International Journal of Population Data Science

Journal Website: www.ijpds.org

Evaluating pregnancy reporting in Siaya Health and Demographic Surveillance System through record linkage with ANC clinics

Hallie Eilerts-Spinelli[1,*], Julio Romero Prieto[1], Julie Ambia[2], Sammy Khagayi[2], Chodziwadziwa Kabudula[3], Jeffrey W. Eaton[4], and Georges Reniers[1]

Submission History

Submitted:	31/05/2022
Accepted:	28/07/2022
Published:	29/09/2022

[1]Population Studies Group, Department of Population Health, London School of Hygiene & Tropical Medicine, London, UK
[2]Kenya Medical Research Institute-Center for Global Health Research (KEMRI-CGHR), Kisumu, Kenya
[3]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[4]MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK

Abstract

Introduction
Health and Demographic Surveillance Systems (HDSS) are important sources of population health data in sub-Saharan Africa, but the recording of pregnancies, pregnancy outcomes, and early mortality is often incomplete.

Objective
This study assessed HDSS pregnancy reporting completeness and identified predictors of unreported pregnancies that likely ended in adverse outcomes.

Methods
The analysis utilized individually-linked HDSS and antenatal care (ANC) data from Siaya, Kenya for pregnancies in 2018-2020. We cross-checked ANC records with HDSS pregnancy registrations and outcomes. Pregnancies observed in the ANC that were missing reports in the HDSS despite a data collection round following the expected delivery date were identified as likely adverse outcomes, and we investigated the characteristics of such individuals. Clinical data were used to investigate the timing of HDSS pregnancy registration relative to care seeking and gestational age, and examine misclassification of miscarriages and stillbirths.

Results
From an analytical sample of 2,475 pregnancies observed in the ANC registers, 46% had pregnancy registrations in the HDSS, and 89% had retrospectively reported pregnancy outcomes. 1% of registered pregnancies were missing outcomes, compared to 10% of those lacking registration. Registered pregnancies had higher rates of stillbirth and perinatal mortality than those lacking registration. In 77% of cases, women accessed ANC prior to registering the pregnancy in the HDSS. Half of reported miscarriages were misclassified stillbirths. We identified 141 unreported pregnancies that likely ended in adverse outcomes. Such cases were more common among those who visited ANC clinics during the first trimester, made fewer overall visits, were HIV-positive, and outside of formal union.

Conclusions
Record linkage with ANC clinics revealed pregnancy underreporting in HDSS, resulting in biased measurement of perinatal mortality. Integrating records of ANC usage into routine data collection can augment HDSS pregnancy surveillance and improve monitoring of adverse pregnancy outcomes and early mortality.

Keywords
sub-Saharan Africa; HDSS; pregnancy; stillbirth; neonatal mortality; antenatal care; record linkage

*Corresponding Author:
Email Address: hceilerts@gmail.com (Hallie Eilerts-Spinelli)

https://doi.org/10.23889/ijpds.v7i4.1762

Introduction

There have been substantial improvements in maternal and newborn health since the beginning of global monitoring, however stark regional disparities persist. Sub-Saharan Africa is the region with the highest burden of stillbirths and early mortality [1, 2]. The United Nations Sustainable Development Goals call for reducing neonatal mortality (deaths within the first 28 days of life) to 12 deaths per 1,000 live births by 2030 [3]. While the reduction of stillbirths has not been a focus of international policy and investment agendas [4], visibility has increased in recent years. The Early Newborn Action Plan, launched through a World Health Assembly resolution in 2014, established the target of reducing the stillbirth rate to 10 per 1,000 total births in all countries by 2035 [5]. Accurate measurement is seen as key to accelerating progress towards both targets. Health management information systems (HMIS) or civil registration and vital statistics (CRVS) systems are the ideal sources for this information, but they are incomplete or non-existent in most low- and lower-middle-income countries [6].

Nationally representative household surveys are the primary sources of data on pregnancy loss and early mortality in sub-Saharan Africa [7]. The most prominent among these, the Demographic and Health Surveys (DHS), have been instrumental to monitoring maternal and newborn health. However, the omission of stillbirths and early deaths is a serious limitation, which causes downward bias in mortality estimates [8, 9].[1] There are persistent challenges associated with collecting survey data on such topics due to their sensitive nature [7]. Qualitative research in rural Tanzania and Uganda found that strong social norms discouraged adverse pregnancy outcomes such as miscarriage and stillbirth from being discussed publicly [12, 13]. The use of male interviewers has been found to negatively impact data quality and completeness, especially for such topics as pregnancy and childbirth [11, 14]. Furthermore, distinguishing between reported instances of miscarriage and stillbirth can be impacted by survey translation, and overlap in the local language terms for such events [7, 12]. It is worth noting that these classification issues are not unique to lower-income settings. A lack of uniformity in benchmarks of gestational age, birthweight, and body length used to distinguish between stillbirth and miscarriage has complicated global monitoring efforts [4, 15].

Health and Demographic Surveillance Systems (HDSS) are important sources of longitudinal population health data throughout sub-Saharan Africa. HDSS sites routinely collect information on demographic events such as births, deaths, and migrations through recurring household interviews of contiguous populations [16]. Though this prospective data collection process is highly effective at tracking the vital status of established residents, it is less reliable for collecting information on pregnancy outcomes and newborns. In the case that a pregnancy ends in a live birth, the newborn will likely be enumerated in the HDSS at the subsequent household interview round. However, pregnancies resulting in adverse outcomes such as miscarriage, stillbirth, and early deaths (before the next household enumeration) are subject to underreporting [16].

Some HDSS sites register ongoing pregnancies to facilitate follow-up on pregnancy outcomes and early mortality [17]. We previously found that mortality during the first year of life was higher for cohorts of births with registered pregnancies compared to those where the pregnancy was not observed, implying omission of early deaths from the latter group [18]. While pregnancy registration appears to be an important tool for improving measurement of early mortality, it is often incomplete in HDSS sites [19]. Furthermore, little is known about the number of pregnancies that are entirely missing from HDSS records, lacking both pregnancy registration and outcome reports. Given indications that pregnancies ending in adverse outcomes are vulnerable to underreporting [19–21], missing pregnancy reports are a likely source of downward bias in HDSS measurement of stillbirths and neonatal mortality. As HDSS are often deliberately set up in locations where there is limited availability of other population-based data [16], external validation of mortality estimates is difficult.

Record linkage with antenatal care (ANC) data is a promising avenue for improving HDSS pregnancy data, and can help address the data gap on pregnancy loss and early mortality in sub-Saharan Africa. The World Health Organization recommends that all pregnant women have at least eight ANC assessments [22]. While only 52% of women in sub-Saharan Africa make four or more visits, close to 90% of women seek ANC services at least once during pregnancy [23]. Integrating records of ANC usage into HDSS data could reduce frequency of surveillance rounds required to comprehensively capture pregnancies and births. ANC data also have more reliable information on the underlying conditions of mothers affecting the survival of newborns. Record linkage has been widely practised to support epidemiological research and health services evaluation in high-income settings [24–29]. It has been less common in sub-Saharan Africa, but the field is growing, with several studies demonstrating its feasibility [30–32].

In this work, we leveraged linked HDSS and ANC data to shed light on pregnancy reporting and potential downward bias in measurement of stillbirths, perinatal, and neonatal mortality in an HDSS in western Kenya. This work provides a framework for how individually-linked HDSS and ANC records can be used to complement HDSS data and improve information collected on pregnancies and their outcomes in settings lacking adequate HMIS and CRVS.

Methods

Data

The study area is located in Siaya County in western Kenya. The Siaya HDSS was established in 2001 as a collaboration between the Kenya Medical Research Institute (KEMRI) and the Centers for Disease Control [33]. The site includes the rural communities of Karemo, Asembo, and Gem; covering a total area of approximately 700 km^2 with around 224,000 residents in 2020. Data collection has been conducted through household interviews every four months up to 2015, and every six months thereafter. A household proxy respondent reports

[1] The DHS recently updated their core reproductive history questionnaire to address data quality issues related to information collected on pregnancy loss and early mortality [10, 11].

to the HDSS fieldworker on behalf of all household members, providing information on births, deaths, and migrations which have occurred since the previous data collection round. Information on pregnancy status is collected from women of reproductive age directly, though a proxy respondent may be used if the individual is not present at the time of the interview. In the case that a pregnancy is registered, the HDSS fieldworker is prompted to follow up on its outcome at subsequent data collection rounds. Local community reporters have also been trained to collect data on births and deaths in their villages. HDSS fieldworkers refer to these records during data collection to assure data completeness. More information on the data collection protocols of the HDSS is available elsewhere [33].

Beginning in February 2018, the HDSS initiated Point-of-contact Interactive Record Linkage (PIRL) with 14 ANC clinics in the Gem District of Siaya County. We will briefly describe the record linkage process in Siaya, and more detail on the PIRL approach is available elsewhere [32, 34, 35]. Data clerks stationed in clinic waiting rooms invited pregnant women aged 13 or older who were seeking ANC to participate in the record linkage study. After obtaining written informed consent, women were enrolled in the study, and PIRL was attempted for those who self-reported residence in the HDSS. The data clerk collected information on up to three names for the individual, date of birth, location of household in the HDSS, and up to three names of another household member. This information was entered into the PIRL software to search for the individual in the HDSS database. Using the probabilistic framework developed by Fellegi and Sunter [36], HDSS and ANC record pairs were compared using a series of identifiers. Each comparison contributed an agreement or disagreement weight towards the total match score of the record pair, with weights calculated as a ratio of the probabilities that true and false matches agreed on the identifier [37, 38]. The match probabilities used in the PIRL software were adapted from a previous study which conducted record linkage between ANC clinics and an HDSS in rural Tanzania [32]. The software returned the 20 highest scoring potential matches from the HDSS ranked in descending order, and the data clerk then consulted the subject to identify the matching record.

Analysis

Sample of linked records

This analysis pertains to individuals who visited ANC clinics after the start of record linkage on February 7, 2018 and had an EDD prior to October 1, 2020. At the time of this analysis, the most recent HDSS data was collected between October and December 2020. The upper bound for EDD thus ensured that at least one HDSS data collection round occurred following the expected completion of all pregnancies included in the sample. For women making their first ANC visit for a given pregnancy during the first trimester, the EDD recorded in the clinic data was typically estimated as 40 weeks after their last menstrual period (LMP). For those with a later gestational age at first visit, the EDD was either estimated from the LMP or by the clinic nurse through fundal height palpation. In cases where the EDD was updated at subsequent ANC visits, we used the latest recorded EDD to evaluate the sample inclusion criteria.

Individuals who were not residing in the HDSS at the time of ANC linkage (as per HDSS records) were excluded from the analysis. As PIRL was only attempted for those who self-reported as current residents of the HDSS, it is possible that the individual incorrectly reported their residency status or that their HDSS residency record was not up-to-date. Alternatively, it could also be the case that the individual had not resided in the HDSS long enough to be considered a permanent resident. New arrivals to the HDSS are subject to a four-month preliminary registration period. Depending on the timing of the individual's in-migration relative to the HDSS data collection rounds, their status as a permanent resident may not be confirmed for several months more. HDSS pregnancy reporting could not be evaluated for such individuals.

External validation of HDSS pregnancy reporting

For linked individuals, we applied a series of deterministic rules to assess whether pregnancy registrations and outcomes in the HDSS pertained to the same pregnancy observed in the ANC records. ANC pregnancies were matched to HDSS records if (i) the registration occurred between the individual's LMP and EDD, and if there was an outcome, it was no more than 36 weeks after the first ANC visit, (ii) the outcome was between 8 weeks before and 20 weeks after the EDD, and the total duration from pregnancy registration or first ANC visit (whichever was earliest) to outcome was no more than 36 weeks, (iii) the outcome was between the first and last ANC visit for a given pregnancy, and if there was a pregnancy registration, the total duration between pregnancy registration and the last ANC visit was no more than 36 weeks.

Some of the matches meeting the above criteria had HDSS pregnancy outcomes which predated ANC visits for a given pregnancy. This was likely due to pregnancy outcome date misreporting. Reported dates of pregnancy outcomes may be subject to recall errors and rounding, especially if the outcome occurred early on in the interval between data collection rounds, and several months passed before the next household interview. Dates of pregnancy registrations, on the other hand, are more reliable as they simply indicate whether the individual in question was pregnant on the date of the household interview. As such, the matching criteria were set to be sufficiently lenient to allow for some date reporting errors in pregnancy outcomes, but not pregnancy registrations.

We calculated the proportions of pregnancies observed in the ANC clinics that were matched to pregnancy registrations and pregnancy outcomes in the HDSS. For those matched to pregnancy outcomes, we calculated the rates of stillbirth, perinatal, and neonatal mortality by pregnancy registration status, and for all pregnancies combined. Perinatal mortality compromises stillbirths and deaths occurring within the first week of life (i.e. early neonatal period). The denominators for the rates of stillbirth and perinatal mortality were total births. For neonatal mortality, deaths and exposure time were aggregated for the first 28 days of life. The abridged life table mortality rate was converted to the probability of dying by exact age 28 days, known as *q(28d)*. In some cases, a pregnancy was reported to have resulted in a live birth, but the child was not enumerated in the HDSS. The status of such newborns was considered unknown, and they were excluded from calculations of perinatal and neonatal mortality.

For pregnancies observed in the ANC records that were matched to HDSS pregnancy registrations, we assessed the timing of registration relative to care seeking and gestational age. Among those matched to a miscarriage or stillbirth in the HDSS, we evaluated potential misclassification between the two using a threshold of 28 weeks gestation, consistent with the International Classification of Diseases 10 system [39]. HDSS pregnancy outcomes were also used to examine uncertainty in ANC estimates of gestational age. We compared the ANC clinic estimates of gestational age against that which could be inferred from the date of the pregnancy outcome, under the assumption that the pregnancy came to term at 40 weeks gestation.

Characteristics associated with missing pregnancies

For linked cases that were missing pregnancy outcomes in the HDSS, we investigated whether the individual was residing in the site for the data collection round immediately following their EDD. We identified individuals that had out-migrated, died, or been lost to follow-up prior to this data collection round, and those whose households had not yet been visited for data collection. These missing outcomes were distinguished from cases where the individual continued to reside in the site, and had no pregnancy outcome despite the occurrence of a household interview following their EDD. This was considered evidence of a potential unreported adverse pregnancy outcome. We investigated the characteristics of such individuals using Chi-squared tests and logistic regression.

Covariates of interest from the HDSS included age, area of residence, duration of residency, education level, household wealth, and marital status. We generated a household wealth index from a principal component analysis of variables denoting socio-economic status including type of toilet facility, water supply source, and ownership of assets such as a radio, television, and cooking stove [40]. From the clinical data, we included information from the ANC register on total number of ANC visits for the given pregnancy (taken from a pregnancy-specific visit order number), gestational age at first visit, parity, and HIV status. Variables strongly associated with having a missing pregnancy outcome in Chi-squared tests were added to multivariable logistic regression models. Variables were removed from the model if they did not contribute significantly to model fit, as measured by Akaike Information Criterion. Age, household wealth quintile, number of ANC visits, and parity were tested as both categorical and continuous variables.

Results

The ANC dataset consisted of 6,626 individuals who had visited clinics since the start of record linkage in February 2018 to December 2020 (flowchart: Figure 1). There were 5,794 individuals that self-reported to reside in the HDSS area. From this group, 3,173 were linked to their record in the HDSS through PIRL, resulting in an overall match rate of 54.8%. Matches were independently verified by three HDSS personnel through manual review, and no false matches were identified.

The match rate was highest (1,162/1,867; 62.2%) for individuals who were enrolled in the study between February and December 2018. The match rate was 52.9% (1,497/2,830) for those enrolled from January 2019 to March 2020. Linkage activities were interrupted in the third week of March 2020 due to the COVID-19 pandemic. Once linkage was resumed in September 2020, the match rate was 46.9% (514/1,097) for those enrolled prior to December 2020. Women who were younger, HIV-negative, in union, and had given birth to fewer children were less likely to be matched to the HDSS. Those residing in Gem were more likely to be matched than residents of Asembo and Karemo. More details on the match rates stratified by socio-demographic characteristics are provided in Supplementary Table 1.

External validation of HDSS pregnancy reporting

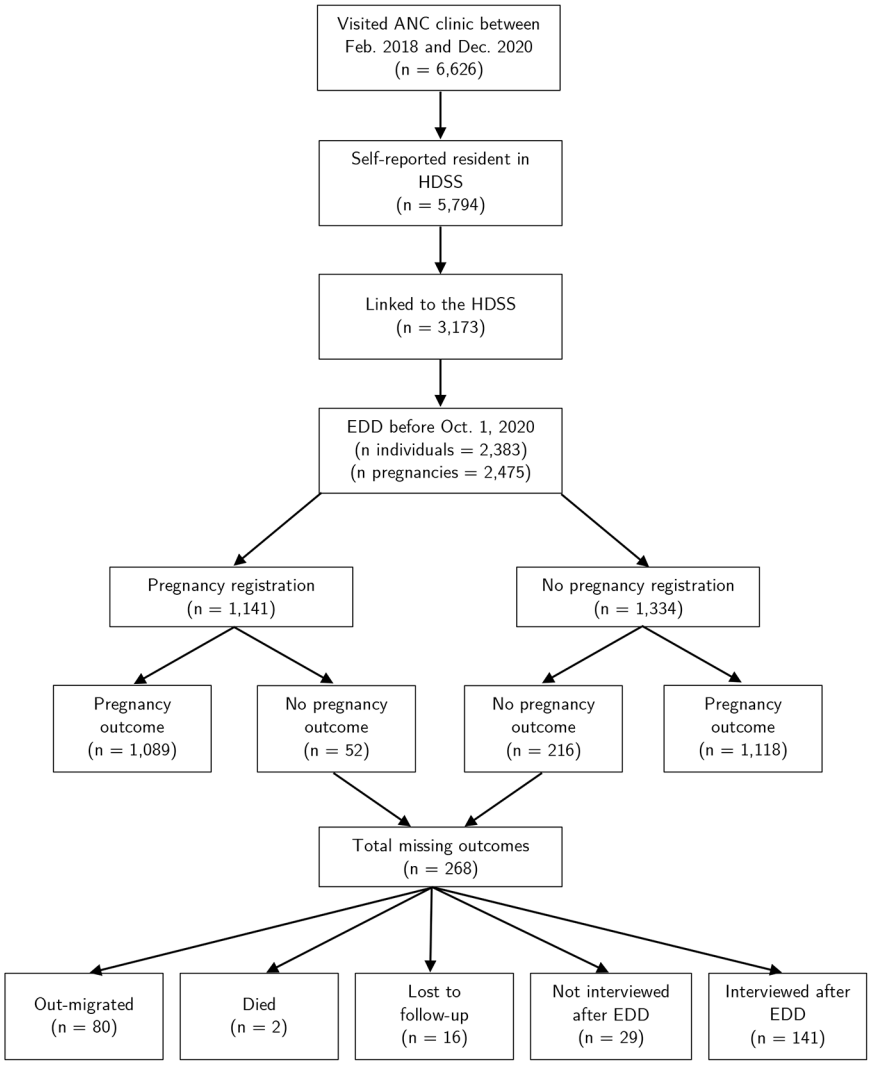
Among the individuals with linked HDSS-ANC records, there were 2,475 pregnancies with EDDs prior to October 1, 2020, attributed to 2,383 individuals. From this total, there were 1,141 (46.1%) pregnancies that were registered in the HDSS. Among those that were registered, 1,089 (95.4%) also had reported pregnancy outcomes. There were 1,334 (53.9%) pregnancies that were not registered in the HDSS. In this group, missing outcomes were more common (n = 216; 16.2%). In total, 268 pregnancies were missing outcome reports in the HDSS.

For pregnancies with missing outcomes, we investigated the residency status of the women in question in the HDSS as of the first household interview following their EDD. In 29.9% (n = 80) of cases, the individual had out-migrated from the HDSS area prior to the occurrence of a post-delivery data collection round. There were two individuals (0.01%) who died prior to the occurrence of a post-delivery data collection round and 16 (6.0%) who were lost to follow-up. For 29 (10.8%) cases, the individual's household had not been visited in a routine HDSS interview round following their EDD. For the largest share (n = 141; 52.6%), the individual was residing in the HDSS during the first household interview following their EDD, and no outcome was reported.

HDSS reporting for pregnancies observed in the ANC clinics is shown in greater detail in Table 1. After excluding cases where a woman's missing pregnancy outcome was attributable to out-migration, death, censoring, or the lack of a subsequent household interview; the remaining missing outcomes consisted of those identified as likely cases of unreported adverse pregnancy outcomes. The share of such outcomes was smaller for pregnancies that were registered in the HDSS (n = 11; 1%) compared to those that were not (n = 130; 10.4%), and made up 6% (n = 141) of the total analytical sample.

Reporting of adverse pregnancy outcomes was more common for pregnancies that had been registered in the HDSS. There were three miscarriages and 14 stillbirths among registered pregnancies. This yielded a stillbirth rate of 12.6 per 1,000 total births (95% confidence interval [CI] 6.3–18.9) and perinatal mortality rate of 27.1 (95% CI 18.1–37.1). There was one miscarriage and no reported stillbirths among pregnancies lacking registration, which had a perinatal mortality rate of 14.0 (95% CI 7.9–21.0). Neonatal mortality for registered pregnancies was 18.4 per 1,000 (95% CI 11.0–25.8), compared

Figure 1: Flowchart of the sample included in the analysis



The flowchart displays the number of pregnancies observed in the antenatal care (ANC) clinics that had pregnancy registrations and outcome reports in Siaya Health and Demographic Surveillance System (HDSS). For pregnancies that were missing outcomes in the HDSS, information is provided on the status of the individual in question in the HDSS as of the first data collection round that occurred following their expected delivery date (EDD).

to 20.2 (95% CI 12.3–29.2) for those that were not registered. In both samples, close to three-quarters of neonatal deaths took place in the first week of life.

For pregnancies in the ANC records that were registered in the HDSS, Figure 2A displays the timing of registration relative to the individual's gestational age. The median gestational age at pregnancy registration in the HDSS was 26.4 weeks, and half of all pregnancy registrations took place between 20 and 33 weeks gestation. For 19 cases (1.7%), pregnancy registration was estimated to have taken place at implausible gestational ages, such as negative values or at more than 42 weeks gestation. These unrealistic values were indicative of error in the ANC clinic estimates of gestational age. Figure 2B shows the timing of pregnancy registration relative to the individual's first ANC clinic visit for the given pregnancy. There were 446 pregnancies that were excluded from this assessment due to the individual's first ANC visit preceding the start of the study or there being missing data for ANC visit number. For the remaining 695 cases, 76.7% ($n = 533$) received ANC prior to registering the pregnancy in the HDSS. Close to 62.2% ($n = 432$) of pregnancy registrations took place in the 16 weeks following the first ANC visit, and 14.5% ($n = 101$) were registered more than 16 weeks later.

Figure 3 shows the results of the investigation into misclassification of miscarriages and stillbirths in the HDSS. Among the reported stillbirths, 92.9% ($n = 13$) had gestational ages of >28 weeks, indicating accurate classification. There were fewer reported miscarriages ($n = 4$), though only half occurred prior to 28 weeks gestation. The other two reported miscarriages appear to have been stillbirths that were incorrectly classified.

Gestational age in the ANC records tended to correspond fairly well with reported dates of HDSS pregnancy outcomes. In Figure 4A, we compared gestational ages recorded in the ANC data to that which could be inferred from pregnancy end dates in the HDSS. For the first ANC visit for a given pregnancy, the median ANC estimate for gestational age was 1.9 weeks lower than inferred gestational age (interquartile

Table 1: Distribution of HDSS pregnancy reporting for women with linked ANC records

	Pregnancy registration		
	Yes	No	All
	n (%)	n (%)	n (%)
Pregnancy outcome reporting			
Missing - potential adverse outcome	11 (1.0)	130 (10.4)	141 (6.0)
Reported	1,089 (99.0)	1,118 (89.6)	2,207 (94.0)
Total	1,100 (100.0)	1,248 (100.0)	2,348 (100.0)
Pregnancy outcomes			
Live birth	1,097[a] (98.4)	1,144[b] (99.9)	2,241 (99.2)
Miscarriage	3 (0.3)	1 (0.1)	4 (0.2)
Stillbirth	14[c] (1.3)	0 (0.0)	14 (0.6)
Total	1,114 (100.0)	1,145 (100.0)	2,259 (100.0)
Neonate status			
Early neonatal death	16 (1.5)	16 (1.4)	32 (1.4)
Late neonatal death	4 (0.4)	7 (0.6)	11 (0.5)
Unknown	9 (0.8)	12 (1.0)	21 (0.9)
Survived	1,068 (97.4)	1,109 (96.9)	2,177 (97.1)
Total	1,097 (100.0)	1,144 (100.0)	2,241 (100.0)
Mortality estimates	*Est. (95% CI)*	*Est. (95% CI)*	*Est. (95% CI)*
Stillbirth rate	12.6 (6.3, 18.9)	0.0 (0.0, 0.0)	6.2 (3.1, 9.8)
Perinatal mortality rate	27.1 (18.1, 37.1)	14.0 (7.9, 21.0)	20.4 (14.7, 26.2)
q(28d)	18.4 (11.0, 25.8)	20.2 (12.3, 29.2)	19.3 (13.9, 25.2)

Notes: Bootstrap 95% confidence intervals (CI) calculated from 10,000 samples with replacement. Stillbirth and perinatal mortality rates shown per 1,000 births, q(28d) shown per 1,000 live births.
[a]Includes 49 multiple births.
[b]Includes 52 multiple births.
[c]Includes 1 multiple birth.

range [IQR] = 5.3). This decreased to a median difference of 1.6 (IQR = 4.1) for the fifth ANC visit, and 0.3 (IQR = 2.6) for the eighth visit (n = 25 pregnancies). In Figure 4B, the ANC record gestational age was plotted against the number of weeks to the HDSS pregnancy end date. A linear regression was fit to the data, showing that with each real increase of one week, the recorded gestational age in the ANC register increased by an average of approximately 0.85 weeks (95% CI 0.83–0.85).

Characteristics associated with missing pregnancies

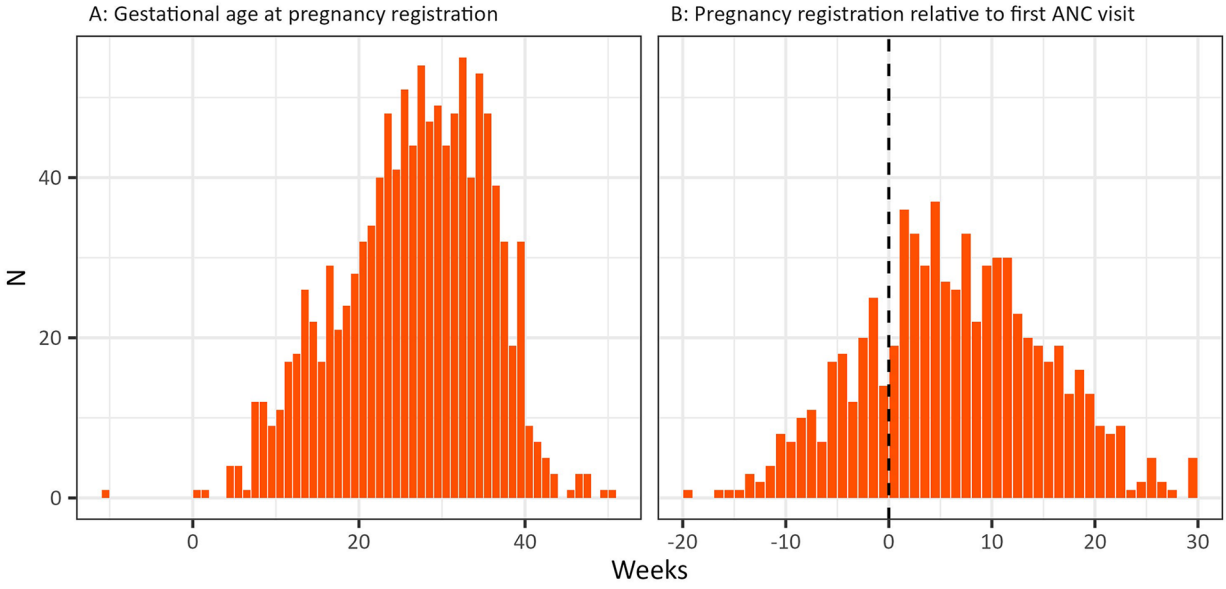
Table 2 presents descriptive statistics for women without a recorded pregnancy outcome in the HDSS, despite their continued residence during a household interview occurring after their EDD. The variables for number of ANC clinic visits, age, and parity were strongly associated with having a missing pregnancy outcome. Underreporting was more common among those making a single as opposed to multiple ANC visits. Women who were in the youngest and oldest age groups also had proportionately more missing outcomes than those aged 20–34 years. Women who had no previous births were more likely to have missing outcomes, as well as those for whom parity was not known.

The variables of gestational age at first ANC visit, HIV status, and household wealth had moderate to weak associations with HDSS pregnancy outcome reporting. Those making their first ANC visit during the first trimester were more likely to be missing pregnancy outcomes in the HDSS. The same was true for women who were HIV-positive or of unknown status. Women belonging to households in the top three wealth quintiles had a higher proportion of missing pregnancy outcome reports than those in the bottom quintiles, though there was not a clear pattern across quintiles. Duration of residency in the HDSS, education, and region of residency within the HDSS were not significantly associated with having a missing pregnancy outcome.

Characteristics associated with HDSS pregnancy outcome reporting were investigated further using logistic regression. The results of the full (1) and final (2) models are shown in Table 3. Controlling for all other variables in the full model, age was not significantly associated with having a missing pregnancy outcome, and was excluded from the final model. In the case of parity, women with an unknown number of previous births had 5.82 (95% CI 1.84–17.37) times the odds of HDSS pregnancy outcome underreporting compared to those with one previous birth. There were not significant differences in the odds of pregnancy outcome reporting for individuals of known parity, and this variable was also left out of the final model.

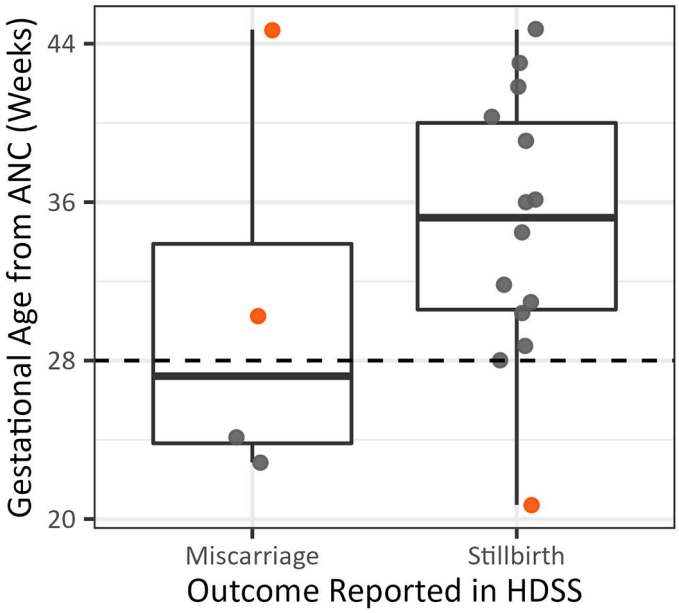
In the final model, a one visit increase in an individual's total number of ANC visits for a given pregnancy was associated with 30% lower odds of pregnancy outcome

Figure 2: Timing of pregnancy registration in HDSS relative to gestational age and seeking antenatal care (ANC)



(A) Histogram of gestational age at the time of HDSS pregnancy registration (n = 1,141). Gestational age was calculated from the ANC register using the value from the individual's latest recorded ANC clinic visit. (B) Histogram of number of weeks between HDSS pregnancy registration and the individual's first ANC clinic visit for the pregnancy. Pregnancies where the first ANC visit preceded the start of PIRL or those missing information for visit number were excluded (n = 446). Negative values indicate that the pregnancy was registered in the HDSS prior to the individual's first ANC clinic visit (represented by the vertical dashed line), while positive values show the reverse.

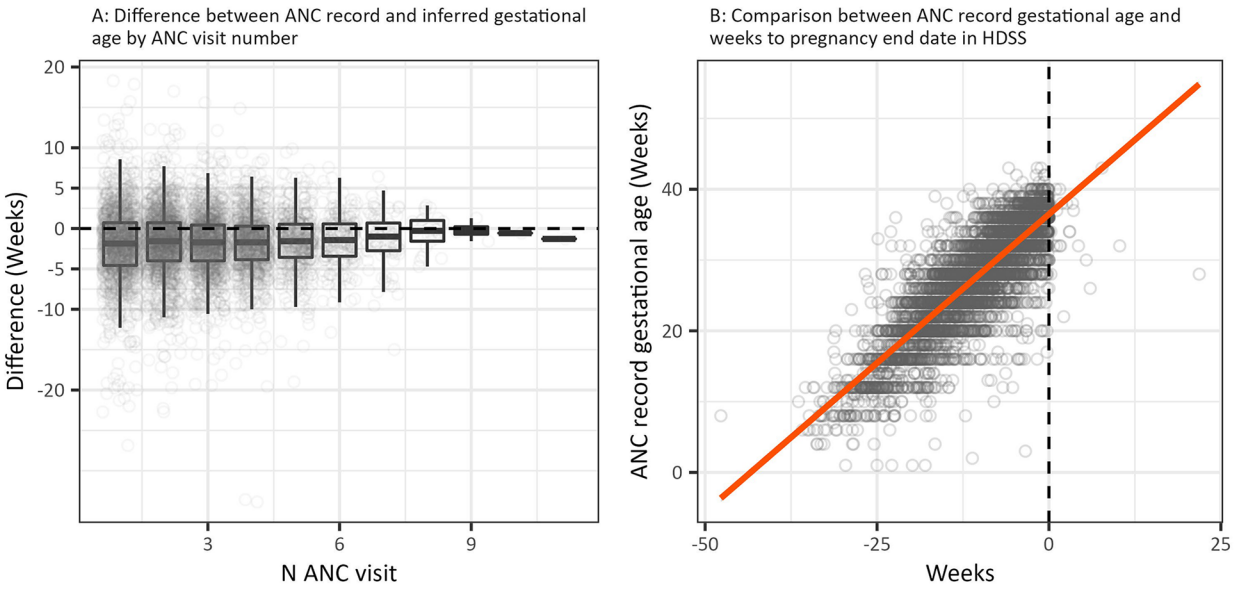
Figure 3: Gestational age for adverse pregnancy outcomes reported in the Health and Demographic Surveillance System (HDSS)



Gestational age at the date of the pregnancy outcome was calculated from the latest recorded value in the ANC data. ANC-estimated gestational age has been plotted separately for pregnancies reported as miscarriages and stillbirths in the HDSS. The horizontal dashed line indicates the 28 week gestational age threshold used to distinguish between miscarriage and stillbirth.

underreporting (odds ratio [OR] 0.70, 95% CI 0.60–0.80). Additionally, women making their first ANC visit during the first trimester had 2.64 (95% CI 1.52–4.51) times the odds of pregnancy outcome underreporting compared to those making their first visit in the second trimester. Those testing HIV-positive during an ANC clinic visit had 86% higher odds of having a missing pregnancy outcome in the HDSS compared to those that tested HIV-negative (OR 1.86, 95% CI 1.22–2.78). Compared to those in the median household wealth quintile,

Figure 4: Comparison of gestational age fromantenatal care (ANC) register and gestational age inferred from the pregnancy end date in the Health and Demographic Surveillance System (HDSS)



(A) Difference between the gestational age recorded in ANC records and the gestational age inferred from the pregnancy end date (assuming a 40 week pregnancy) by ANC visit number. The horizontal dashed line denotes absolute agreement between ANC-estimated gestational age and the gestational age inferred from the HDSS pregnancy end date. Pregnancies missing outcomes or those missing visit numbers were excluded ($n = 354$). (B) Comparison between gestational age from ANC register and weeks to pregnancy end date in HDSS. The pregnancy end date is represented by the vertical dashed line. The orange line has been plotted from a linear regression of weeks to pregnancy outcome in the HDSS on ANC-estimated gestational age.

the odds of having a missing pregnancy outcome were 79% lower for women in the fourth quintile (OR 0.21, 95% CI 0.05–0.73). In the first model, the odds of pregnancy underreporting were 54% higher for unmarried women compared to those in formal union (OR 1.54, 95% CI 0.91–2.60). This increased to an odds ratio of 2.10 (95% CI 1.45–3.04) in the final model.

Discussion

In this work, we compared individually-linked ANC clinic data with records of pregnancies and their outcomes in Siaya HDSS. HDSS are often set up in regions where availability of other population health data is limited [16] and this offered a unique opportunity to externally validate HDSS pregnancy reporting completeness, identify predictors of underreporting, and examine potential bias in the reporting of adverse outcomes. Linkage with ANC is a particularly high value extension of HDSS data, given the almost universal coverage of ANC services in many parts of sub-Saharan Africa. It has been estimated that close to 90% of women in sub-Saharan Africa and 94% in Siaya HDSS receive ANC at least once during pregnancy [23, 41].

Of the 2,475 pregnancies observed in ANC records that were included in the analysis, there were 268 that did not have a corresponding pregnancy outcome report in Siaya HDSS. Close to 36% of these missing outcome reports were attributable to the individual exiting the study area or being lost to follow-up prior to the next HDSS data collection round. A household interview had not yet taken place in another 11% of cases. The remaining 53% did not have a corresponding pregnancy outcome report in the HDSS, despite the occurrence of a household interview following their expected delivery. This subset of missing outcomes is cause for concern, as such underreporting is more likely when an adverse pregnancy outcome or early death has occurred.

It is worth noting that outcomes which were not reported at the first interview following delivery could still be reported in later data collection rounds. However, delayed reports are only likely for pregnancies that were registered or resulted in a live birth. In the case of the former, the HDSS fieldworker will be prompted to inquire about the status of the pregnancy until an outcome is recorded or the individual is no longer resident in the study site. Alternatively, in the case of a live birth, the presence of a new child in the household helps to ensure that the birth is recorded when the child is enumerated in the HDSS. Otherwise, the likelihood that adverse outcomes are reported diminishes with time, amplifying the downward bias in measurements of perinatal and neonatal mortality.

Our results suggest that pregnancy registration can improve ascertainment of adverse outcomes that are otherwise vulnerable to underreporting. The observed rate of stillbirths for pregnancies that were registered in the HDSS was 12.6 (95% CI 6.3–18.9), while zero stillbirths were reported among those lacking pregnancy registration. The observed neonatal mortality for registered and unregistered pregnancies was similar, though significant differences in neonatal mortality by pregnancy registration status have been found in other research [18, 20]. If such differences were attributable to differential usage of ANC services, this could explain the lack of

Table 2: Characteristics of linked women that were missing pregnancy outcome reports in the HDSS

Variable	Value	*Missing pregnancy outcomes*	
		n (%)	p-value
ANC clinic visits	1	57 (17.2)	<0.01
	2	27 (6.1)	
	3	22 (4.4)	
	4+	35 (3.3)	
Age	10–19	34 (9.4)	0.01
	20–24	36 (6.6)	
	25–29	25 (4.6)	
	30–34	20 (4.0)	
	35+	26 (6.4)	
Duration of residency in HDSS	<2 years	27 (6.7)	0.59
	2+ years	114 (5.9)	
Education	None	1 (5.3)	0.99
	Primary	102 (6.0)	
	Secondary/Higher	37 (5.9)	
	Unknown	1 (11.1)	
Gestational age at 1st ANC visit	1st trimester	24 (10.3)	<0.01
	2nd trimester	44 (4.9)	
	3rd trimester	18 (5.7)	
	Unknown	55 (6.1)	
HDSS region	Asembo	11 (6.2)	0.98
	Gem	126 (6.0)	
	Karemo	4 (5.6)	
HIV status	Negative	100 (5.4)	0.04
	Positive	36 (8.1)	
	Unknown	5 (13.2)	
Household wealth quintile	1	21 (6.1)	0.02
	2	17 (11.1)	
	3	10 (8.7)	
	4	3 (2.2)	
	5	6 (4.4)	
	Unknown	84 (5.7)	
Marital status	In union	83 (4.7)	<0.01
	Not in union	58 (9.8)	
Parity	0	46 (8.7)	<0.01
	1	22 (6.1)	
	2+	66 (4.6)	
	Unknown	7 (35.0)	
Total		141 (6.0)	

Notes: Table provides row percentages, denoting the share of pregnancies observed in the ANC records that were missing outcomes in the HDSS out of the total number that were assessed for reporting. P-values display the results of Chi-squared tests of independence. Tests were performed on complete case data, with "Unknown" values excluded.

significant difference in our own analysis, where the sample was entirely composed of pregnancies for which the mother had accessed ANC. It is also possible that the lack of a significant difference in neonatal mortality in this work was related to the relatively small number of observations.

We found that women who began ANC in the first trimester and had fewer visits overall were more likely to be missing pregnancy outcomes in the HDSS. It is possible that some of these pregnancies were not carried to full term, and such women made fewer ANC visits as a result. This proposition is further supported by the finding that women who were unmarried and HIV-positive were more likely to have missing pregnancy outcomes. In both cases, these characteristics are independently associated with elevated risk

Table 3: Logistic regression results for having a missing pregnancy outcome in the HDSS

Variable		*Dependent variable: Missing pregnancy outcome*	
		(1)	(2)
		OR (95% CI)	OR (95% CI)
Age	10–19	1.26 (0.72, 2.24)	
	20–24	1	
	25–29	0.93 (0.50, 1.71)	
	30–34	0.79 (0.39, 1.57)	
	35+	1.11 (0.57, 2.20)	
ANC clinic visits		0.71 (0.62, 0.82)	0.70 (0.60, 0.80)
Gestational age at 1st ANC visit	1st trimester	2.70 (1.55, 4.62)	2.64 (1.52, 4.51)
	2nd trimester	1	1
	3rd trimester	0.89 (0.49, 1.57)	0.87 (0.48, 1.53)
	Unknown	0.85 (0.55, 1.32)	0.94 (0.61, 1.45)
HIV status	Negative	1	1
	Positive	1.89 (1.20, 2.94)	1.86 (1.22, 2.78)
	Unknown	2.39 (0.78, 6.00)	2.19 (0.72, 5.45)
Household wealth quintile	1	0.60 (0.27, 1.40)	0.58 (0.27, 1.35)
	2	1.27 (0.55, 3.08)	1.18 (0.51, 2.83)
	3	1	1
	4	0.22 (0.05, 0.75)	0.21 (0.05, 0.73)
	5	0.48 (0.16, 1.37)	0.47 (0.15, 1.35)
	Unknown	0.67 (0.34, 1.46)	0.66 (0.34, 1.41)
Marital status	In union	1	1
	Not in union	1.54 (0.91, 2.60)	2.10 (1.45, 3.04)
Parity	0	1.08 (0.56, 2.11)	
	1	1	
	2+	0.83 (0.46, 1.55)	
	Unknown	5.82 (1.84, 17.37)	
Observations		2,348	2,348
Log Likelihood		−487.66	−495.23
Akaike Inf. Criterion		1,015.32	1,016.46

Notes: Sample only includes individuals with missing outcomes who were resident in the HDSS for the first household interview following their EDD. ANC – antenatal care, OR – odds ratio, CI – confidence interval.

of perinatal mortality [42, 43], as well as under-5 mortality [44–46]. However, it could also be the case that women with fewer clinic visits out-migrated from the study area midway through pregnancy, and were thus also less likely to have reported pregnancy outcomes. While our evaluation was restricted to women who were resident in the HDSS during their ANC usage and for the first household interview following their EDD, some absences from the study area may not have been accurately reflected in HDSS records.

Tracking frequent in- and out-migrations that vary in their destination and duration is a complex task in HDSS [16]. Pregnancy and the postpartum period can be a time of increased mobility, when women often travel to seek medical care or the support of family members [47–49]. If such migrations were not captured by the HDSS, it may have appeared some individuals were present in the site for the data collection round following their delivery when they were in fact not. As this relates to our results, it is possible that the share of missing pregnancy outcomes attributable to migration was larger than observed, and the share that were unreported adverse outcomes was smaller.

Nevertheless, our results are consistent with previous findings of downward bias in HDSS estimates of early mortality. The tendency for deaths occurring soon after birth to be under-counted in population-based surveillance data was noted as early as the 1950s and 60s [50, 51]. More recently, rates of neonatal mortality in African HDSS have been found to be lower than corresponding estimates from retrospective household surveys such as DHS and MICS [19, 52]. Additionally, while both sources exhibit higher levels of child mortality (deaths between one and four years) than found in the historic record of high-quality data, HDSS are alone in their downward deviation at early ages [53]. Reliably monitoring pregnancy outcomes and early mortality has been identified as one of the most serious challenges faced by HDSS [16].

The magnitude of downward bias in HDSS estimates of adverse pregnancy outcomes and early mortality ultimately depends on the rate of such events among missing pregnancy outcomes, which cannot be determined without additional data collection. This is an area in which record linkage with ANC data could be very useful. Leveraging information on ANC usage to organize follow-up interviews with those who are missing pregnancy outcome reports is an important avenue for future research. Such work has been piloted using record-linked ANC data in HDSS in The Gambia, where it identified elevated rates of perinatal and neonatal mortality among pregnancies that were missing outcome reports in the HDSS [21].

Ongoing record linkage between HDSS and ANC clinics could allow for earlier pregnancy detection and more accurate outcome reporting. We found that individuals in Siaya HDSS had visited an ANC clinic prior to registering pregnancies in the site in approximately 77% of cases. Additionally, ANC clinic information on gestational age indicated that half of reported miscarriages were potentially misclassified stillbirths. However, it is worth noting that the assessment of misclassification between miscarriage and stillbirth in the HDSS relied on the accurate estimation of gestational age at the ANC clinics. Given the uncertainty of estimating gestational age through fundal height palpitation and reported LMP [54], it is important to interpret these results cautiously.

This study is subject to a few important limitations. First, it would be beneficial to conduct follow-up data collection on pregnancies that were observed in the ANC but missing outcome reports in the HDSS. This analysis identified pregnancies that were likely to have ended in adverse outcomes given that they were not reported in a data collection round occurring after the EDD. This shed light on the potential for bias in HDSS estimates of adverse pregnancy outcomes and early mortality, however, additional information is needed in order to provide a more precise estimate of its magnitude. It is also important to acknowledge that close to 45% of women seeking ANC who self-reported residence in the HDSS were not successfully linked to their record. While some characteristics associated with lower HDSS-ANC match rates were also associated with having a missing pregnancy outcome in the regression analysis (e.g. being younger, seeking ANC services in the first trimester), others were positively associated with pregnancy outcome reporting (e.g. being HIV negative, in union). It is thus unclear how an improvement in linkage rates would affect the proportion of pregnancy outcomes missing from HDSS records.

Conclusion

HDSS are valuable sources of empirical demographic and epidemiological data for much of sub-Saharan Africa where current systems of CRVS and HMIS are deficient. However, HDSS data on pregnancies, pregnancy outcomes and neonatal mortality is often incomplete. This research demonstrates the potential of using record linkage with ANC clinics to evaluate pregnancy reporting completeness in HDSS and investigate bias in estimates of adverse pregnancy outcomes and early mortality. Record linkage between HDSS and routine programme data is an efficient manner of augmenting population health information in sub-Saharan Africa, and addressing the lack of good quality data on pregnancy and early mortality. Such efforts have the potential to both improve our understanding of population health and our ability to accurately measure it.

Acknowledgments

This work constitutes PhD research funded by the UK Economic and Social Research Council (ESRC). The PhD project is nested within a National Institute of Child Health and Human Development grant on Age patterns of under-five mortality in human populations (NICHD R01HD090082).

Statement on conflicts of interest

The authors declare no conflict of interest.

Ethics statement

Ethical approval for the record linkage study was obtained from the KEMRI Scientific Ethics Review Unit (SERU application 3589) and the institutional review board of the London School of Hygiene and Tropical Medicine under protocol 14458. Ethical approval for this analysis was obtained from the Institutional Review Board of the London School of Hygiene & Tropical Medicine under protocol 25133.

References

1. UN IGME. A neglected tragedy: the global burden of stillbirths. United Nations Children's Fund; 2020.

2. UN IGME. Levels & Trends in Child Mortality: Report 2021, Estimates developed by the United Nations Inter-agency Group for Child Mortality Estimation. New York: United Nations Children's Fund; 2021.

3. United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development. In: Rosa W, editor. A New Era in Global Health. New York, NY: Springer Publishing Company; 2017. https://doi.org/10.1891/9780826190123.0013

4. Lawn JE, Blencowe H, Waiswa P, Amouzou A, Mathers C, Hogan D, et al. Stillbirths: rates, risk factors, and acceleration towards 2030. The Lancet. 2016;387(10018):587–603. https://doi.org/10.1016/S0140-6736(15)00837-5

5. World Health Organization. Every newborn: an action plan to end preventable deaths. Geneva: World Health Organization; 2014.

6. AbouZahr C, de Savigny D, Mikkelsen L, Setel PW, Lozano R, Lopez AD. Towards universal civil registration and vital statistics systems: the time is now. The Lancet. 2015;386(10001):1407–18. https://doi.org/10.1016/S0140-6736(15)60170-2

7. Kwesiga D, Tawiah C, Imam MA, Tesega AK, Nareeba T, Enuameh YAK, et al. Barriers and enablers to reporting pregnancy and adverse pregnancy outcomes in population-based surveys: EN-INDEPTH study. Popul Health Metr. 2021;19(1):15. https://doi.org/10.1186/s12963-020-00228-x

8. Bradley SEK, Winfrey W, Croft TN. Contraceptive use and perinatal mortality in the DHS: An assessment of the quality and consistency of calendars and histories. Rockville, Maryland, USA: ICF International; 2015. Report No.: 17.

9. Pullum T, Becker S. DHS Methodological Reports 11: Evidence Omission and Displacement in DHS Birth Histories. Rockville, Maryland, USA: ICF International; 2014.

10. DHS. DHS-8 questionnaires: revision process and new content [Internet]. ICF International; 2019. Available from: https://dhsprogram.com/pubs/pdf/DHSM11/DHSM11.pdf

11. Akuze J, Blencowe H, Waiswa P, Baschieri A, Gordeev VS, Kwesiga D, et al. Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in five countries: the Every Newborn-INDEPTH study. Lancet Glob Health. 2020;8(4):e555–66. https://doi.org/10.1016/S2214-109X(20)30044-9

12. Haws RA, Mashasi I, Mrisho M, Schellenberg JA, Darmstadt GL, Winch PJ. "These are not good things for other people to know": How rural Tanzanian women's experiences of pregnancy loss and early neonatal death may impact survey data quality. Soc Sci Med. 2010;71(10):1764–72. https://doi.org/10.1016/j.socscimed.2010.03.051

13. Kiguli J, Namusoko S, Kerber K, Peterson S, Waiswa P. Weeping in silence: community experiences of stillbirths in rural eastern Uganda. Glob Health Action. 2015;8(1). https://doi.org/10.3402/gha.v8.24011

14. Kadobera D, Waiswa P, Peterson S, Blencowe H, Lawn J, Kerber K, et al. Comparing performance of methods used to identify pregnant women, pregnancy outcomes, and child mortality in the Iganga-Mayuge Health and Demographic Surveillance Site, Uganda. Glob Health Action. 2017;10(1). https://doi.org/10.1080/16549716.2017.1356641

15. Blencowe H, Cousens S, Jassir FB, Say L, Chou D, Mathers C, et al. National, regional, and worldwide estimates of stillbirth rates in 2015, with trends from 2000: a systematic analysis. Lancet Glob Health. 2016;4(2):e98–108. https://doi.org/10.1016/S2214-109X(15)00275-2

16. Sankoh O, Byass P. The INDEPTH Network: filling vital gaps in global epidemiology. Int J Epidemiol. 2012;41(3):579–88. https://doi.org/10.1093/ije/dys081

17. Kwon C, Naser AM, Eilerts H, Reniers G, Cunningham SA. Pregnancy Surveillance Methods within Health and Demographic Surveillance Systems. Gates Open Res. 2021;5(144). https://doi.org/10.12688/gatesopenres.13332.1

18. Eilerts H, Romero Prieto J, Herbst K, Gareta D, Jasseh M, Khagayi S, et al. Pregnancy Reporting and Biases in Under-5 Mortality in Three HDSS in Western, Eastern, and Southern Africa. In: Session 136, Data Quality: From Pregnancy to Death. 2021. (International Union for the Scientific Study of Population).

19. Waiswa P, Moyer C, Kwesiga D, Arthur S, Sankoh O, Welaga P, et al. Status of birth and pregnancy outcome capture in Health Demographic Surveillance Sites in 13 countries. Int J Public Health. 2019;64(6):909–20. https://doi.org/10.1007/s00038-019-01241-0

20. Nareeba T, Dzabeng F, Alam N, Biks GA, Thysen SM, Akuze J, et al. Neonatal and child mortality data in retrospective population-based surveys compared with prospective demographic surveillance: EN-INDEPTH study. Popul Health Metr. 2021;19(1):7. https://doi.org/10.1186/s12963-020-00232-1

21. Rerimoi AJ. Development and assessment of systems for population-based estimation of neonatal and perinatal mortality in The Gambia [PhD thesis]. 2019.

22. World Health Organization. WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience: Summary. 2018.

23. Adedokun ST, Yaya S. Correlates of antenatal care utilization among women of reproductive age in sub-Saharan Africa: evidence from multinomial analysis of demographic and health surveys (2010-2018) from 31 countries. Arch Public Health. 2020;78:134. https://doi.org/10.1186/s13690-020-00516-w

24. Ford JB, Roberts CL, Taylor LK. Characteristics of unmatched maternal and baby records in linked birth records and hospital discharge data. Paediatr Perinat Epidemiol. 2006;20(4):329–37. https://doi.org/10.1111/j.1365-3016.2006.00715.x

25. Chen J, Fair M, Wilkins R, Cyr M. Maternal education and fetal and infant mortality in Quebec. Fetal and Infant Mortality Study Group of the Canadian Perinatal Surveillance System. Health Rep. 1998;10(2):53–64 (Eng); 57–70 (Fre).

26. Howard L. The psychosocial outcome of pregnancy in women with psychotic disorders. Schizophr Res. 2004;71(1):49–60. https://doi.org/10.1016/j.schres.2004.01.003

27. Riordan DV, Morris C, Hattie J, Stark C. Family size and perinatal circumstances, as mental health risk factors in a Scottish birth cohort. Soc Psychiatry Psychiatr Epidemiol. 2012;47(6):975–83. https://doi.org/10.1007/s00127-011-0405-5

28. Holian J, Mallick MJ, Zaremba CM. Maternity and infant care, race and birth outcomes. J Health Soc Policy. 2004;18(4):1–11. https://doi.org/10.1300/J045v18n04_01

29. Holian J. Live birth and infant death record linkage: methodological and policy issues. J Health Soc Policy. 2000;12(1):1–10. https://doi.org/10.1300/J045v12n01_01

30. Joubert J, Bradshaw D, Kabudula C, Rao C, Kahn K, Mee P, et al. Record-linkage comparison of verbal autopsy and routine civil registration death certification in rural north-east South Africa: 2006-09. Int J Epidemiol. 2014;43(6):1945–58. https://doi.org/10.1093/ije/dyu156

31. Kabudula CW, Clark BD, Gómez-Olivé FX, Tollman S, Menken J, Reniers G. The promise of record linkage for assessing the uptake of health services in resource constrained settings: a pilot study from South Africa. BMC Med Res Methodol. 2014;14(1):71. https://doi.org/10.1186/1471-2288-14-71

32. Rentsch CT, Reniers G, Kabudula C, Machemba R, Mtenga B, Harron K, et al. Point-of-contact interactive record linkage (PIRL) between demographic surveillance and health facility data in rural Tanzania. Int J Popul Data Sci [Internet]. 2017 [cited 2019 May 15];2(1). https://doi.org/10.23889/ijpds.v2i1.408

33. Odhiambo FO, Laserson KF, Sewe M, Hamel MJ, Feikin DR, Adazu K, et al. Profile: The KEMRI/CDC Health and Demographic Surveillance System–Western Kenya. Int J Epidemiol. 2012;41(4):977–87. https://doi.org/10.1093/ije/dys108

34. Kabudula C, Rentsch CT, Catlett J, Beckles D, Masilela N, Zaba B, et al. LSHTM ALPHA network / PIRL RecordLinkageSoftware: Initial public release of Point-of-contact Interactive Record Linkage (PIRL) software [Internet]. Zenodo; 2017. https://doi.org/10.5281/zenodo.998868

35. Rentsch CT, Kabudula CW, Catlett J, Beckles D, Machemba R, Mtenga B, et al. Point-of-contact Interactive Record Linkage (PIRL): A software tool to prospectively link demographic surveillance and health facility data. Gates Open Res. 2017;1:8. https://doi.org/10.12688/gatesopenres.12751.1

36. Fellegi IP, Sunter AB. A Theory for Record Linkage. J Am Stat Assoc. 1969;64.

37. Cook LJ, Olson LM, Dean JM. Probabilistic record linkage: relationships between file sizes, identifiers and match weights. Methods Inf Med. 2001;40(3):196–203. https://doi.org/10.1055/S-0038-1634155

38. Jaro MA. Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida. J Am Stat Assoc. 1989;84(406):414–20. https://doi.org/10.1080/01621459.1989.10478785

39. World Health Organization. International Statistical Classification of Diseases and Related Health Problems. 2010th ed. Vol. 2. 2013.

40. Amek N, Vounatsou P, Obonyo B, Hamel M, Odhiambo F, Slutsker L, et al. Using health and demographic surveillance system (HDSS) data to analyze geographical distribution of socio-economic status; an experience from KEMRI/CDC HDSS. Acta Trop. 2015 Apr;144:24–30. https://doi.org/10.1016/j.actatropica.2015.01.006

41. Kohler PK, Okanda J, Kinuthia J, Mills LA, Olilo G, Odhiambo F, et al. Community-Based Evaluation of PMTCT Uptake in Nyanza Province, Kenya. PLOS ONE. 2014;9(10):e110110. https://doi.org/10.1371/journal.pone.0110110

42. Aminu M, Unkels R, Mdegela M, Utz B, Adaji S, van den Broek N. Causes of and factors associated with stillbirth in low- and middle-income countries: a systematic literature review. BJOG Int J Obstet Gynaecol. 2014;121(s4):141–53. https://doi.org/10.1111/1471-0528.12995

43. Kupka R, Kassaye T, Saathoff E, Hertzmark E, Msamanga GI, Fawzi WW. Predictors of stillbirth among HIV-infected Tanzanian women. Acta Obstet Gynecol Scand. 2009;88(5):584–92. https://doi.org/10.1080/00016340902835901

44. Clark S, Hamplová D. Single Motherhood and Child Mortality in Sub-Saharan Africa: A Life Course Perspective. Demography. 2013;50(5):1521–49. https://doi.org/10.1007/s13524-013-0220-6

45. Marston M, Becquet R, Zaba B, Moulton LH, Gray G, Coovadia H, et al. Net survival of perinatally and postnatally HIV-infected children: a pooled analysis of individual data from sub-Saharan Africa. Int J Epidemiol. 2011;40(2):385–96. https://doi.org/10.1093/ije/dyq255

46. Zaba B, Marston M, Nakiyingi J, Whitworth J, Ruberantwari A, Urassa M, et al. HIV and Child Mortality: Evidence from Surveillance Studies in Uganda, Tanzania and Malawi. University of North Carolina at Chapel Hill: Carolina Population Center; 2003 p. 26. (MEASURE Evaluation).

47. Clouse K, Pettifor A, Shearer K, Maskew M, Bassett J, Larson B, et al. Loss to follow-up before and after delivery among women testing HIV positive during pregnancy in Johannesburg, South Africa. Trop Med Int Health. 2013;18(4):451–60. https://doi.org/10.1111/tmi.12072

48. Kaplan R, Orrell C, Zwane E, Bekker LG, Wood R. Loss to follow-up and mortality amongst pregnant women referred to a community clinic for antiretroviral treatment. AIDS Lond Engl. 2008;22(13):1679–81. https://doi.org/10.1097/QAD.0b013/e32830ebcee

49. Wang B, Losina E, Stark R, Munro A, Walensky RP, Wilke M, et al. Loss to Follow-Up in a Community Clinic in South Africa – Roles of Gender, Pregnancy and CD4 count. South Afr Med J Suid-Afr Tydskr Vir Geneeskd. 2011;101(4):253–7. https://doi.org/10.7196/samj.4078

50. Billewicz WZ, McGregor IA. The Demography of Two West African (Gambian) Villages, 1951–75. J Biosoc Sci. 1981;13(2):219–40. https://doi.org/10.1017/S0021932000013390

51. Cantrelle P. Etude démographique dans la région du Sine-Saloum (Sénégal): état civil et observation démographique 1963–1965. Paris: ORSTOM; 1969. (Travaux et Documents de l'O.R.S.T.O.M.).

52. Eilerts H, Romero Prieto J, Eaton JW, Reniers G. Age patterns of under-5 mortality in sub-Saharan Africa during 1990-2018: A comparison of estimates from demographic surveillance with full birth histories and the historic record. Demogr Res. 2021;44(18):415–42. https://doi.org/10.4054/DemRes.2021.44.18

53. Verhulst A, Romero Prieto J, Alam N, Eilerts-Spinelli H, Erchick DJ, Gerland P, et al. Divergent age patterns of under-5 mortality in south Asia and sub-Saharan Africa: a modelling study. Lancet Glob Health. 2022. https://doi.org/10.1016/S2214-109X(22)00337-0

54. White LJ, Lee SJ, Stepniewska K, Simpson JA, Dwell SLM, Arunjerdja R, et al. Estimation of gestational age from fundal height: a solution for resource-poor settings. J R Soc Interface. 2012 Mar 7;9(68):503–10. https://doi.org/10.1098/rsif.2011.0376

Abbreviations

ANC:	Antenatal Care
CRVS:	Civil Registration and Vital Statistics
DHS:	Demographic and Health Survey
EDD:	Expected Delivery Date
HDSS:	Health and Demographic Surveillance System
HMIS:	Health Management Information Systems
KEMRI:	Kenya Medical Research Institute
LMP:	Last Menstrual Period
PIRL:	Point-of-contact Interactive Record Linkage

Supplementary Table 1: Characteristics of individuals visiting antenatal care (ANC) clinics that were linked to Siaya Health and Demographic Surveillance System (HDSS)

	Linkage with HDSS	
Variable	Value	n (%)
Age	10–19	538 (51.2)
	20–24	874 (44.9)
	25–29	730 (56.6)
	30–34	595 (64.5)
	35+	436 (74.3)
Date of enrolment	February–December 2018	1,162 (62.2)
	January 2019–March 2020	1,497 (52.9)
	September–December 2020	514 (46.9)
Gestational age at 1st ANC visit	1st trimester	390 (49.4)
	2nd trimester	1,535 (55.4)
	3rd trimester	1,060 (56.3)
	Unknown	188 (53.4)
HIV status	Negative	2,568 (53.6)
	Positive	605 (60.7)
HDSS region	Asembo	365 (48.0)
	Gem	2,693 (56.2)
	Karemo	15 (48.5)
Marital status	In union	2,247 (53.3)
	Not in union	924 (58.7)
	Unknown	2 (66.7)
Parity	0	799 (46.8)
	1	534 (39.6)
	2+	1,806 (68.2)
	Unknown	34 (38.2)
Total		3,173 (54.8)