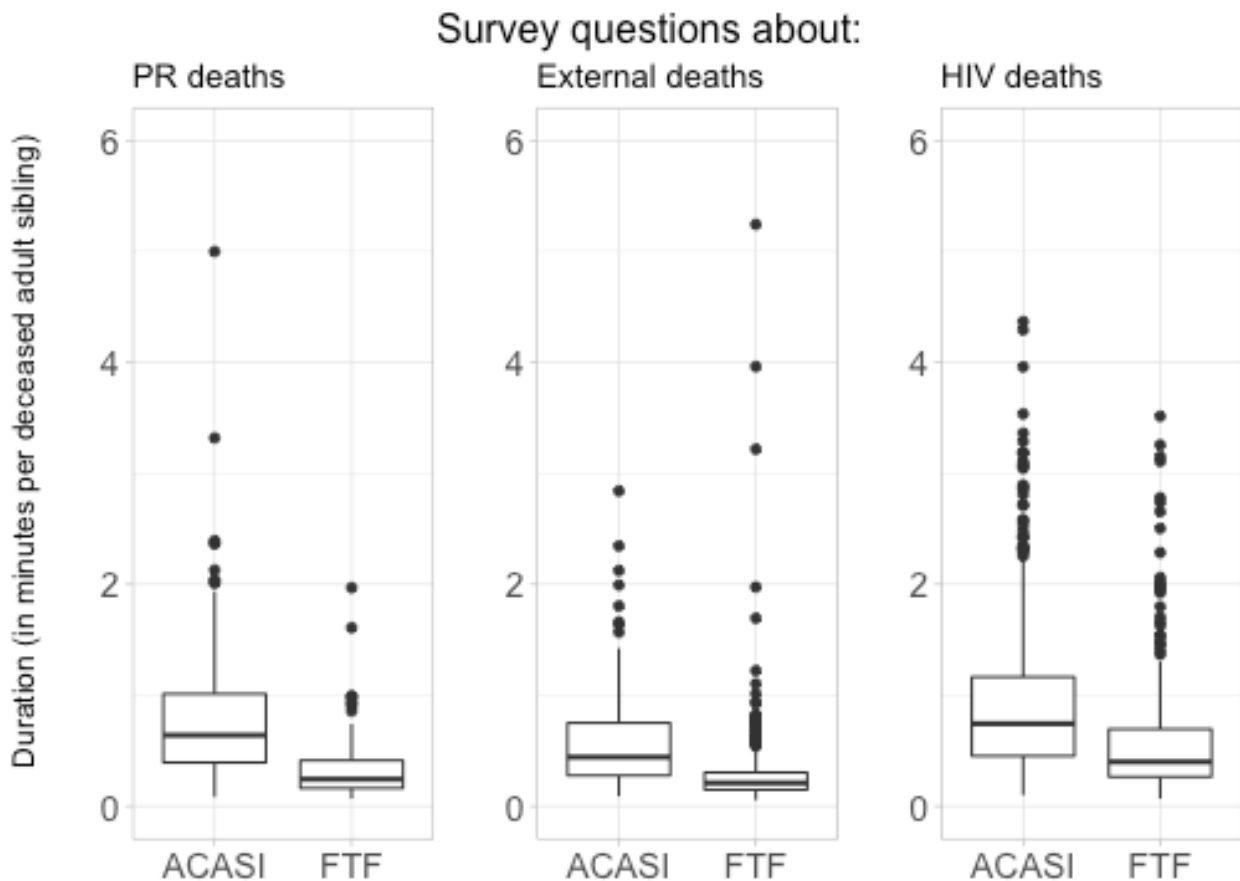
**Figure 1: Enrollment process**

*Notes:* <sup>a</sup> These errors were first detected at the data editing stage, due to large discrepancies in age of the respondent between survey and HDSS data (>20 years); they were subsequently confirmed during household re-visits by study supervisors.



**Figure 2: percentage of deceased adult siblings with missing data on HIV status, by assigned mode of interview and reported timing of the death (n=885).**

*Notes:* Recent deaths are deaths that have occurred over the past 8 years according to reports from survey respondents. Earlier deaths are deaths that have occurred more than 8 years ago according to reports from survey respondents. Vertical bars represent 95% confidence intervals. The calculations of these confidence intervals accounted for the clustering of adult deaths within families/respondents, and for the stratified survey design. All calculations used the “survey” package in R to incorporate survey weights. One respondent did not report the time since the death of her adult sibling; as a result, the analyses by reported timing of the death were conducted with n = 884.



**Figure 3: Distribution of time required to collect data on circumstances of reported adult deaths, by study group.**

*Notes:* PR = pregnancy-related; For each circumstance of death (i.e., PR deaths, external deaths and HIV deaths), based on rank sum tests, we rejected the null hypothesis that the distribution of interviewing time was similar in the ACASI and FTF arms ( $p < 0.001$ ). All estimates reported in this figure were calculated per reported adult death (i.e.,  $n = 885$ ) for external deaths and HIV deaths; for PR deaths, these estimates were calculated for deaths of women of reproductive age ( $n = 384$ ).

	ACASI	FTF
<b>Respondents (n=535)</b>		
<i>Sex</i>		
Male	130 (49.2)	132 (48.7)
Female	134 (50.8)	139 (51.3)
<i>Age</i>		
15-24y	40 (15.2)	30 (11.1)
25-34y	60 (22.7)	68 (25.1)
35-44y	69 (26.1)	85 (31.4)
≥45y	95 (36.0)	88 (32.5)
<i>Education<sup>a</sup></i>		
No schooling/incomplete primary	95 (36.3)	97 (36.5)
Completed primary	80 (30.5)	76 (28.6)
Secondary schooling or higher	87 (33.2)	93 (34.9)
<i>Mobile phone<sup>b</sup></i>		
Owens a mobile phone	163 (61.7)	174 (64.2)
Does not own a mobile phone	101 (38.3)	97 (35.8)
<i>Internet usage</i>		
Has ever used internet	31 (11.7)	37 (13.7)
Has never used internet	103 (39.0)	102 (37.6)
Does not know about internet	130 (49.3)	132 (48.7)
<i>Number of known/expected siblings<sup>c</sup></i>		
0	13 (4.9)	14 (5.2)
1	32 (12.1)	41 (15.1)
2	45 (17.1)	55 (20.3)
3	46 (17.4)	41 (15.1)
4+	128 (48.5)	120 (44.3)
<b>All reported siblings (n=3,414)</b>		
<i>Sex</i>		
Male	853 (50.5)	913 (52.9)
Female	836 (49.5)	812 (47.1)
<i>Vital status</i>		
Alive	954 (56.5)	953 (55.2)
Deceased	728 (43.1)	769 (44.6)
Unknown	7 (0.4)	3 (0.2)
<i>Age at death</i>		
0-14y	274 (37.6)	289 (37.6)
≥15y	427 (58.7)	458 (59.6)
Unknown	27 (3.7)	22 (2.8)
<b>Reported adult deaths (n=885)</b>		
<i>Age at death</i>		
15-24y	63 (14.8)	61 (13.3)
25-34y	123 (28.8)	148 (32.3)
35-44y	147 (34.4)	145 (31.7)
45-54y	67 (15.7)	77 (16.8)
≥55y	27 (6.3)	27 (5.9)
<i>Date of death<sup>d</sup></i>		
Recent deaths (<8y before survey)	190 (44.6)	175 (38.2)
Earlier deaths (≥8y before survey)	236 (55.4)	283 (61.8)

**Table 1: Characteristics of study participants, by assigned study group.**

Notes: <sup>a</sup> 5 respondents reported never attending school; <sup>b</sup> mobile phone ownership was defined as “personal ownership”, rather than “household ownership”; <sup>c</sup> siblings of all ages; <sup>d</sup> The date of one reported adult death was missing.

	ACASI	FTF	P-value
<b><u>Sensitivity</u></b>			
<i>Reference data</i> <sup>a</sup>	0.89 (0.85 to 0.92)	0.91 (0.88 to 0.93)	0.40
<i>Imputed data</i> <sup>b</sup>	0.82 (0.46 to 0.96)	0.78 (0.44 to 0.94)	0.28
<b><u>Specificity</u></b>			
<i>Reference data</i> <sup>a,c</sup>	0.94 (0.90 to 0.97)	0.91 (0.86 to 0.95)	0.71
<i>Imputed data</i> <sup>b</sup>	0.98 (0.79 to 1.00)	0.96 (0.81 to 0.99)	0.39

**Table 2: sensitivity/specificity of survey data in recording HIV status of deceased adult siblings, by assigned mode of interview.**

*Notes:* <sup>a</sup> Reference HDSS data were available for 262 deaths reported during the survey, 180 of which were deaths of persons who were HIV-positive and 82 of which were deaths of persons classified as HIV-negative.

<sup>b</sup> We imputed the reference HIV status when survey data on HIV status were not missing (n=794). We generated 20 datasets through chained equations, in which we imputed missing values of reference HIV status. We did so using the “mice” package in R.

<sup>c</sup> The null hypothesis of a one-sided test of the non-inferiority of ACASI relative to FTF with  $\delta = 0.08$  was rejected with  $p = 0.029$ . In this test, the null hypothesis assumes that the specificity of ACASI is inferior to the specificity of FTF by  $\delta = 0.08$ .



## Appendix A1: mathematical framework

We outline a framework detailing how the data collected through the HIV questions that we added to siblings' survival histories might help better assess the contribution of HIV to adult mortality in affected SSA countries.

We note  $S$  the number of siblings (including selected respondents) listed during a survey that included a sibling survival history module (e.g., a DHS). We note  $D$  a binary variable indicating that one of these siblings died during a given timeframe, from any cause of death. We note  $H$  a binary variable that indicates whether a sibling has been infected with HIV. For simplicity, we consider that all deaths occur exactly at the same time (e.g., during a reference period). This assumption can easily be relaxed in future refinements of this framework, given that data on the timing of deaths are collected during siblings' survival histories.

The addition of the HIV questions to siblings' survival histories then allows estimating  $P(H|D)$ , i.e., the proportion of deaths that occurred to PHIV. From Bayes' rule, we have:

$$P(D|H) = \frac{P(D) \times P(H|D)}{P(H)}, \text{ and } P(D|\bar{H}) = \frac{P(D) \times P(\bar{H}|D)}{P(\bar{H})} = \frac{P(D) \times [1 - P(H|D)]}{1 - P(H)}, \text{ where } P(H) \text{ is the}$$

proportion of all siblings who have become infected with HIV irrespective of survival status, and

$P(\bar{H})$  is the proportion of all siblings that have not become infected with HIV. Since  $H$  is a

binary variable, we have  $P(\bar{H}) = 1 - P(H)$ . The mortality ratio associated with HIV infection,

i.e.,  $\frac{P(D|H)}{P(D|\bar{H})}$  can thus be re-written as:

$$\sigma = \frac{P(H|D) \times [1 - P(H)]}{P(H) \times [1 - P(H|D)]}$$

When  $\sigma > 1$ , PHIV die at faster rate than individuals who are HIV-negative, and  $\sigma$  measures the excess mortality faced by PHIV in the population of interest. Success in prevention, care and treatment of HIV can then be measured by the pace at which  $\sigma$  converges to 1 over time (e.g., from one survey to the next).

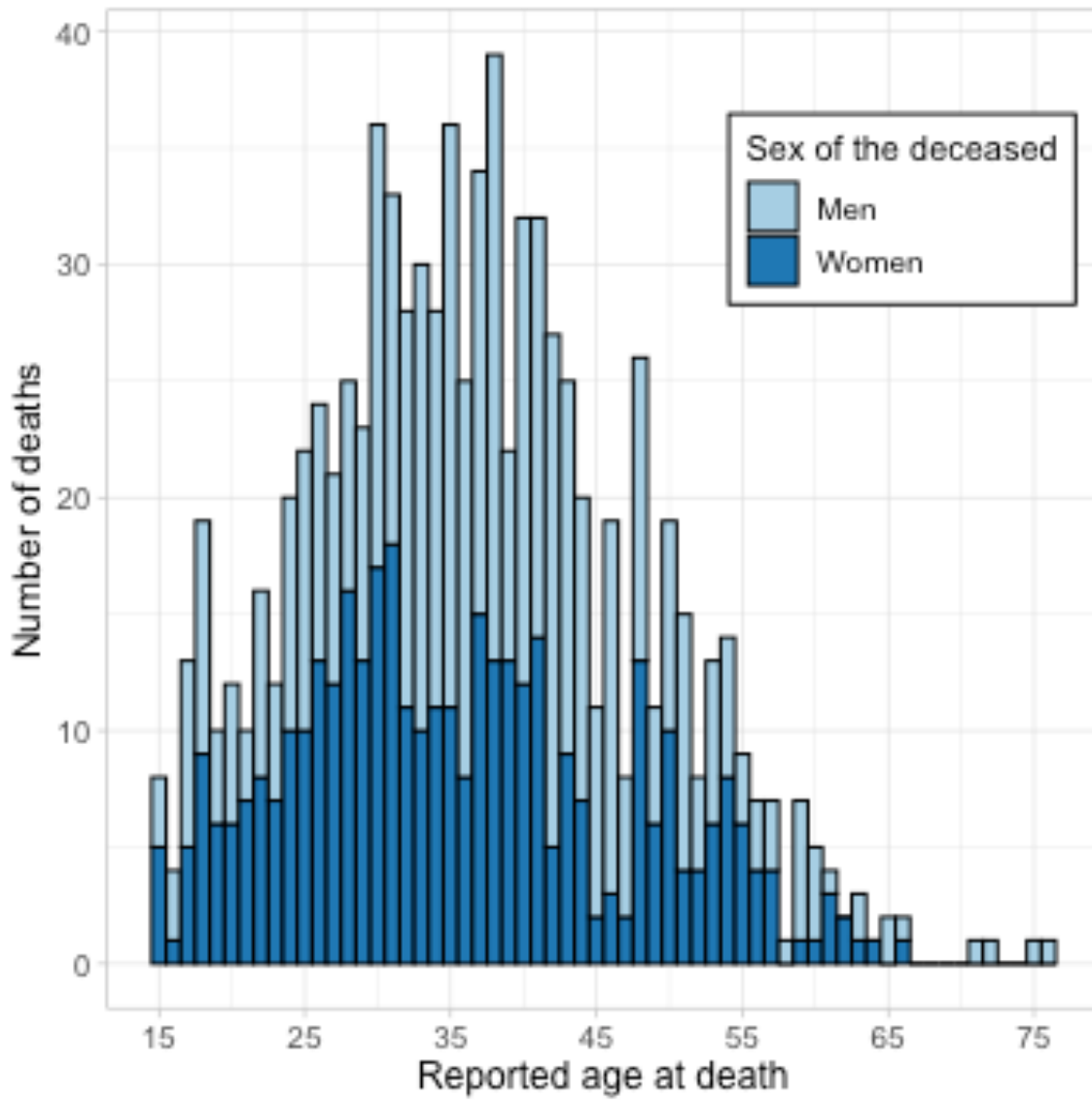
From the formula of total probabilities, we can expand  $P(H)$  and rewrite  $\sigma$  as follows:

$$\sigma = \frac{P(H|D) \times (1 - [(1 - P(D)) \times P(H|\bar{D}) + P(D) \times P(H|D)])}{[(1 - P(D)) \times P(H|\bar{D}) + P(D) \times P(H|D)] \times [1 - P(H|D)]}$$

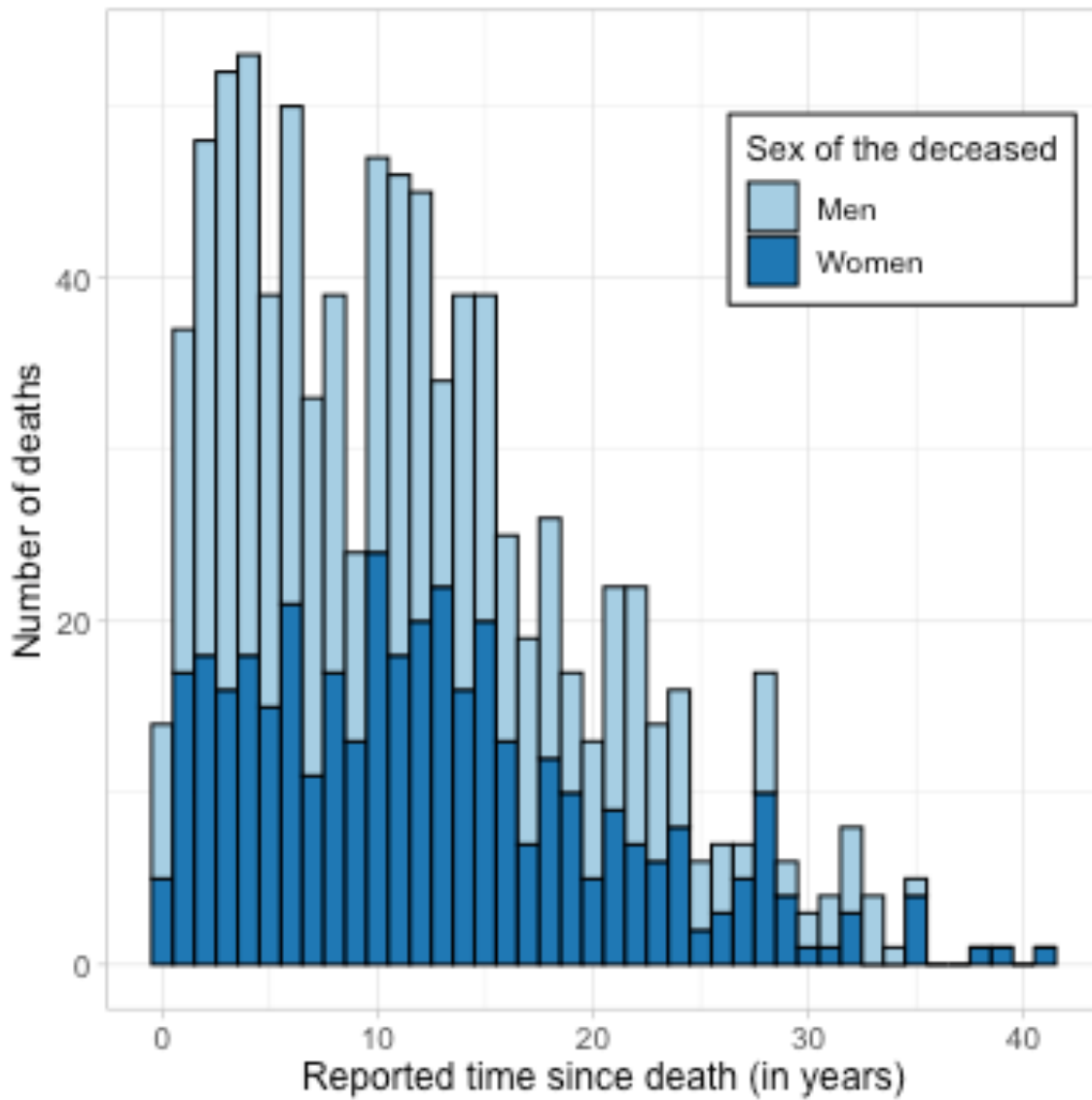
In our expression of  $\sigma$ ,  $P(H|D)$  can be estimated from survey data, if the 3 HIV questions yield accurate data. Estimates of  $P(H|D)$  can be adjusted for potential misclassifications and other reporting errors, using data on sensitivity and specificity from validation studies such as the one reported in this paper.

$P(D)$  can be estimated directly from the siblings' survival histories. On the other hand, data on  $P(H|\bar{D})$ , the prevalence of HIV among surviving siblings, are available through several modalities. In some surveys,  $P(H|\bar{D})$  can be measured directly, through sero-testing of survey respondents (e.g., some DHS). In other surveys,  $P(H|\bar{D})$  can also be measured through self-reported data, by asking respondents to report their most recent test results. These approaches require the assumption that the sample of survey respondents is a random sample of surviving siblings. In addition, data on  $P(H|\bar{D})$  can also be potentially collected during siblings' survival histories, by extending the HIV-related questions to all siblings, whether they are alive or deceased at the time of the survey.

**Appendix A2: Distributions of ages at death and time since death**

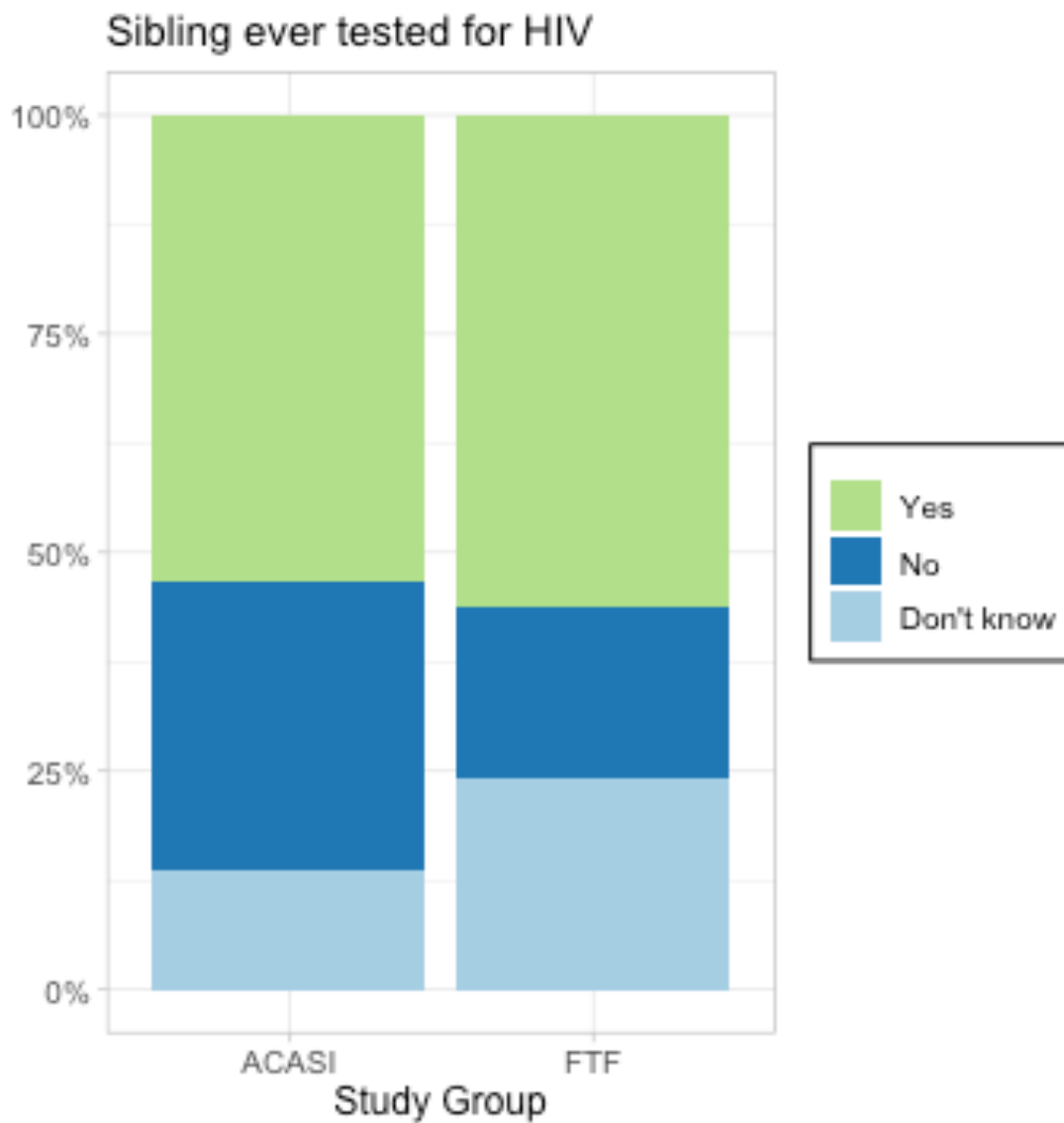


**Figure A2-1: Reported ages at death, by sex of the deceased (n=885)**

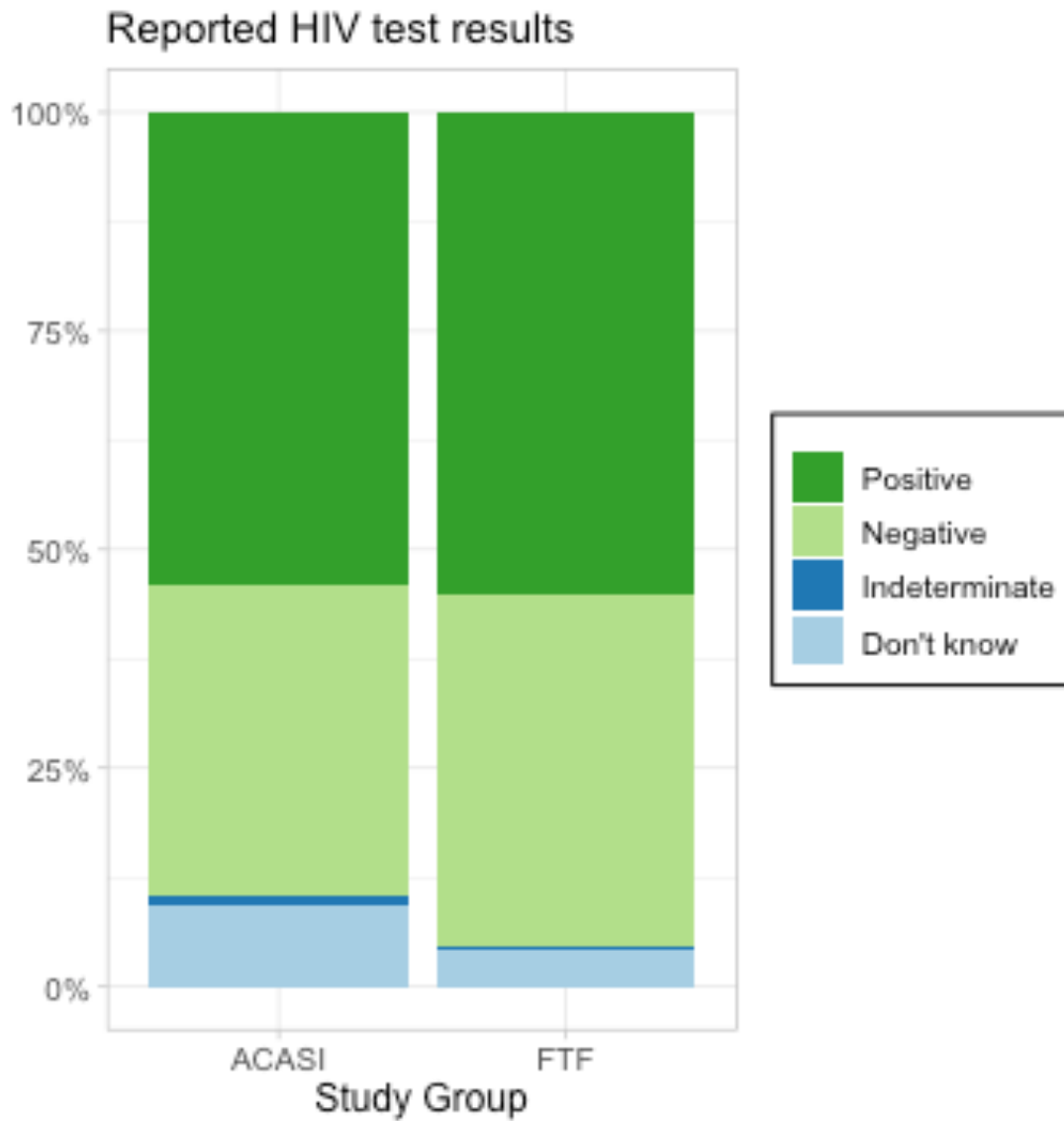


**Figure A2-2: Reported numbers of years since the death, by sex of the deceased (n = 884).**  
*Notes:* the date of the death was not reported for one of the reported adult deaths.

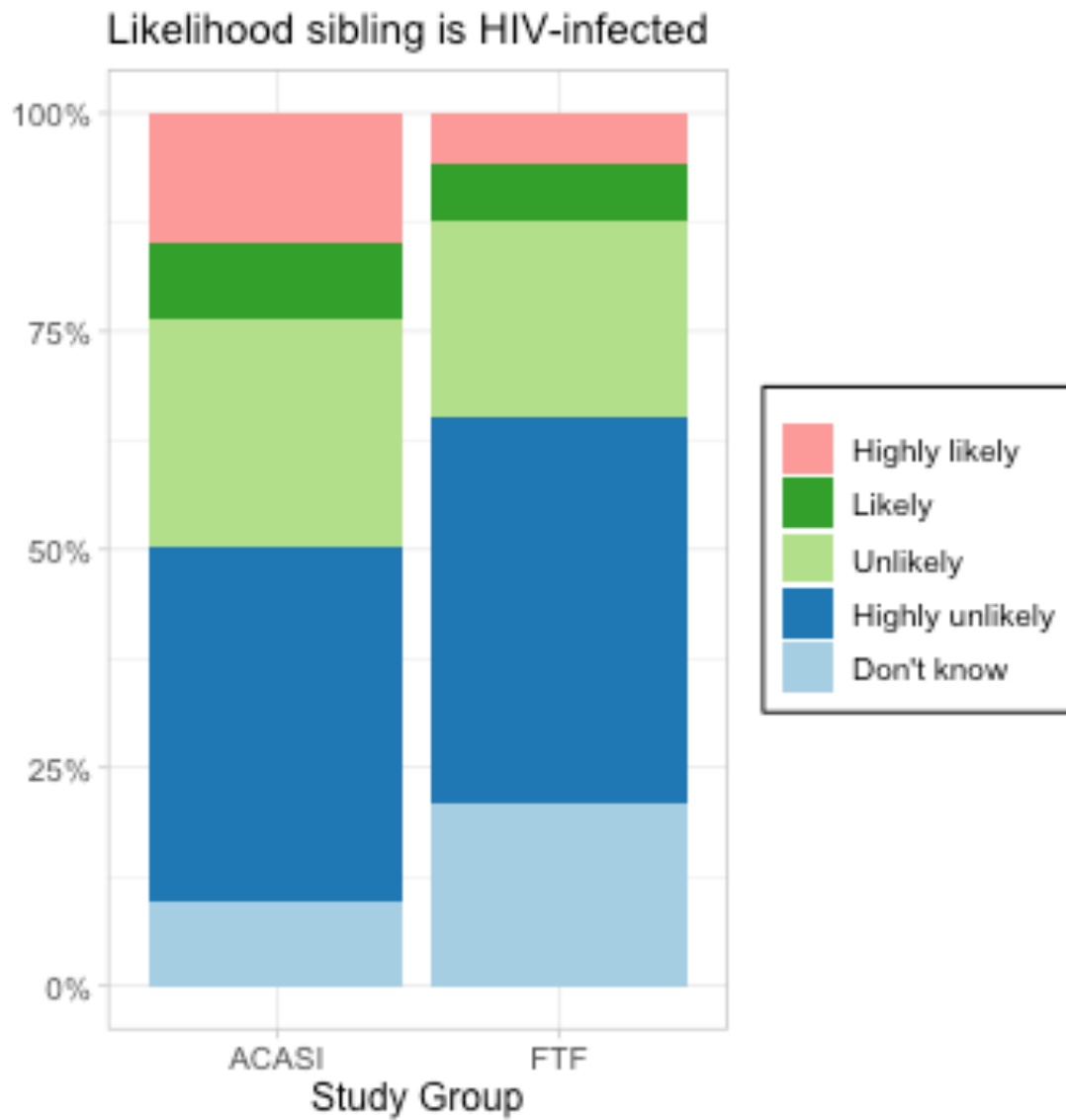
**Appendix A3: data generated by HIV questions**



**Figure A3-1: Distribution of answers to question “Has [deceased sibling] ever been tested for HIV?”, by study group (n=885)**

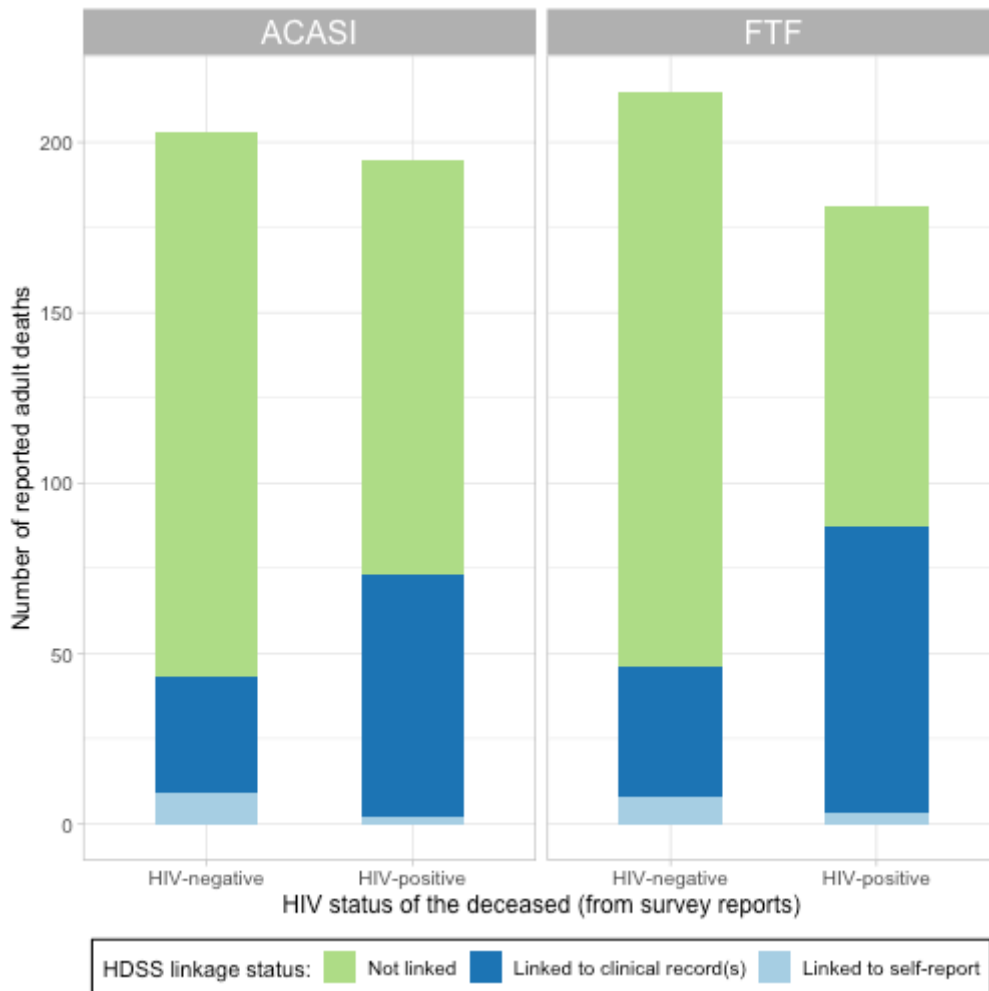


**Figure A3-2: Distribution of answers to question “What were the results of [deceased sibling]’s latest HIV test”, by study group (n=486)**



**Figure A3-3: Distribution of answers to question “How likely do you think it is that [deceased sibling] was infected with HIV?”, by study group (n=620)**

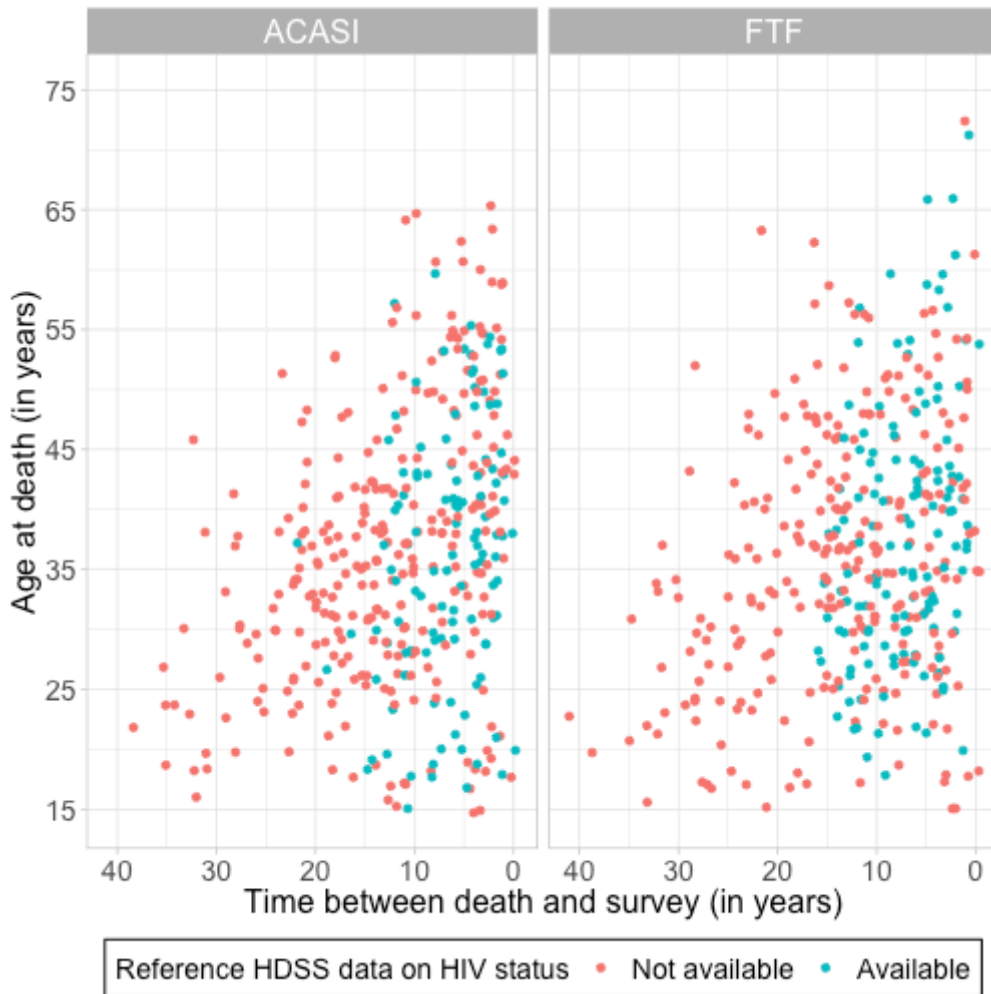
## Appendix A4: Linkages between survey and reference data



**Figure A4.1: Outcomes of linkages between reported survey data and HDSS datasets, by reported HIV status of the deceased and study group.**

*Notes:* Clinical records in the HDSS datasets include results of the serosurveys conducted between 2005 and 2011, as well as data routinely extracted from clinical registers at facilities serving the HDSS population.

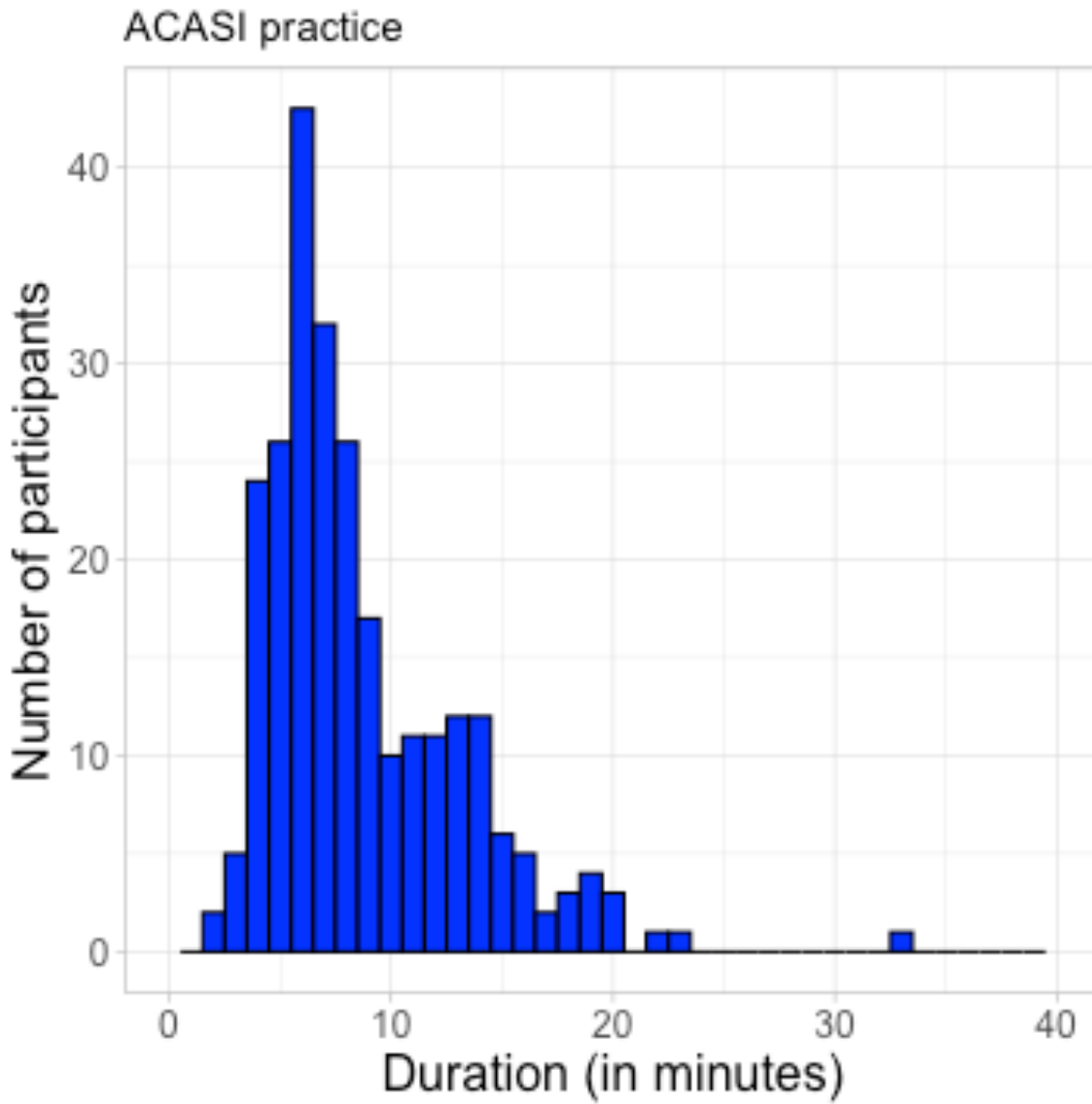




**Figure A4.2: Joint distribution of reported ages at, and time since, death (n=884), by study group and availability of reference HDSS data.**

*Notes:* one death was not included on the graph due to missing data on date of death.

**Appendix A5: duration of ACASI training**



**Figure A5-1: Duration of ACASI practice**

Notes: duration estimates were obtained using the audit feature in ODK.