Misreporting of pregnancy-related deaths in siblings' survival histories: a validation study in Niakhar, Senegal.

Estimates of maternal mortality in countries with limited vital registration increasingly rely on data from siblings' survival histories (SSH). These data are collected during retrospective household-based surveys such as the DHS. Demographers have expressed concern that such data may underestimate the proportion of deaths due to maternal causes. We conducted a validation study of SSH in Niakhar (Senegal), a locality where prospective data on adult mortality has been collected during demographic surveillance since 1962. We found that SSH collected using the DHS questionnaire significantly overestimated the proportion of deaths due to pregnancy-related causes. On the other hand, a modified SSH questionnaire (which incorporated recall cues and an event history calendar) yielded unbiased estimates in settings where more than 15% of deaths are due to pregnancy-related causes.

Background

Maternal deaths are defined as 'the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes" (ICD-10). The Millennium Development Goals (MDGs) call for a reduction of the maternal mortality ratio (i.e., the number of maternal deaths per 1,000 live births) by three quarters between 1990 and 2015 (MDG5). It is however difficult to estimate how fast various countries are making progress towards MDG5. In a large number of low-income countries, the coverage of death registration is very low and selective [1, 2]. When deaths are registered, the cause of death is often poorly ascertained [2-5].

In the absence of high-quality vital registration data on maternal mortality, the United Nations Maternal Mortality Estimation Group (UN MMEIG) and the Institute for Health Metrics and Evaluation (IHME) use statistical models to produce estimates of maternal mortality [6-11]. Both models have the same structure. In a first step, estimates of all-cause adult mortality are produced. This is referred to as the "mortality envelope". In a second step, the number of maternal deaths is estimated by multiplying the mortality envelope by the proportion of deaths among women of reproductive age due to maternal causes. Through this process, the IHME estimated that the number of maternal deaths worldwide declined from 376,000 in 1990 to 293,000 in 2013 [11], corresponding to an annualized rate of change of -1.1%. By comparison, the UN MMEIG estimated that maternal deaths declined much more rapidly between 1990 and 2013 at a rate of -3.1% per year [8, 12] .

In countries with limited vital registration data, the statistical models of the UN MMEIG and IHME are informed by a variety of data sources. To estimate the mortality envelope, they may rely on model life tables, censuses, burial or mortuary surveillance, and surveys [13-17]. To estimate the proportion of maternal deaths, they increasingly rely on siblings' survival histories (SSH) collected during nationally representative surveys such as the Demographic and Health Surveys (DHS). During SSH [18], respondents are asked to report the complete list of their maternal siblings by birth order. They are then asked to report the sex, survival status and age of each sibling. Current

age is recorded for surviving siblings, whereas age at death and years since death are recorded for the deceased ones. If the sister of a respondent died at age 12 years or older, respondents are asked whether she died while pregnant, at the time of delivery or within 42 days of her most recent delivery. If the answer is 'yes' to any one of these three questions, the sister's death is classified as "pregnancy related" (PR). Pregnancy related deaths differ from maternal deaths as defined by ICD-10 because they also include deaths from accidents and injuries. But in the absence of other data, PR deaths are used as a proxy for maternal deaths. In the most recent IHME estimates, SSH data on PR mortality were used to produce MMR estimates for 38 African countries [11]. SSH data also represented 81% of the observations included in the dataset used by the UN MMEIG to estimate maternal mortality in sub-Saharan countries [7].

The quality of SSH data on PR mortality is thus a critical determinant of the validity of maternal mortality estimates. Demographers consider that SSH data may under-estimate the proportion of PR deaths [19], in particular if SSH respondents are not aware of their sister's pregnancy. In a validation study conducted in Matlab (Bangladesh), which compared SSH to prospective data on adult mortality from a Health and Demographic Surveillance System (HDSS, [20, 21]), SSH data had high specificity but low sensitivity: respondents omitted, or misclassified as non-PR, deaths following induced abortions or deaths due to indirect maternal causes [22, 23]. In estimating MMRs, SSH data thus need to be adjusted upwards "to correct for the likely under-identification of deaths from maternal causes that is thought to occur almost universally" [7]. In the UN MMEIG model, for example, the raw SSH estimates of the proportion of PR deaths were multiplied by a factor of 1.1 prior to inclusion in the MMR model.

In sub-Saharan countries, assessments of the quality of SSH data [24-26] have focused on aggregate indicators of maternal mortality (e.g., MMRs), rather than on the proportion of PR deaths. In 2010, we conducted a SSH/HDSS comparison in Bandafassi HDSS in Senegal [27] which suggested that unlike in Bangladesh SSH may overestimate the proportion of PR deaths [28]. A number of SSH respondents reported deaths as having occurred during pregnancy or within 42 days of delivery, when the HDSS recorded these deaths as not PR. This study was however limited by small sample sizes. In this paper, we report further investigation of biases in SSH estimates of the proportion of PR deaths in a different rural population in Senegal. We also

investigated whether a modified SSH questionnaire, the siblings' survival calendar (SSC), permits obtaining more accurate SSH data on PR mortality than the standard questionnaire used in particular during the DHS [29].

Data and Methods

***Reference dataset*:** The reference dataset for the validation of SSH data on maternal mortality comes from the Niakhar HDSS, located in the Fatick region of Senegal. Activities of the Niakhar HDSS started in 1962 in eight villages of the Niakhar area and were later expanded to 30 villages in 1983 [30]. An initial baseline census was carried out in 1962, followed by another census in 1983, when the study area was expanded. Since then, data on demographic events – births, deaths, marriages, pregnancies and migrations - have been collected from household informants during household visits. Study interviewers use a roster of household residents and inquire about the vital status of each household member, as well as possible changes in marital status and births since the previous household visit. For each death, interviewers conduct a verbal autopsy (VA), during which a relative of the deceased (or someone who knew the deceased well) is asked about the circumstances of the death. In particular, VA respondents are asked whether the deceased was pregnant at the time of death. Physicians then review each VA questionnaire and attribute a cause of death using ICD-9 codes [31-33]. The HDSS data have been used to estimate levels and determinants of maternal mortality in the Niakhar area [31, 32, 34]. In the HDSS dataset, every population member is potentially linked to his/her biological mother through a mother ID number. This number is attributed either at the time of birth (if the mother gave birth in the HDSS area) or the first time an individual enters the HDSS population (i.e., initial census or after in-migration). By using the mother ID number, we can identify any population member's maternal siblings (i.e., his/her sibship) within the Niakhar HDSS population.

SSH data collection: in 2013, we conducted a SSH survey in the Niakhar HDSS area. We selected a stratified sample of population members, which included a) members of 592 sibships in which there was at least one adult death since the beginning of the HDSS, and b) members of 500 other sibships, e.g., sibships in which all siblings were still alive or sibships in which deaths took place before age 15 We then randomized half

of this sample to an SSH interview with a standard DHS questionnaire and the other half to an interview with the siblings' survival calendar (SSC). Compared to the standard DHS questionnaire, the SSC includes several supplementary interviewing techniques (e.g., recall cues) to prevent omissions of siblings during SSH. It also uses an event history calendar to improving the reporting of ages and dates during SSH. SSH respondents included both men and women aged 15-59 years old. Former residents of the Niakhar HDSS area who had migrated to Dakar (Senegal's capital), Mbour (a large coastal town) or within 50 kilometers of the Niakhar HDSS were traced and interviewed. A team of 8 interviewers who had previously worked on one the Senegal DHS conducted SSH interviews. In total, we interviewed 609 respondents from the DHS group and 580 from the SSC group. The SSC trial is registered with ISRCTN (http://www.controlled-trials.com/ISRCTN06849961/). Further details on the SSH data collection have been provided elsewhere [29]. Initial results showed that the SSC improved the quality of SSH data on all-cause adult mortality over the standard DHS questionnaire.

Pregnancy-related deaths in reference dataset: To be consistent with the definition used in SSH, our classification of deaths as PR or not PR in the reference dataset is solely based on the recorded timing of deaths. It uses data from the VA questionnaire about the pregnancy status of the deceased, but it does not rely on physician review of VA data. Among the 592 sibships in which there was at least one adult death since the beginning of the HDSS, we identified the subset of sibships in which a woman died at reproductive age, i.e., between 15-49 years old. We then used HDSS data to assess whether the deceased sibling had given birth within 42 days of her death. To do so, we calculated the difference (in days) between her date of death and the date of her most recent delivery. Among deaths that had not occurred within 42 days of a delivery, we then reviewed VA questionnaires to determine whether the deceased was reported as pregnant at the time of the death. We focused on deaths having occurred within the past 15 years to match the reference period used by IHME in producing maternal mortality estimates [9, 11].

Analytical framework: SSH respondents belong to three groups according to the HDSS dataset: they are either members of 1) a sibship in which there was one PR death within the past 15 years, 2) a sibship in which one adult woman died of non-PR causes within

the past 15 years or 3) a sibship in which no adult woman died in the past 15 years. According to the HDSS, there was at most one PR death per sibship over the past 15 years, whereas there were some sibships with more than one non-PR deaths. Based on this classification, we can then write the proportion of PR deaths reported during SSH as:

$$P(M) = \frac{D_{PR}}{D_{PR} + D_{NPR}} \quad (1)$$

where D_{PR} is the number of deaths reported as PR during SSH and D_{NPR} is the number of deaths reported as non-PR. We can write

$$D_{PR} = N_{PR} \times P(R|N_{PR}) \times P(C|R, N_{PR}) + N_{NPR} \times P(R|N_{NPR}) \times [1 - P(C|R, N_{NPR})] \quad (2)$$

and

$$D_{NPR} = N_{PR} \times P(R|N_{PR}) \times [1 - P(C|R, N_{PR})] + N_{NPR} \times P(R|N_{NPR}) \times P(C|R, N_{NPR}) \quad (3)$$

where N_{PR} is the number of SSH respondents in sibships with a PR death according to the HDSS; N_{NPR} is the number of SSH respondents in sibships with a non-PR death; $P(R|N_{PR})$ and $P(R|N_{NPR})$ are the probabilities of reporting an adult death (irrespective of cause) among respondents in sibships with a PR death and in sibships with a non-PR death, respectively; $P(C|R, N_{PR})$ is the probability of correctly classifying the cause of a reported adult death among respondents in sibship with a PR death (i.e., the sensitivity of SSH); and $P(C|R, N_{NPR})$ is the probability of correctly classifying the cause of a reported adult death among respondents in sibship with a non-PR death (i.e., the specificity of SSH).

Data analysis: we calculated each parameter for the DHS and SSC questionnaires separately. For each type of questionnaire, we tested for differences in the probability of reporting an adult death (irrespective of cause) between sibships with a PR death and sibships with a non-PR death. To do so we used a χ^2test of the association between categorical variables. We also tested for differences in sensitivity and specificity between the DHS and SSC questionnaires using similar tests. Standard errors were adjusted for

the clustering of respondents within sibships. For cross-tabulations in which some cells have n<5, we used exact tests. We then used these estimates to assess the extent of bias in SSH measures of the proportion of PR deaths. We considered samples of 1,000 respondents in which the true proportion of PR deaths varied between 0.1 and 0.5. We calculated the reported proportion of PR deaths for each type of questionnaire using (1), (2) and (3). The relative bias was defined as the difference between the reported proportion of PR deaths and the true proportion of PR deaths, divided by the true proportion.

Robustness tests: HDSS data do not constitute a "gold standard" for the measurement of maternal mortality. Validation studies in which HDSS data were compared against post-mortem data from health facilities indicated that some PR deaths may be misclassified as non-PR by HDSS and VA data [35, 36]. To assess whether this potential bias affects our study results, we conducted the following series of robustness tests. We considered that HDSS was highly specific, i.e., all non-PR related deaths were classified as such by the HDSS. But we assumed various sensitivity levels for the HDSS data, i.e., 95%, 90%, 80% and 70%. Based on these levels, we reclassified as PR a random sample of deaths initially classified as non-PR by the HDSS dataset. We repeated this reclassification 1,000 times, and we recalculated all parameters in each bootstrap sample. We plotted the median, as well as the 2.5^{th} and 97.5^{th} percentiles of the bootstrap distribution of each parameter to assess the robustness of our findings.

Results

Descriptive statistics:

There were 136 deaths among women aged 15-49 years in the past 15 years according to the HDSS dataset. Among those, we could not determine whether the death was PR in 13 cases (9.6%) because the VA questionnaire was missing. Among deaths for which VA data were available, 38 were PR deaths (30.9%) and 85 were non-PR deaths (69.1%). The characteristics of these deaths are described in table 1. There were no differences in family size between PR and non-PR deaths (6.7 vs. 6.2). The average time since death was 7.9 years for PR deaths vs. 6.9 years for non-PR deaths (p=0.19). The average age at death, on the other hand, was 26.2 years for PR deaths vs. 27.9 years for non-PR deaths (p = 0.34). According to the HDSS, 21% of PR deaths occurred

during pregnancy vs. 50% at the time of the delivery and 29% between 1 and 42 days of the delivery. According to physician review of the VA data, 22 out 28 deaths classified as PR were due to complications of pregnancy, childbirth and the puerperium (i.e., “maternal causes”, ICD-9 codes 630-674). Twenty-nine out of 85 deaths classified as non-PR were due to infectious & parasitic diseases (ICD-9 codes 001 to 139). The father or mother of the deceased primarily provided VA information including whether the deceased was pregnant at the time of death. Other respondents included one of the in-laws of the deceased, her husband or a co-wife. Very few VA interviews were conducted with a sibling of the deceased (2 out of 38 for PR deaths vs. 6 out of 85 for non-PR deaths).

[table 1 about here]

In total, we interviewed 50 respondents who were siblings of one of the 38 women who had died of PR causes over the past 15 years. Among those, 22 were assigned to the SSC group vs. 28 assigned to the DHS group (figure 1). We also interviewed 114 respondents who were siblings of women who had died of non-PR causes. Among those, 59 were assigned to the SSC group vs. 55 assigned to the DHS group.

[Figure 1 about here]

Reporting of adult deaths by cause of death:

In table 2, we report significant differences in the reporting of adult death by cause of death among respondents interviewed with the DHS questionnaire. Twenty-seven out of 28 respondents whose sister had died of PR death according to the HDSS reported an adult female death during the DHS interview (96.4%). On the other hand, only 39 out of 55 (70.9%) respondents whose sister had died of non-PR death according to the HDSS reported an adult female death during the DHS interview (p=0.007). This difference was even larger among respondents whose sister had died at ages below 25 years old (93.8% vs. 48%, p = 0.003). Among respondents interviewed with the SSC questionnaire, there were no significant differences in reporting of adult deaths by cause of death.

[table 2 about here]

Classification of causes of death by SSH data:

Table 3 describes measures of the sensitivity of SSH data in identifying PR deaths. We found no differences in sensitivity between the DHS and the SSC questionnaires. Respondents correctly classified the death of their sister as PR in more than 90% in both questionnaires (25/27 for the DHS questionnaire, i.e., 92.6% vs. 18/20 for the SSC questionnaire, i.e., 90%).

[Table 3 about here]

Table 4 describes measures of the specificity of SSH data. We found that the SSC had high specificity: 48/50 SSC respondents correctly classified the death of their sister as non-PR (95.4%). The specificity of the DHS questionnaire was significantly lower: only 31 out of 39 DHS respondents correctly classified the death of their sister as non-PR (79.5%, $p = 0.014$).

[Table 4 about here]

Bias in estimates of the proportion of pregnancy-related deaths:

Figure 1 reports estimates of the extent bias in estimates of the proportion of PR deaths by type of questionnaire. It indicates that the DHS questionnaire over-estimates this proportion. When the true proportion of PR deaths is 0.1, the DHS questionnaire yields an estimate of the proportion of 0.299 (i.e., a 198.6% bias); when the true proportion of PR deaths is 0.3, the DHS questionnaire yields an estimate of 0.467 (i.e., a 55% bias); finally, when the true proportion of PR deaths is 0.5, the DHS questionnaire yields an estimate of 0.619 (i.e., a 24% bias). The extent of bias in SSC estimates of the proportion of PR deaths was much lower. In settings where the true proportion was 0.1, the SSC yields an estimate of 0.127 (i.e., a 27% bias). In settings where the true proportion of PR deaths was > 0.15, the extent of bias in SSC estimates was consistently less than 20%.

[figure 2 about here]

Robustness tests:

In tables A1 and A2, we report results of robustness tests in which we reclassify as PR a subset of deaths previously classified as non-PR by the HDSS. These tests indicate that our study results are robust to potential misclassifications of causes of death in the HDSS data. The observed difference in reporting of adult deaths between DHS respondents in sibships with a PR death vs. sibships with a non-PR death remained significant in 3 out of 4 data accuracy scenarios (figure A1). Similarly, the difference in specificity between the SSC and DHS questionnaires remained significant in all 4 data accuracy scenarios (figure A2). Finally, there were no significant differences in sensitivity between the DHS and SSC questionnaires in any of the 4 data accuracy scenarios.

Discussion

In this study, we compared data collected during SSH interviews to prospective data on adult mortality collected by the Niakhar HDSS since 1962. In doing so, we found that SSH data collected by the DHS questionnaire yielded significant biases: DHS respondents were significantly more likely to report an adult death among their sisters during the SSH when their sister had died from PR causes than when she had died of non-PR causes. In particular, we found that study respondents interviewed with the DHS questionnaire very frequently omitted the deaths of sisters who had died before age 25 from non-PR causes. In addition, DHS respondents also classified as PR a number of deaths that were classified as non-PR by the HDSS. We already observed similar biases during a smaller study we conducted in 2010 in Bandafassi, in southeastern Senegal [28]. Simple calculations indicate that these omissions and misclassifications lead to large over-estimates of the proportion of PR deaths when respondents are interviewed with the DHS questionnaire.

These findings contrast with prior validation studies of SSH data collected using the standard DHS questionnaire. In Bangladesh, for example, a similar comparison of SSH and HDSS data found that SSH data *under*estimated the proportion of PR deaths. Deaths from indirect maternal causes or deaths due to induced abortions were frequently misclassified by SSH data in that setting. The pattern of misreporting may be different in Senegal if abortion-related deaths and deaths from indirect maternal causes are a) less common or b) more frequently known by the siblings of the deceased.

In our study, SSH data collected through a new questionnaire, the siblings' survival calendar (SSC), was not affected by similar biases. Among respondents interviewed with the SSC, the likelihood of reporting an adult death was not associated with the cause of death: non-PR deaths were as likely to be reported as PR deaths. Similarly, data collected through the SSC had much higher specificity than the DHS questionnaire, while maintaining comparable sensitivity. As a result, SSH data collected yielded virtually unbiased estimates of the proportion of PR deaths for settings in which more than 15% of deaths were due to PR causes.

There are several important limitations. First, the HDSS dataset does not constitute a gold standard measure of maternal mortality. Some PR deaths may be classified as non-PR by the HDSS. We have however assessed the robustness of our findings to misclassifications in the HDSS dataset (figures A1 and A2). These tests indicated that differences in reporting patterns between the SSC and the DHS questionnaires remained, even after reclassifying a number of deaths classified as non-PR by the HDSS. Second, some sub-group analyses were limited by small sample sizes. This was the case for example of our analyses by age at death of the deceased. Third, we were only able to trace former residents of the Niakhar HDSS who had migrated to a small subset of accessible localities. If the patterns of SSH reporting are associated with geographic distance between siblings, our estimates of the extent of misreporting in SSH data may be biased.

The generalizability of our findings may also be limited. We found similar SSH reporting patterns using the DHS questionnaire both in Niakhar and Bandafassi, two localities inhabited by different ethnic groups. However, these localities may not be representative of other parts of Senegal, or more broadly western African settings. On the one hand, both localities have been undergoing demographic surveillance for an extended period of time. Individuals may thus be more aware of ages, dates and causes of deaths than respondents in other settings. On the other hand, these are localities with significantly lower educational average than other parts of Senegal. Further validation studies should be conducted in different societal settings. Finally, the underlying causes of PR and non-PR deaths may differ in Niakhar and Bandafassi, relative to other Senegalese or West African communities. In particular, the frequency of induced abortions and deaths from indirect maternal causes may be higher elsewhere, thus leading to different reporting

patterns.

Our results nonetheless have important implications for the estimation of maternal mortality and progress towards MDG5. They suggest that SSH data should not be systematically inflated before being entered into statistical models. Instead, for Senegal and similar countries, SSH data may also need to be adjusted downwards prior to use in maternal mortality estimation.

REFERENCES

1. Zachariah R, Mwagomba B, Misinde D, Mandere BC, Bemeyani A, Ginindza T*, et al.* Vital registration in rural Africa: is there a way forward to report on health targets of the Millennium Development Goals? *Trans R Soc Trop Med Hyg* 2011,**105**:301-309.
2. AbouZahr C, Rampatige R, Lopez A, deSavigny D. When civil registration is inadequate: interim methods for generating vital statistics. *Pac Health Dialog* 2012,**18**:215-230.
3. Hill K, Lopez AD, Shibuya K, Jha P, Monitoring of Vital E. Interim measures for meeting needs for health sector data: births, deaths, and causes of death. *Lancet* 2007,**370**:1726-1735.
4. Naghavi M, Makela S, Foreman K, O'Brien J, Pourmalek F, Lozano R. Algorithms for enhancing public health utility of national causes-of-death data. *Popul Health Metr* 2010,**8**:9.
5. Hernandez B, Ramirez-Villalobos D, Romero M, Gomez S, Atkinson C, Lozano R. Assessing quality of medical death certification: Concordance between gold standard diagnosis and underlying cause of death in selected Mexican hospitals. *Popul Health Metr* 2011,**9**:38.
6. Wilmoth J, Mathers C, Say L, Mills S. Maternal deaths drop by one-third from 1990 to 2008: a United Nations analysis. *Bull World Health Organ* 2010,**88**:718-718A.
7. Wilmoth JR, Mizoguchi N, Oestergaard MZ, Say L, Mathers CD, Zureick-Brown S*, et al.* A New Method for Deriving Global Estimates of Maternal Mortality. *Stat Politics Policy* 2012,**3**.
8. Zureick-Brown S, Newby H, Chou D, Mizoguchi N, Say L, Suzuki E*, et al.* Understanding global trends in maternal mortality. *Int Perspect Sex Reprod Health* 2013,**39**:32-41.
9. Hogan MC, Foreman KJ, Naghavi M, Ahn SY, Wang M, Makela SM*, et al.* Maternal mortality for 181 countries, 1980-2008: a systematic analysis of progress towards Millennium Development Goal 5. *Lancet* 2010,**375**:1609-1623.
10. Lozano R, Wang H, Foreman KJ, Rajaratnam JK, Naghavi M, Marcus JR*, et al.* Progress towards Millennium Development Goals 4 and 5 on maternal and child mortality: an updated systematic analysis. *Lancet* 2011,**378**:1139-1165.
11. Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR*, et al.* Global, regional, and national levels and causes of maternal mortality during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2014,**384**:980-1004.
12. Trends in Maternal Mortality: 1990 to 2013. In. Edited by Organization WH.
13. Zaba B, Paes N. An alternative procedure for fitting relational model life tables. *Genus* 1995,**51**:19-43.
14. Gakidou E, Hogan M, Lopez AD. Adult mortality: time for a reappraisal. *Int J Epidemiol* 2004,**33**:710-717.

15. Rajaratnam JK, Marcus JR, Levin-Rector A, Chalupka AN, Wang H, Dwyer L*, et al.* Worldwide mortality in men and women aged 15-59 years from 1970 to 2010: a systematic analysis. *Lancet* 2010,**375**:1704-1720.
16. Obermeyer Z, Rajaratnam JK, Park CH, Gakidou E, Hogan MC, Lopez AD*, et al.* Measuring adult mortality using sibling survival: a new analytical method and new results for 44 countries, 1974-2006. *PLoS Med* 2010,**7**:e1000260.
17. Wang H, Dwyer-Lindgren L, Lofgren KT, Rajaratnam JK, Marcus JR, Levin-Rector A*, et al.* Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012,**380**:2071-2094.
18. Hosegood V, Vanneste AM, Timaeus IM. Levels and causes of adult mortality in rural South Africa: the impact of AIDS. *AIDS* 2004,**18**:663-671.
19. Hill K, El Arifeen S, Koenig M, Al-Sabir A, Jamil K, Raggers 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**:173-180.
20. Ye Y, Wamukoya M, Ezeh A, Emina JB, Sankoh O. Health and demographic surveillance systems: a step towards full civil registration and vital statistics system in sub-Sahara Africa? *BMC Public Health* 2012,**12**:741.
21. Pison G. Population observatories as sources of information on mortality in developing countries. *Demographic Research* 2005,**13**:301-333.
22. 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**:937-942.
23. Shahidullah M. The sisterhood method of estimating maternal mortality: the Matlab experience. *Stud Fam Plann* 1995,**26**:101-106.
24. Garenne M, Sauerborn R, Nougtara A, Borchert M, Benzler J, Diesfeld J. Direct and indirect estimates of maternal mortality in rural Burkina Faso. *Stud Fam Plann* 1997,**28**:54-61.
25. Ngom P, Akweongo P, Adongo P, Bawah AA, Binka F. Maternal mortality among the Kassena-Nankana of northern Ghana. *Stud Fam Plann* 1999,**30**:142-147.
26. Stanton C, Abderrahim N, Hill K. An assessment of DHS maternal mortality indicators. *Stud Fam Plann* 2000,**31**:111-123.
27. Pison G, Douillot L, Kante AM, Ndiaye O, Diouf PN, Senghor P*, et al.* Health & demographic surveillance system profile: Bandafassi Health and Demographic Surveillance System (Bandafassi HDSS), Senegal. *Int J Epidemiol* 2014,**43**:739-748.
28. Helleringer S, Duthe G, Kante AM, Andro A, Sokhna C, Trape JF*, et al.* Misclassification of pregnancy-related deaths in adult mortality surveys: case study in Senegal. *Trop Med Int Health* 2013,**18**:27-34.
29. Helleringer S, Pison G, Masquelier B, Kante AM, Douillot L, Duthe G*, 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**:e1001652.

30. Delaunay V, Douillot L, Diallo A, Dione D, Trape JF, Medianikov O*, et al.* Profile: the Niakhar Health and Demographic Surveillance System. *Int J Epidemiol* 2013,**42**:1002-1011.
31. Kodio B, de Bernis L, Ba M, Ronsmans C, Pison G, Etard JF. Levels and causes of maternal mortality in Senegal. *Trop Med Int Health* 2002,**7**:499-505.
32. Ba MG, Kodio B, Etard JF. [Verbal autopsy to measure maternal mortality in rural Senegal]. *J Gynecol Obstet Biol Reprod (Paris)* 2003,**32**:728-735.
33. Trape JF, Sauvage C, Ndiaye O, Douillot L, Marra A, Diallo A*, et al.* New malaria-control policies and child mortality in senegal: reaching millennium development goal 4. *J Infect Dis* 2012,**205**:672-679.
34. Ronsmans C, Etard JF, Walraven G, Hoj L, Dumont A, de Bernis L*, et al.* Maternal mortality and access to obstetric services in West Africa. *Trop Med Int Health* 2003,**8**:940-948.
35. Chandramohan D, Rodrigues LC, Maude GH, Hayes RJ. The validity of verbal autopsies for assessing the causes of institutional maternal death. *Stud Fam Plann* 1998,**29**:414-422.
36. Setel PW, Whiting DR, Hemed Y, Chandramohan D, Wolfson LJ, Alberti KG*, et al.* Validity of verbal autopsy procedures for determining cause of death in Tanzania. *Trop Med Int Health* 2006,**11**:681-696.

	Deaths of women aged 15-49 years old recorded by the HDSS over the past 15 years		
	Pregnancy-related deaths (n=38)	Other deaths (n=85)	P-value
Family characteristics			
Family size*			
Number of maternal siblings	6.7 (2.1)	6.2 (2.6)	0.25
Event characteristics			
Time since death*			
Years	7.9 (4.0)	6.9 (4.0)	0.19
Age at death*			
Years	26.2 (7.7)	27.9 (10.2)	0.34
Timing of death			
During pregnancy	8 (21.0)	--	
At delivery	19 (50.0)	--	
Within 42 days of delivery	11 (29.0)	--	
Cause of death**			<0.01
Infectious & parasitic disease	2 (5.3)	29 (34.1)	
Neoplasm	--	6 (7.1)	
Disease of nervous system	--	7 (8.2)	
Disease of the circulatory system	2 (5.3)	12 (14.1)	
Disease of respiratory system	--	2 (2.3)	
Disease of digestive system	--	9 (10.6)	
Disease of genitourinary system	3 (7.9)	3 (3.5)	
Complications of pregnancy, childbirth and the puerperium	22 (57.9)	--	
Injury and poisoning	2 (5.3)	2 (2.4)	
Other cause	5 (13.2)	12 (14.1)	
Non-coded	2 (5.3)	3 (3.5)	
Verbal autopsy (VA) respondent			0.15
Father or mother	11 (29.0)	36 (42.4)	
In-law	7 (18.4)	13 (15.3)	
Sibling	2 (5.3)	6 (7.1)	
Husband or co-wife	5 (13.2)	12 (14.1)	
Other	10 (26.3)	18 (21.2)	
Missing	3 (7.9)	--	

Table 1: Characteristics of deaths recorded by the Health and Demographic Surveillance System in Niakhar and included in the validation study

Notes: causes of death were recorded using ICD-9 classification. P-values were obtained using t-tests (continuous variables) or chi-square tests (categorical variables). * numbers in parentheses are standard deviations; other numbers in parentheses are column percentages; ** we used an exact chi-square test to account for low numbers of observations (n<5) in various cells.

	DHS Questionnaire					SSC Questionnaire				
	Sibships with one pregnancy-related death		Sibships with no pregnancy-related deaths		P-value[a]	Sibships with one pregnancy-related death		Sibships with no pregnancy-related deaths		P-value[a]
	Reported/ expected	% (95% CI)	Reported/ expected	% (95% CI)		Reported/ expected	% (95% CI)	Reported/ expected	% (95% CI)	
All deaths	27/28	96.4 (78.9, 99.5)	39/55	70.9 (56.6, 82.0)	0.007	20/22	90.9 (71.4, 97.6)	50/59	84.8 (70.6, 92.8)	0,469
15-24 years old	15/16	93.8 (66.9, 99.1)	12/25	48.0 (29.9, 66.7)	0.003	7/7	100.0 (59.0, 100)	29/36	80.6 (59.8, 92.0)	0.288
≥ 25 years old	12/12	100.0 (73.5, 100)	27/30	90.0 (73,5, 96.7)	0.251	13/15	86.7 (61.4, 96.4)	21/23	91.3 (70.6, 97.9)	0.637

Table 2: Proportion of respondents reporting an adult death among their sisters, by questionnaire type and sibship composition.

Notes: [a] P-value test the difference in proportion of respondents reporting an adult death between families in which there was one pregnancy-related deaths and other families in which all adult deaths were not pregnancy-related. This p-value is based on a χ^2test. Standard errors are adjusted for the clustering of observations within sibships.

	Sensitivity				
	DHS Questionnaire		SSC Questionnaire		P-value[a]
	Correct/ expected	% (95% CI)	Correct/ expected	% (95% CI)	
All deaths	25/27	92.6 (74.6, 98.2)	18/20	90.0 (69.5, 97.3)	0.744
15-24 years old	15/15	100.0 (78.2, 100)	6/7	85.7 (46.7, 97.6)	0.318
25 years old and above	10/12	83.3 (52.3, 95.8)	12/13	92.3 (61.9, 98.9)	0.593

Table 3: Sensitivity of reports of adult female deaths collected during SSH survey, by type of questionnaire

Notes: sensitivity refers to the proportion of deaths classified as pregnancy-related using the HDSS data, which were also reported as pregnancy-related during the SSH survey. P-values measure differences in sensitivity between the SSC and DHS questionnaires. They were based on a χ^2test of association. Standard errors were adjusted for the clustering of observations within families/sibships, except for reporting by age groups for which we used exact χ^2tests due to limited sample sizes.

	Specificity				
	DHS Questionnaire		SSC Questionnaire		P-value[a]
	Correct/ expected	% (95% CI)	correct/ expected	% (95% CI)	
All deaths	31/39	79.5 (64.1, 89.4)	48/50	96.0 (85.5, 99.0)	0.014
15-24 years old	11/12	91.7 (56.2, 98.9)	29/29	100.0 (88.1, 100)	0.293
25 years old and above	20/27	74.1 (54.5, 87.2)	19/21	90.5 (70.1, 97.5)	0.264

Table 4: Specificity of reports of adult female deaths collected during SSH survey, by type of questionnaire

Notes: Specificity refers to the proportion of deaths that were not pregnancy-related according to the HDSS dataset and that were also reported as not related to pregnancy during the SSH survey. P-values measure differences in specificity between the SSC and DHS questionnaires. They were based on a χ^2test of association. Standard errors were adjusted for the clustering of observations within families/sibships, except for reporting by age groups for which we used exact χ^2tests due to limited sample sizes.

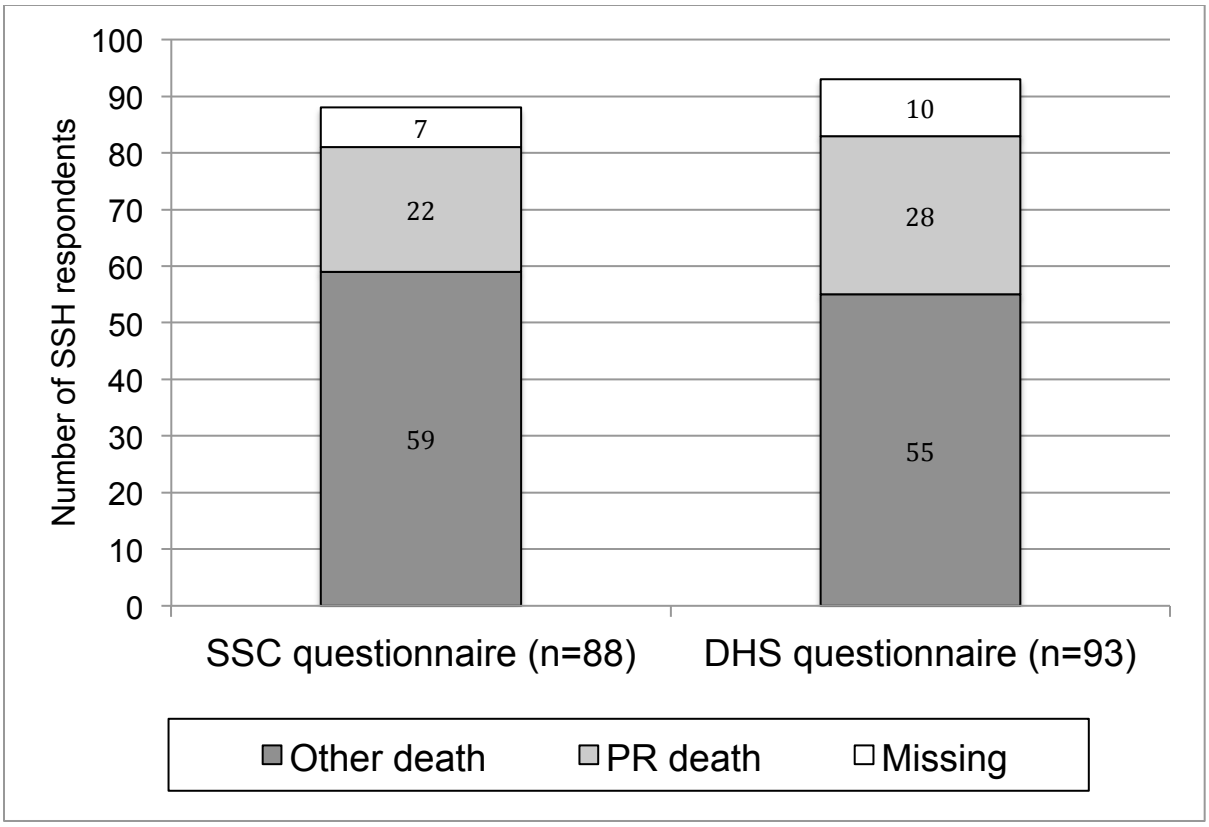

Figure 1: distribution of study respondents by type of questionnaire and cause of death

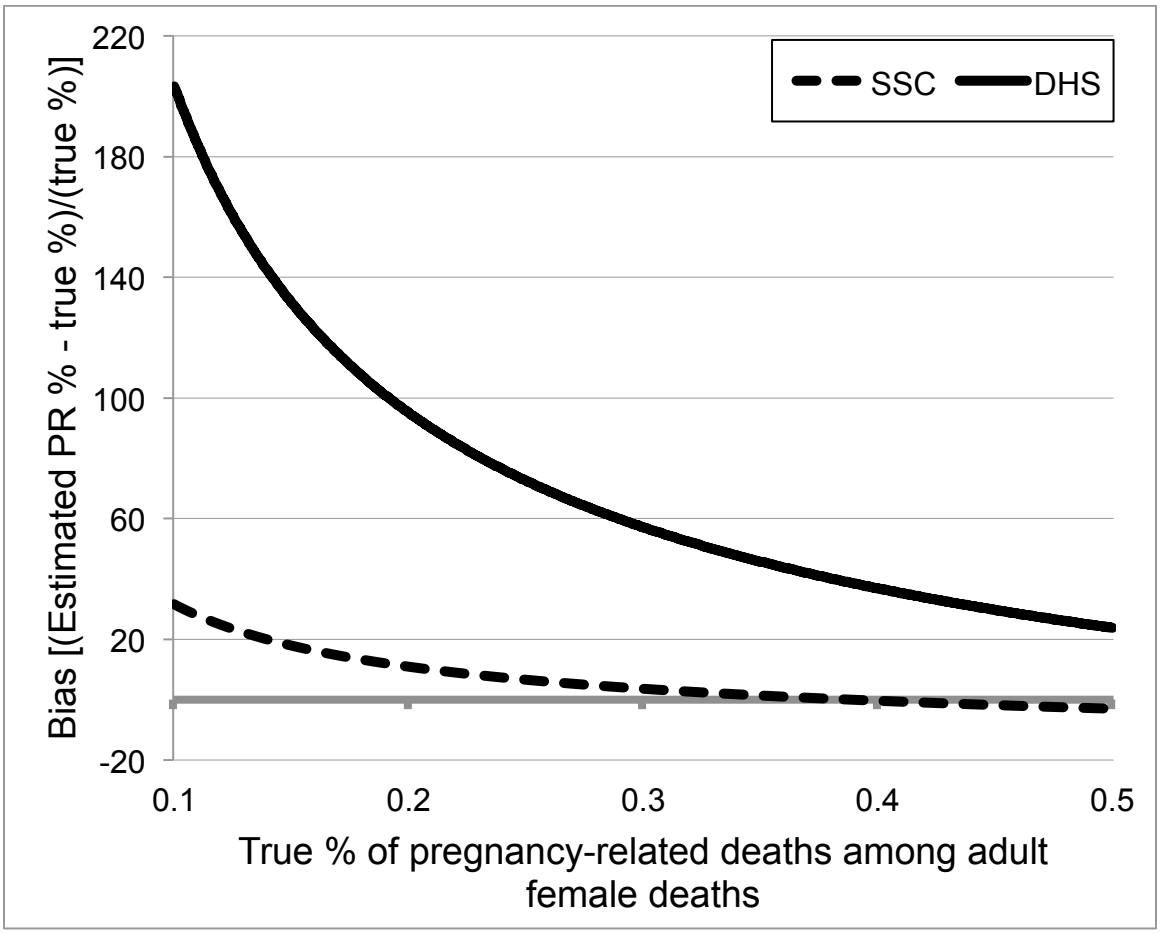

Figure 2: Bias in estimates of the proportion of deaths due to pregnancy-related deaths, by type of questionnaire.

Notes: the extent of bias was calculated on the basis of the formula listed in the text, and using parameters described in tables 2-4.

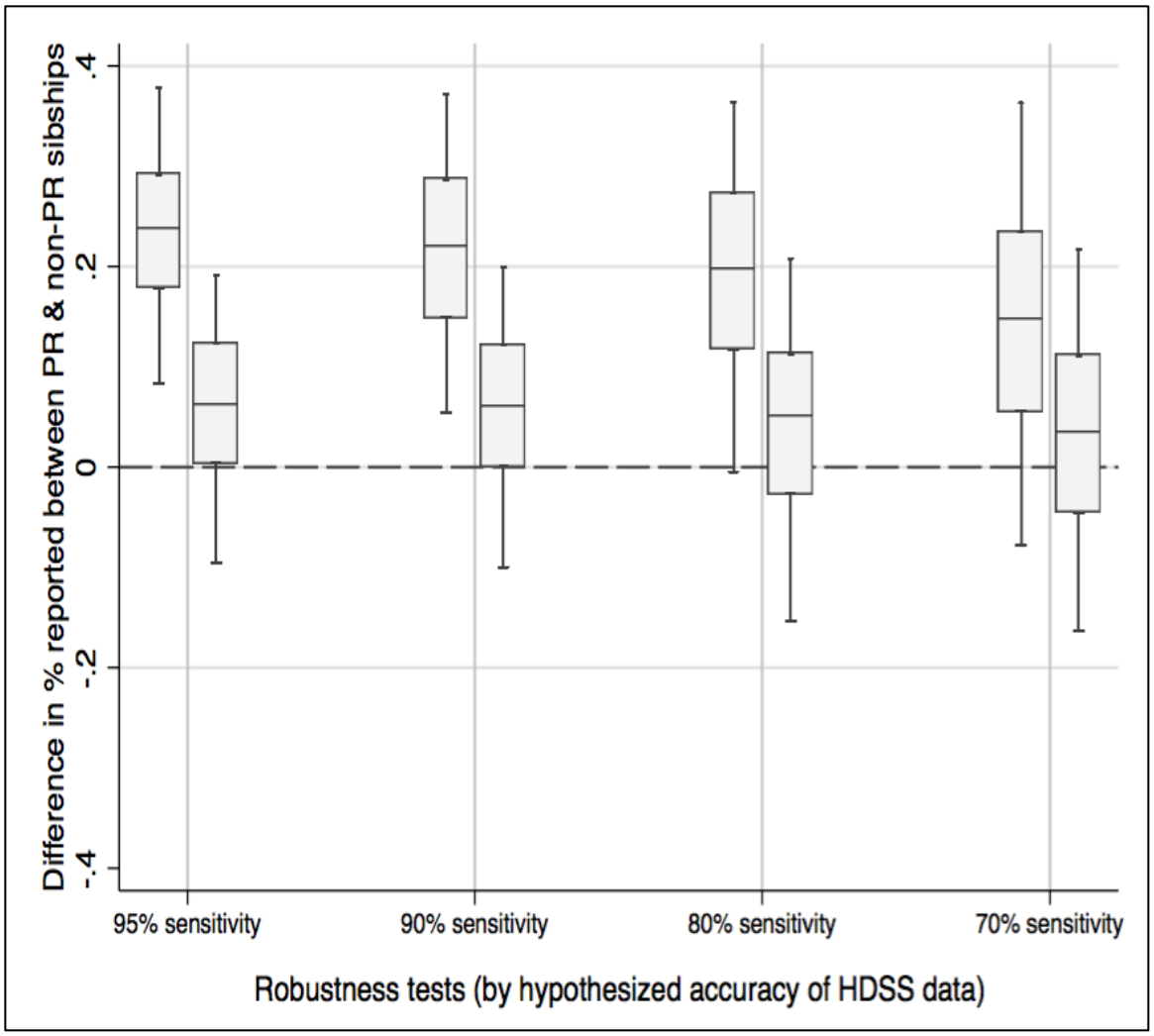

Figure A1: Robustness analyses of the difference in reporting of adult deaths among a respondent's sisters, by type of questionnaire and cause of death.

Notes: the figure tests the robustness of findings reported in table 2, under the heading "all deaths". It shows the difference in proportion of respondents reporting an adult death among their sisters between sibships in which one adult sister died of non pregnancy-related causes and sibships in which adult sister died of pregnancy-related causes according to the HDSS. The robustness test we conduct consists of reclassifying as pregnancy-related a subset of deaths classified as non pregnancy-related by the HDSS. We do so by hypothesizing various sensitivity levels for the HDSS data: 95%, 90%, 80% and 70%. In these scenarios, we reclassify as pregnancy-related 2,4, 9, and 17 deaths previously classified by the HDSS as non-pregnancy related, respectively. We select the deaths to reclassify at random, and we draw 1,000 bootstrap samples from which we recalculate the figures in table 2. The box plots shown in the figure represent the distribution of these bootstrap samples: the contours of the box represent the interquartile range, the middle line is the median and the whiskers represent the 2.5^{th} and 97.5^{th} percentiles. As a result, if the lower whisker is above 0, it indicates that the difference in reporting of adult deaths between families with non-pregnancy related deaths vs. families with pregnancy-related deaths is significant at the $p<0.05$ level. In each robustness test, the leftmost box plot represent values obtained for the DHS questionnaire, whereas the rightmost box plot represents values obtained for the SSC questionnaire.

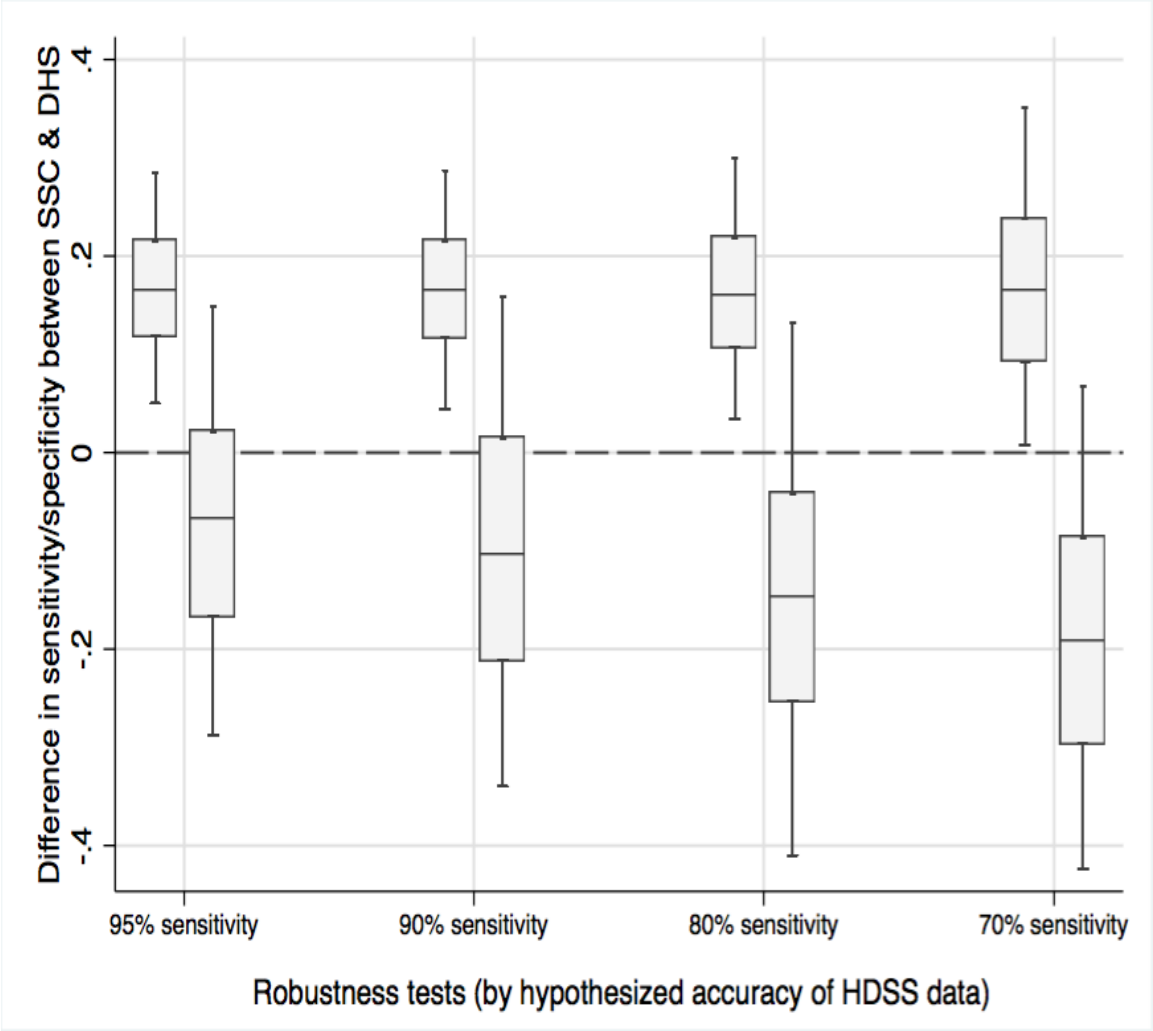

Figure A2: Robustness analyses of the difference in sensitivity/specificity between the DHS and SSC questionnaires.

Notes: the figure tests the robustness of findings reported in tables 3 and 4, under the heading "all deaths". It shows the difference in sensitivity/specificity between the SSC and DHS questionnaires. The robustness test we conduct consists of reclassifying as pregnancy-related a subset of deaths classified as non pregnancy-related by the HDSS. We do so by hypothesizing various sensitivity levels for the HDSS data: 95%, 90%, 80% and 70%. In these scenarios, we reclassify as pregnancy-related 2,4, 9 and 17 deaths previously classified by the HDSS as non-pregnancy related, respectively. We select the deaths to reclassify at random, and we draw 1,000 bootstrap samples from which we recalculate the figures in tables 3 and 4. The box plots shown in the figure represent the distribution of these bootstrap samples: the contours of the box represent the interquartile range, the middle line is the median and the whiskers represent the 2.5^{th} and 97.5^{th} percentiles. As a result, if the lower whisker is above 0, it indicates that the difference in sensitivity/specific between the SSC and the DHS questionnaires is significant at the $p<0.05$ level. In each robustness test, the leftmost box plot represents differences in specificity values between the SSC and the DHS, whereas the rightmost box plot represents differences in sensitivity values between the SSC and the DHS.