Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in five countries: the Every Newborn-INDEPTH study

Joseph Akuze, Hannah Blencowe, Peter Waiswa, Angela Baschieri, Vladimir S Gordeev, Doris Kwesiga, Ane B Fisker, Sanne M Thysen, Amabelia Rodrigues, Gashaw A Biks, Solomon M Abebe, Kassahun A Gelaye, Mezgebu Y Mengistu, Bisrat M Geremew, Tadesse G Delele, Adane K Tesega, Temesgen A Yitayew, Simon Kasasa, Edward Galiwango, Davis Natukwatsa, Dan Kajungu, Yeetey AK Enuameh, Obed E Nettey, Francis Dzabeng, Seeba Amenga-Etego, Sam K Newton, Charlotte Tawiah, Kwaku P Asante, Seth Owusu-Agyei, Nurul Alam, Moinuddin M Haider, Ali Imam, Kaiser Mahmud, Simon Cousens, Joy E Lawn*, on behalf of the Every Newborn-INDEPTH Study Collaborative Group†*

Summary

Background An estimated 5·1 million stillbirths and neonatal deaths occur annually. Household surveys, most notably the Demographic and Health Survey (DHS), run in more than 90 countries and are the main data source from the highest burden regions, but data-quality concerns remain. We aimed to compare two questionnaires: a full birth history module with additional questions on pregnancy losses (FBH+; the current DHS standard) and a full pregnancy history module (FPH), which collects information on all livebirths, stillbirths, miscarriages, and neonatal deaths.

Methods Women residing in five Health and Demographic Surveillance System sites within the INDEPTH Network (Bandim in Guinea-Bissau, Dabat in Ethiopia, IgangaMayuge in Uganda, Matlab in Bangladesh, and Kintampo in Ghana) were randomly assigned (individually) to be interviewed using either FBH+ or FPH between July 28, 2017, and Aug 13, 2018. The primary outcomes were stillbirths and neonatal deaths in the 5 years before the survey interview (measured by stillbirth rate [SBR] and neonatal mortality rate [NMR]) and mean time taken to complete the maternity history section of the questionnaire. We also assessed between-site heterogeneity. This study is registered with the Research Registry, 4720.

Findings 69 176 women were allocated to be interviewed by either FBH+ (n=34 805) or FPH (n=34 371). The mean time taken to complete FPH (10·5 min) was longer than for FBH+ (9·1 min; p<0·0001). Using FPH, the estimated SBR was 17·4 per 1000 total births, 21% (95% CI –10 to 62) higher than with FBH+ (15·2 per 1000 total births; p=0·20) in the 5 years preceding the survey interview. There was strong evidence of between-site heterogeneity (I^2=80·9%; p<0·0001), with SBR higher for FPH than for FBH+ in four of five sites. The estimated NMR did not differ between modules (FPH 25·1 per 1000 livebirths *vs* FBH+ 25·4 per 1000 livebirths), with no evidence of between-site heterogeneity (I^2=0·7%; p=0·40).

Interpretation FPH takes an average of 1·4 min longer to complete than does FBH+, but has the potential to increase reporting of stillbirths in high burden contexts. The between-site heterogeneity we found might reflect variations in interviewer training and survey implementation, emphasising the importance of interviewer skills, training, and consistent implementation in data quality.

Funding Children's Investment Fund Foundation.

Introduction

Around 2·6 million stillbirths and 2·5 million neonatal deaths are estimated to occur worldwide each year, 98% of these are in low-income and middle-income countries (LMICs) and the majority are preventable.[1,2] Sustainable Development Goal 3 for healthy lives and wellbeing specifies that, by 2030, every country should reach a neonatal mortality rate (NMR) of 12 per 1000 livebirths or lower, and the Every Newborn Action Plan includes a similar target for stillbirths (stillbirth rate [SBR] ≤12 stillbirths per 1000 total births).[3] Monitoring progress towards these targets requires regular, timely, and reliable data.

The 2% of global neonatal deaths and stillbirths that occur in high-income countries are recorded by high-quality Civil and Vital Registration Statistics systems. In these settings, robust data on causes and care are also recorded within the health sector and are linked to perinatal audit systems to further reduce preventable deaths. However, most stillbirths and neonatal deaths occur in LMICs, with about 75% in sub-Saharan Africa and southern Asia, where often few or no reliable data on these events are captured in Civil and Vital Registration Statistics or other health data systems.[4,5] Hence, although investments in data systems are increasing, LMICs still largely rely on nationally

Lancet Glob Health 2020; 8: e555–66

See **Comment** page e464

*Contributed equally

†Collaborators listed at the end of the Article

Maternal, Adolescent, Reproductive, and Child Health Centre (J Akuze DSTA, H Blencowe PhD, A Baschieri PhD, V S Gordeev PhD, Prof S Cousens DipMathStat, Prof J E Lawn FRCPCH), **and Malaria Centre** (Prof S Owusu-Agyei PhD), **London School of Hygiene & Tropical Medicine, London, UK; Center of Excellence for Maternal Newborn and Child Health Research, School of Public Health, Makerere University, Kampala, Uganda** (J Akuze, P Waiswa PhD, D Kwesiga MSc, S Kasasa PhD); **Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden** (P Waiswa); **Bandim Health Project, Bissau, Guinea-Bissau** (A B Fisker PhD, S M Thysen PhD, A Rodrigues PhD); **Research Center for Vitamins and Vaccines, Bandim Health Project, Statens Serum Institut, Copenhagen, Denmark** (A B Fisker, S M Thysen); **Odense Patient data Explorative Network, Odense University Hospital/Institute of Clinical Research, University of Southern Denmark, Odense, Denmark** (A B Fisker); **Center for Global Health, Department of Public Health, Aarhus University, Aarhus, Denmark** (S M Thysen); **Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia** (G A Biks PhD, S M Abebe PhD, K A Gelaye PhD, M Y Mengistu PhD, B M Geremew PhD, T G Delele VMD, A K Tesega MPH, T A Yitayew MSc); **IgangaMayuge Health and Demographic Surveillance System, Iganga, Uganda** (S Kasasa, E Galiwango MA, D Natukwatsa BSc, D Kajungu PhD); **Kwame Nkrumah University of Science and Technology, Kumasi, Ashanti, Ghana** (Y A K Enuameh PhD, Prof S K Newton PhD); **Kintampo Health Research Centre, Kintampo, Ghana** (Prof S Owusu-Agyei, Y A K Enuameh, O E Nettey MPhil, F Dzabeng MSc, S Amenga-Etego MSc, Prof S K Newton, C Tawiah MSc, K P Asante PhD); **University of Health and Allied Sciences, Kintampo Health Research Centre, Kintampo, Ghana** (Prof S Owusu-Agyei); **Health Systems and Population Studies Division, icddr,b, Dhaka, Bangladesh** (N Alam PhD, M M Haider MPH, A Imam MPH, K Mahmud MSc); **and The Institute of Population Health Sciences, Queen Mary University of London, London, UK** (V S Gordeev)

Correspondence to: Mr Joseph Akuze, Maternal, Adolescent, Reproductive, and Child Health Centre, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK **jakuze@musph.ac.ug**

Research in context

Evidence before this study

An estimated 2·6 million stillbirths and 2·5 million neonatal deaths occur worldwide every year. The majority (98%) occur in low-income and middle-income countries, which remain largely reliant on population-based household surveys—notably the Demographic and Health Survey (DHS), which has a 40-year history in more than 90 countries—to measure these deaths. However, reporting of stillbirths in many of these surveys is low, and there has been little research on how to improve the reporting of stillbirths in household surveys. Some evidence from previous studies and grey literature has suggested a full pregnancy history (FPH) approach might be associated with increased reporting of births and early neonatal deaths compared with a full birth history (FBH+), which has been the standard approach in DHS. No study has directly compared these two methods for mortality rate reporting or for time taken to complete each survey module.

Added value of this study

The Every Newborn-INDEPTH study is the first randomised comparison of two survey methods for the measurement of stillbirth and neonatal death and the time taken to complete these modules. 69 176 women from five countries were individually randomised between study groups. FPH recorded a higher stillbirth rate than did FBH+ in four of five sites, being 21% (95% CI –10 to 62) higher on average than FBH+. There was heterogeneity between sites, with one site recording fewer stillbirths with the FPH approach, possibly explained by different training on this aspect (2–3 h rather than 2–3 days). Neonatal mortality did not differ between the two modules. The mean completion time was slightly longer for FPH than for FBH+ (10·5 min *vs* 9·1 min).

Implications of all the available evidence

Estimated stillbirth rates were higher using the FPH approach, with a small difference in time to administer the survey. Population-based household surveys should consider adopting the FPH approach to improve the reporting of stillbirths. The DHS programme has recently altered their core survey module to change to FPH based on this study's findings. More research is required to inform further refinements to survey questions, context-specific adaptation, and implementation, including interviewer training, assessing stillbirth rate data, and examining the measurement of other pregnancy outcomes.

representative household surveys to obtain data on these indicators.[6,7] Additionally, many of the highest risk countries are those with humanitarian crises, and these contexts are likely to be dependent on survey data for much longer.[8] The largest multicountry platform for such surveys is the Demographic and Health Survey (DHS) programme, which has run for nearly 40 years in more than 90 countries. Another important multicountry survey platform is UNICEF's Multiple Indicator Cluster Survey (MICS).

These household surveys usually involve asking a nationally representative sample of women about all their previous births, with more detailed information collected on the most recent pregnancies and their outcomes, usually for the 5 years preceding the day of the interview. However, the quality of birth outcome data varies between surveys, and potential underlying reasons for discrepancies (such as variation in tools, training and implementation, or context and barriers to response) are little researched. Surveys use different methods to ask women about their maternity histories, such as including details on all pregnancies or only on livebirths. These questions can be asked backwards or forwards or can be truncated.[9–11]

Since 1984, the DHS programme's questionnaires have used a full birth history (FBH) module, which records each pregnancy ending in a livebirth.[10] However, 17 countries have chosen to use the full pregnancy history (FPH) module, which was previously used in the World Fertility Survey and Contraceptive Prevalence Survey, the predecessors of the DHS.[9,10,12] FBH was preferred by the DHS programme, because the initial focus of the DHS was to improve measurement of child mortality.[9] Until 2013, the core DHS questionnaire used a reproductive calendar to collect information on stillbirths for the 5 years before the survey. The core DHS questionnaire (DHS-7), which has been used in 48 countries since 2013, uses a full livebirth history to gather data on under-5 and neonatal deaths, plus additional questions on pregnancy losses in the past 5 years to document stillbirths (known as the full birth history+ [FBH+]) rather than the reproductive calendar.[9,10] However, analyses have suggested that this approach misses some stillbirths and, to a lesser extent, early neonatal deaths.[9,10,13,14]

Another option is to record a woman's full pregnancy history, including pregnancies that do not end in a livebirth. This strategy has been postulated to improve reporting of pregnancy outcomes, notably stillbirths, but also miscarriages and termination of pregnancy. The FPH is often assumed to take longer to collect, although no studies have been done to compare the time taken to complete each module.[10,11]

Hence, there is a widely recognised need to evaluate the questionnaires used in surveys to improve estimates of the NMR and, particularly, the SBR. We aimed to compare two approaches of collecting maternity history (DHS-7 FBH+ *vs* FPH) to examine whether the two methods yield different estimates of SBR and NMR, and to determine whether there is a difference in completion time for these two approaches.

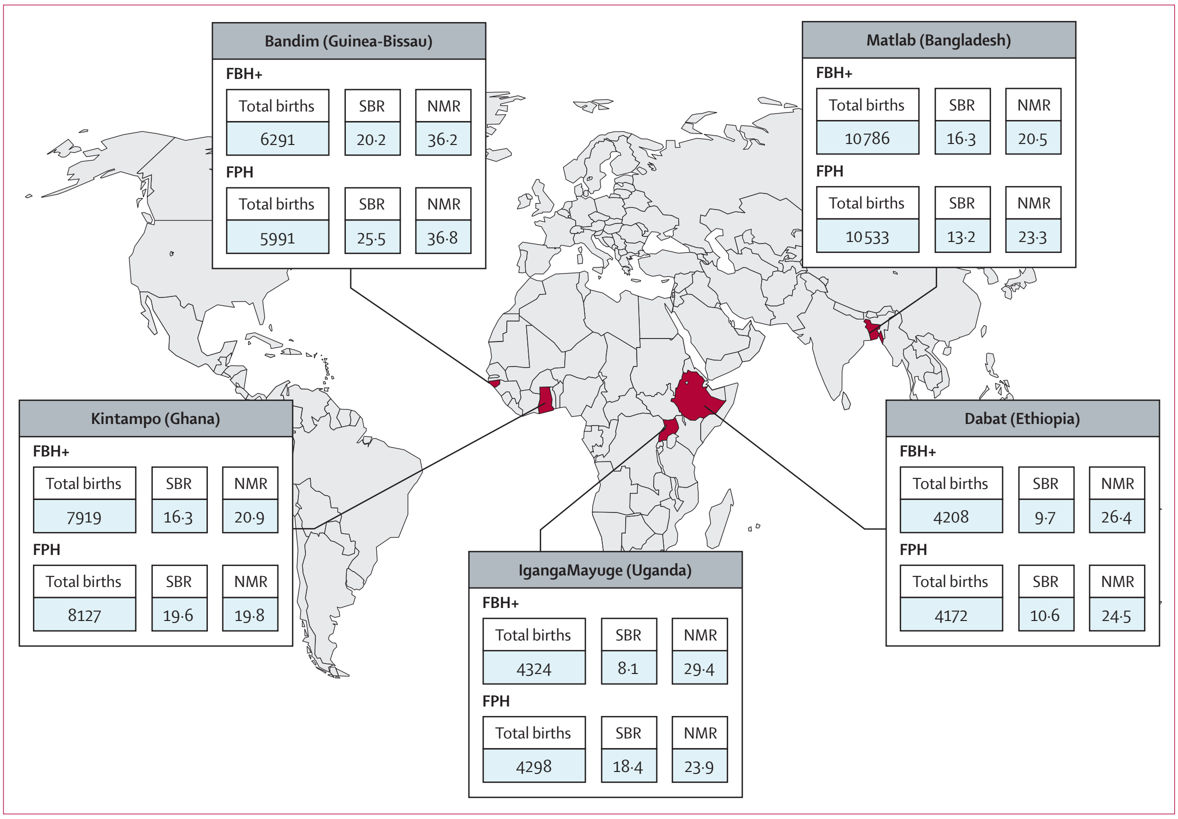

Figure 1: **Five Health and Demographic Surveillance System sites with summary statistics from the Every Newborn-INDEPTH study**
FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. NMR=neonatal mortality rate. SBR=stillbirth rate.

Methods

Study design and participants

The Every Newborn–International Network for the Demographic Evaluation of Populations and their Health (EN-INDEPTH) study was a cross-sectional, multisite study done in five Health and Demographic Surveillance System (HDSS) sites within the INDEPTH Network: Bandim in Guinea-Bissau, Dabat in Ethiopia, IgangaMayuge in Uganda, Matlab in Bangladesh, and Kintampo in Ghana (figure 1).[15] Details of the study protocol, including selection of sites and sample size (at least 68 000 total births with 80% power to detect a difference of ≥15% between proportions of total births that were stillborn between the FBH+ and FPH), have previously been published.[15] We report the results of objective 1 of the main EN-INDEPTH study: to undertake a randomised comparison of the reproductive modules used in DHS-7 (FBH+) versus an FPH module to examine the variation in reporting of stillbirths and neonatal deaths. Results of the other objectives will be published elsewhere.

In each site, we undertook a household survey of women aged 15–49 years who consented to participate. We randomly assigned (1:1) women to be interviewed using a questionnaire containing either an FBH+ or FPH module in a parallel design (appendix 1).[15] In subsection 2.1 of the survey, we asked women to state their lifetime total number of liveborn children (FBH+ and FPH) and total number of pregnancy losses (FPH only). In subsection 2.2, we asked women details about their lifetime livebirths (FBH+) and lifetime pregnancies (FPH). In subsection 2.3, we asked women in the FBH+ group about pregnancy losses in the past 5 years. We visited eligible women up to three times to seek consent for participation. We gathered data on interviewer characteristics through a self-completed questionnaire.

This study gained ethics approval from the London School of Hygiene & Tropical Medicine ethics committee (12218) and the relevant ethics committees in the five countries (appendix 2 p 8).

See Online for appendix 2

Randomisation and masking

We used HDSS listings as the sampling frame. We selected for inclusion in the study all women of reproductive age in Dabat, IgangaMayuge, and Kintampo; all women with a recorded birth outcome in the past 5 years in Matlab and Bandim urban sites; and a random sample of 80% of all women with a recorded birth outcome in the past 5 years in the rural Bandim site.[15] Interviewers attempted to locate all selected women and obtain their consent to take part in the survey. Women who gave consent were randomly assigned (individually) by the Survey Solutions application (versions 5.21, 5.22,

See Online for appendix 1

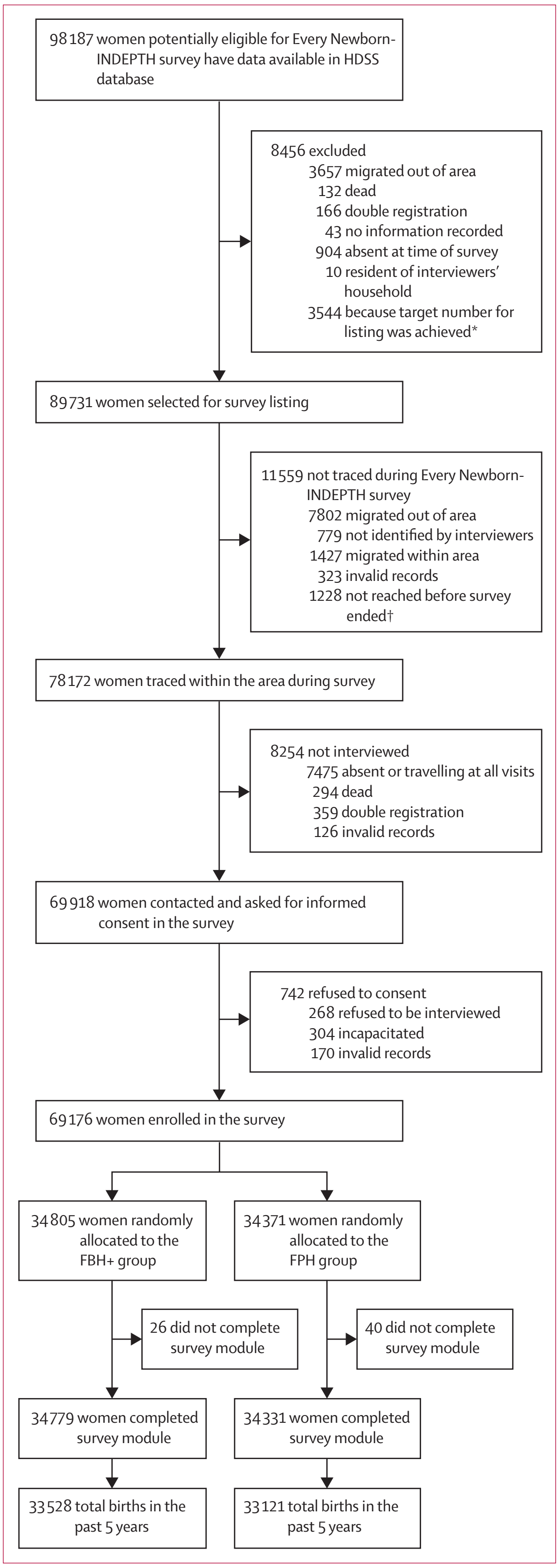

Figure 2: **Every Newborn-INDEPTH study profile**
FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. HDSS=Health and Demographic Surveillance System. *In IgangaMayuge, when creating the Every Newborn-INDEPTH study listing, some women were excluded because the site team had reached the target number of women as per the study protocol. †In IgangaMayuge, it was not possible in the time available to trace all the selected women and 1228 (6·8% of the women in the study listing) were not traced during the data collection period. There is evidence that these women differed with regards to demographic characteristics from included women (appendix 2 p 16).

5.23, 5.24, 5.25, 5.26, 18.04, and 18.06; The World Bank, Washington, DC, USA) to be interviewed using either the FBH+ or FPH module (1:1). Interviewers were not masked to the module but were not informed of the study's hypothesis.

Interviewer selection and training

We selected interviewers from staff already working for the HDSS sites. In IgangaMayuge only, interviewers had an opportunity for pre-training self-study of the interviewer manuals and tools. Interviewers were trained by facilitators for 19 days in Bandim, 21 days in Dabat, 10 days in IgangaMayuge, 6·5 days in Matlab, and 10 days in Kintampo.

Data collection and management

We obtained woman and interviewer data on Android tablets using the World Bank's Survey Solutions data collection and management system, which records data on survey processes (paradata), including timestamps.[16] Data from all five HDSS sites were anonymised by the local HDSS scientists, encrypted, and shared once data collection had been completed in each site.[15] The received datasets were cleaned and merged using Stata (version 15.1; StataCorp, College Station, TX, USA) and R statistical programming software in RStudio (version 1.2.5033; RStudio, Boston, MA, USA). The variables for the analysis were extracted from the pooled datasets and analysed using Stata version 15.1.

Outcomes

The primary outcomes were stillbirths (fetal death with a reported gestational age of ≥7 months) and neonatal deaths (death before 28 days after birth) in the 5 years before the survey interview—and their associated rates, SBR, and NMR—and time taken to complete the maternity history section of the questionnaire.

Time taken to complete the maternity history section was defined as time (in min) taken to complete subsections 2.1, 2.2, and 2.3 for FPH and FBH+; in FPH, questions on termination of pregnancy in subsection 2.3 were excluded (appendix 2 p 7).

We used socioeconomic wealth quintiles to measure the wealth status of households, which we derived from infrastructure, housing, and assets owned using principle components analysis, as used by DHS and

MICS (appendix 2 p 9).[17] Details of other independent variables are given in appendix 2 (p 11).

Statistical analysis

We used the DHS programme's century month code method[18,19] to identify events occurring in the 5 years before the interview (appendix 2 p 23). We computed site-specific crude risk ratios (RRs) with corresponding 95% CIs to compare SBRs and NMRs between the FBH+ and FPH groups. Estimates accounted for clustering of outcomes within women through use of generalised estimating equations (GEEs) with an exchangeable correlation matrix. We fitted GEE models using a log-link function and the binomial probability model. We then combined site-specific estimates using meta-analysis with random effects to obtain an overall estimate of the RR. We assessed evidence for heterogeneity between sites using the I^2 statistic obtained from the meta-analysis.

We did sensitivity analyses by excluding sites that were outliers. We checked for the potential omission of events by assessing consistency of reporting pregnancy losses between subsections 2.1, 2.2, and 2.3. We assessed the possibility that results were affected by differential misclassification between stillbirths and early neonatal deaths by computing site-specific RRs and their 95% CIs for perinatal mortality rates between the study groups.

We stratified our analyses by maternal and interviewer characteristics to produce stratum-specific RRs and corresponding 95% CIs. To study whether the effect of survey module varied with maternal or interviewer characteristics, we estimated site-specific interaction parameters and combined these in random-effects meta-analyses.

We restricted analyses of time taken to complete the maternity history section of the survey to interviews lasting 0·5–180 min to exclude implausible values. We summarised the distribution of times and mean times for survey module completion by survey module, HDSS site, and maternal and interviewer characteristics. We calculated the mean difference in time taken to complete the modules, with a corresponding 95% CI. We used Student's *t* test to compare time taken between the FBH+ and FPH groups because it is robust to departures from normality. We fitted a linear regression model to identify predictors of time taken to complete the survey modules (FBH+ *vs* FPH), with individual interviewers treated as random effects. We checked for multicollinearity between the independent variables by comparing the pairwise correlation coefficients. All covariates had a correlation of less than 0·5.

This study is registered with the Research Registry, 4720.

Role of the funding source

The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all data in the study and had final responsibility for the decision to submit for publication.

Results

Survey data were collected between July 28, 2017, and Aug 13, 2018, by 117 interviewers. Interviewers were predominantly female in Matlab (100%), Bandim (86%), and Dabat (100%), whereas they were mostly male in Kintampo (86%) and half were male in IgangaMayuge (50%). Interviewers had a mean age of 30·8 years, 90% had secondary or higher levels of education, and 80% had previous survey experience (appendix 2 p 12).

	FBH+ group (n=34 805)	FPH group (n=34 371)
Health and Demographic Surveillance System sites		
Bandim	4832 (14%)	4660 (14%)
Dabat	6327 (18%)	6266 (18%)
IgangaMayuge	6788 (20%)	6649 (19%)
Matlab	10 809 (31%)	10 653 (31%)
Kintampo	6049 (17%)	6143 (18%)
Median age (IQR)	28 (22–34)	28 (22–35)
Age, years		
15–19	5007 (14%)	4799 (14%)
20–24	7104 (20%)	7017 (20%)
25–29	7546 (22%)	7572 (22%)
30–34	6430 (19%)	6453 (19%)
≥35	8700 (25%)	8506 (25%)
Missing	18 (<1%)	24 (<1%)
Education level		
No education*	7851 (23%)	7687 (22%)
Primary only	10 860 (31%)	10 843 (32%)
Primary and secondary	13 078 (38%)	12 878 (38%)
Higher	3002 (9%)	2946 (9%)
Missing	14 (<1%)	17 (<1%)
Socioeconomic wealth quintile		
1 (poorest)	7235 (21%)	7087 (21%)
2	6829 (20%)	6608 (20%)
3	6825 (20%)	6948 (20%)
4	6891 (20%)	6926 (20%)
5 (richest)	7025 (20%)	6802 (20%)
Missing	0	0
Parity		
0	4733 (14%)	4662 (14%)
1	6878 (20%)	6692 (20%)
2	7238 (21%)	7019 (20%)
3	5272 (15%)	5185 (15%)
4	3395 (10%)	3541 (10%)
≥5	7289 (21%)	7272 (21%)
Missing	0	0

Data are n (%) or median (IQR). Definitions of variables used in this analysis are provided in appendix 2 (pp 9–10), as are details of background characteristics of the interviewers (p 13). FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. *Never attended school or madrasa.

Table 1: **Baseline characteristics of women in the Every Newborn-INDEPTH study (n=69 176)**

	FBH+ group		FPH group		p value
	Number/total*	Rate per 1000	Number/total*	Rate per 1000	
Stillbirths (n=66 649)					
Bandim	127/6291	20·2 (16·7–23·6)	153/5991	25·6 (21·5–29·5)	0·048
Dabat	41/4208	9·7 (6·7–12·7)	44/4172	10·5 (7·5–13·7)	0·71
IgangaMayuge	35/4324	8·1 (5·4–10·8)	79/4298	18·4 (14·3–22·4)	<0·0001
Matlab	176/10 786	16·3 (13·9–18·7)	140/10 533	13·3 (11·1–15·5)	0·067
Kintampo	129/7919	16·3 (13·5–19·1)	159/8127	19·6 (16·6–22·6)	0·12
Overall	508/33 528	15·2 (13·8–16·5)	575/33 121	17·4 (16·0–18·8)	0·024
Neonatal deaths (n=65 566)					
Bandim	223/6164	36·2 (31·5–40·8)	215/5838	36·8 (32·0–41·7)	0·85
Dabat	110/4167	26·4 (21·5–31·3)	101/4128	24·5 (19·8–29·2)	0·58
IgangaMayuge	126/4289	29·4 (24·3–34·4)	101/4219	23·9 (19·3–28·6)	0·12
Matlab	217/10 610	20·5 (17·8–23·2)	242/10 393	23·3 (20·4–26·2)	0·16
Kintampo	163/7790	20·9 (17·7–24·1)	158/7968	19·8 (16·8–22·9)	0·63
Overall	839/33 020	25·4 (23·7–27·1)	817/32 546	25·1 (23·4–26·8)	0·80

p value calculated using Student's *t* test to compare means (SBR or NMR) obtained between the FBH+ and FPH. FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. *For stillbirths, total refers to total births; for neonatal deaths, total refers to livebirths.

Table 2: **FBH+ versus FPH in the last 5 years of the EN-INDEPTH study, by study site and overall**

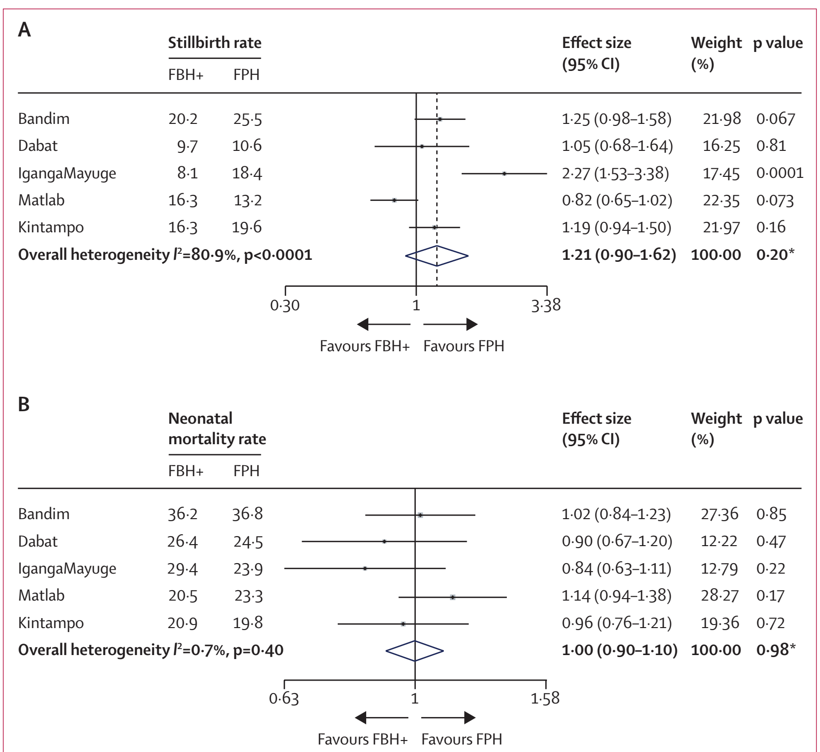

Figure 3: **Forest plot comparing stillbirth (A) and neonatal mortality (B) between FBH+ and FPH modules, by HDSS site**

Data are from random-effects analyses. We adjusted for clustering of neonatal deaths within individual women, showing overall and by the five HDSS sites. FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. HDSS=Health and Demographic Surveillance System. *Pooled p value obtained from point estimate and 95% CI.

98 187 women were identified as potentially eligible for inclusion in the EN-INDEPTH study across the five HDSS sites (figure 2, appendix 2 pp 13–16), and 89 731 women were selected to participate in the EN-INDEPTH survey, of whom 11 559 (12·9%) were not traced. Migration was the most common reason for not being traced (n=9229). Of the 78 172 women who were traced, 8254 (10·6%) were not available for interview after three attempted household visits.

Of the 69 918 women contacted to participate in the survey, 742 (1·1%) refused. Of the 69 176 women who consented to participate, 34 371 were randomly allocated to the FPH group and 34 805 to the FBH+ group (figure 2). Survey completion was high (>99% of women) in both groups. The groups were well balanced in terms of background characteristics; therefore, we would not expect confounding to affect the observed results (table 1).

The EN-INDEPTH survey achieved 98% of the planned sample size (66 649 of 68 000 total births within 5 years before the date of interview). The number of births by HDSS site were 12 282 in Bandim, 8380 in Dabat, 8622 in IgangaMayuge, 21 319 in Matlab, and 16 046 in Kintampo (table 2, appendix 2 p 16).

The FBH+ module recorded 508 stillbirths from 33 528 total births (SBR 15·2 per 1000 total births) and the FPH module recorded 575 stillbirths from 33 121 total births (SBR 17·4 per 1000 total births) in the 5 years preceding the survey interview (table 2). Across HDSS sites, the crude SBR ranged from 8·1 to 20·2 per 1000 total births in the FBH+ group and from 10·5 to 25·6 per 1000 total births in the FPH group (table 2). 11 women in the FBH+ group and 24 women in the FPH group reported more than one stillbirth in the preceding 5 years (appendix 2 p 17).

Combining data across sites in a meta-analysis with random effects showed that, on average, the estimated SBR was 21% (95% CI –10 to 62) higher in the FPH group than in the FBH+ group (p=0·20; figure 3A). There was strong evidence of heterogeneity across the HDSS sites (p<0·0001), with an I^2 of 80·9%. Four of the five sites reported higher SBRs in the FPH group than in the FBH+ group. In Matlab, SBR was lower for FPH than for FBH+. A much higher SBR was seen in the FPH group in IgangaMayuge than in other sites.

In our stratum-specific analysis, there was strong evidence of heterogeneity across the HDSS sites (p<0·0001) even after adjusting for maternal and interviewer characteristics with an I^2 of 81·3% (appendix 2 p 17). We found little evidence of interaction between survey module and any maternal or interviewer covariates (table 3). The variables for which there was the most evidence of interaction, albeit weak, were interviewer gender and maternal education (table 3).

The FBH+ module recorded 839 neonatal deaths from 33 020 livebirths (NMR 25·4 per 1000 livebirths) and the FPH recorded 817 neonatal deaths from 32 546 livebirths

	FBH+ group (base), n	FPH group, n	Stratum-specific RR (95% CI)	Stratum-specific between-site heterogeneity I^2; p value	Overall between-site heterogeneity I^2; p value	Module covariate interaction parameters (95% CI); p value*
Respondent characteristics						
Overall	508	575	1·21 (0·90–1·62)	..	80·9%; p<0·0001	..
Age, years						
<30	253	277	1·19 (0·85–1·66)	70·0%; p=0·010	..	..
≥30	255	298	1·27 (0·86–1·86)	76·7%; p=0·0020	70·5%; p<0·0001	1·08 (0·67–1·74); p=0·77
Education level						
Not educated and primary	270	351	1·38 (0·96–1·98)	77.8%; p=0·0010	..	..
Secondary and higher	238	224	0·98 (0·74–1·28)	70.2%; p=0·13	66·2%; p=0·0030	0·73 (0·49–1·10); p=0·13
Socioeconomic wealth quintile						
Poor (1, 2, and 3)	338	376	1·25 (0·85–1·83)	83·2%; p<0·0001	..	..
Rich (4 and 5)	170	199	1·18 (0·96–1·45)	0·0%; p=0·77	65·2%; p=0·0020	1·02 (0·77–1·36); p=0·90
Parity						
<4	274	285	1·14 (0·79–1·64)	74·0%; p=0·0040	..	..
≥4	234	290	1·28 (0·91–1·81)	69·7%; p=0·010	69·4%; p=0·0010	1·07 (0·73–1·56); p=0·74
Interviewer characteristics						
Gender						
Male	139	212	1·76 (1·01–3·07)	75·1%; p=0·018	..	..
Female	361	354	1·06 (0·83–1·34)	48·6%; p=0·10	70·0%; p=0·0010	0·68 (0·43–1·08); p=0·10
Age, years						
<30	218	221	1·05 (0·80–1·38)	43·9%; p=0·13	..	..
≥30	282	345	1·24 (0·84–1·82)	79·8%; p=0·0010	68·3%; p=0·0010	1·15 (0·79–1·66); p=0·47
Education level						
Primary or secondary	134	180	1·35 (1·07–1·69)	0·0%; p=0·72	..	..
Higher	359	381	1·17 (0·81–1·68)	80·9%; p<0·0001	71·5%; p=0·0010	0·81 (0·59–1·12); p=0·20
Reported previous experience with DHS or MICS						
Yes	367	397	1·20 (0·86–1·68)	78·8%; p=0·0010	..	..
No	133	169	1·28 (1·02–1·61)	0·0%; p=0·57	64·5%; p=0·0040	1·06 (0·77–1·47); p=0·74
Reported previous experience with non-DHS or non-MICS						
Yes	329	370	1·15 (0·90–1·48)	58·4%; p=0·047	..	..
No	171	196	1·30 (0·79–2·14)	77·1%; p=0·0020	66·7%; p=0·0010	1·14 (0·73–1·77); p=0·57

Base refers to the reference category. DHS=Demographic and Health Survey. FBH+= full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. MICS=Multiple Indicator Cluster Survey. RR=risk ratio. *Evidence that the effect of survey module varied with maternal or interviewer characteristics was sought by estimating site-specific interaction parameters and combining these in a random-effects meta-analysis. Details of background characteristics of the interviewers are provided in appendix 2 (p 12). Interviewer information missing for two interviewers (eight stillbirths in FBH+ and nine stillbirths in FPH). Education status missing for five interviewers (15 stillbirths in FBH+ and 14 stillbirths in FPH).

Table 3: **Stratum-specific risk ratios for stillbirths by survey module in the Every Newborn-INDEPTH study (n=69 176)**

(NMR 25·1 per 1000 livebirths) in the 5 years preceding the interview (table 2). Across HDSS sites, the crude NMR ranged from 20·5 to 36·2 per 1000 livebirths in the FBH+ group and from 19·8 to 36·8 per 1000 livebirths in the FPH group (table 2). 66 women in the FBH+ group and 57 in the FPH group reported more than one neonatal death in the last 5 years (appendix 2 p 18).

Combining neonatal death data across sites showed that estimated NMRs did not differ between the FBH+ and FPH groups (difference 0%, 95% CI –10 to 10; p=0·98; figure 3B). There was no evidence of heterogeneity between the HDSS sites (p=0·40), with an I^2 of 0·7%.

In our stratum-specific analysis, there was little evidence of heterogeneity between HDSS sites after stratification by woman's and interviewer's characteristics for all covariates (table 4). We found no evidence of interaction for any maternal or interviewer covariates with the survey modules. Reporting of neonatal deaths by interviewer characteristics did not differ between modules.

In our sensitivity analyses, after excluding Matlab, the estimated SBR was 35% (95% CI 3 to 77) higher in the FPH group than in the FBH+ group (p=0·030; appendix 2 p 18). Excluding IgangaMayuge gave an estimated SBR that was 6% (–15 to 32) higher in the FPH group than in the FBH+ group (p=0·60; appendix 2 p 19).

Among women with at least one recorded livebirth, 29 928 (99%) of 30 255 women reported the same number of livebirths in subsections 2.1 and 2.2. In the FPH group, 1566 (5·1%) of 30 873 women had a different number of livebirths and 1974 (6·4%) had a different

	FBH+ group (base), n	FPH group, n	Stratum-specific RR (95% CI)	Stratum specific between-site heterogeneity I^2; p value	Overall between-site heterogeneity I^2; p value	Module covariate interaction parameter (95% CI); p value*
Respondent characteristics						
Overall	839	817	1·00 (0·90–1·10)	..	0·0%; p=0·48	..
Age, years						
<30	446	465	1·03 (0·84–1·26)	49·9%; p=0·092	..	..
≥30	393	352	0·92 (0·78–1·07)	7·7%; p=0·36	38·1%; p=0·36	0·89 (0·65–1·22); p=0·48
Education level						
Not educated and primary	610	616	0·91 (0·74–1·12)	0·0%; p=0·46	..	..
Secondary and higher	229	201	1·00 (0·84–1·19)	50·6%; p=0·088	32·9%; p=0·16	1·12 (0·74–1·68); p=0·60
Socioeconomic wealth quintile						
Poor (1, 2, and 3)	532	535	1·00 (0·83–1·21)	53·0%; p=0·074	..	..
Rich (4 and 5)	307	282	0·97 (0·77–1·21)	37·8%; p=0·17	41·8%; p=0·079	0·97 (0·68–1·39); p=0·88
Parity						
<4	441	458	1·01 (0·82–1·24)	50·5%; p=0·089	..	..
≥4	398	359	0·90 (0·75–1·09)	30·6%; p=0·22	43·9%; p=0·066	0·90 (0·62–1·31); p=0·59
Interviewer characteristics						
Gender						
Male	253	232	0·90 (0·72–1·12)	23·5%; p=0·27	..	..
Female	571	579	1·05 (0·93–1·19)	0·0%; p=0·64	0·0%; p=0·46	1·19 (0·86–1·66); p=0·30
Age, years						
<30	384	372	0·94 (0·74–1·21)	58·0%; p=0·049	..	..
≥30	440	439	1·03 (0·90–1·18)	0·0%; p>0·99	7·6%; p=0·37	1·08 (0·82–1·41); p=0·59
Education level						
Primary or secondary	238	229	1·00 (0·82–1·20)	0·0%; p=0·37	..	..
Higher	581	556	1·00 (0·89–1·13)	0·0%; p=0·43	0·0%; p=0·56	0·95 (0·72–1·24); p=0·73
Reported previous experience with DHS or MICS						
Yes	604	597	1·02 (0·91–1·15)	0·0%; p=0·43	..	..
No	220	214	0·98 (0·81–1·20)	0·0%; p=0·63	0·0%; p=0·69	0·96 (0·73–1·27); p=0·79
Reported previous experience with non-DHS or non-MICS						
Yes	488	507	1·05 (0·90–1·23)	28·5%; p=0·23	..	..
No	336	304	0·94 (0·80–1·10)	0·0%; p=0·75	0·0%; p=0·45	0·89 (0·72–1·10); p=0·29

Base refers to the reference category. DHS=Demographic and Health Survey. FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. MICS=Multiple Indicator Cluster Survey. RR=risk ratio. *Evidence that the effect of survey module varied with maternal or interviewer characteristics was sought by estimating site-specific interaction parameters and combining these in a random-effects meta-analysis. Details of background characteristics of the interviewers are provided in appendix 2 (p 12). Interviewer information missing for six interviewers (15 neonatal deaths in FBH+ and six neonatal deaths in FPH). Education status missing for seven interviewers (20 neonatal deaths in FBH+ and 22 neonatal deaths in FPH).

Table 4: **Stratum-specific RRs for neonatal deaths by survey module in the Every Newborn-INDEPTH study (n=69 176)**

number of pregnancy losses recorded in subsections 2.1 and 2.2. These inconsistencies were most noticeable in Matlab, where 1546 (14·5%) of 10 653 women had a different number of pregnancy losses recorded in each subsection; 1513 (98%) of these differences were due to more pregnancy losses being recorded in the summary (subsection 2.1) than in the line listing of pregnancies (subsection 2.2). In all other sites, 428 (<3%) of 23 718 women had more pregnancies recorded in subsection 2.1 than in the line listing of pregnancies (subsection 2.2): Bandim, 47 (1%) of 4660 women; Dabat, 61 (1%) of 6266 women; IgangaMayuge, 127 (2%) of 6649 women; and Kintampo, 193 (3%) of 6143 women. In IgangaMayuge, where the largest difference in reported SBRs between the two modules was observed, perinatal mortality rates (ie, both stillbirths and early neonatal deaths) were 23% (95% CI –3 to 56) higher in the FPH group than in the FBH+ group (appendix 2 pp 21–22).

As a result of restricting our analyses of time taken to complete the maternity history section of the survey to interviews lasting 0·5–180 min to exclude implausible values, we excluded 1·5% of women with implausible values less than 0·5 min (222 in FBH+ and 740 in FPH) and 0·6% of women with implausible values greater than 180 min (180 in FBH+ and 199 in FPH). The mean time taken to complete the reproduction module was longer for FPH (mean 10·5 min) than for FBH+ (9·1 min, p<0·0001; table 5, figure 4), a difference of 1·4 min.

The mean time taken to complete both the reproduction modules (FBH+ and FPH) varied by study site and parity (table 5). After adjusting for parity and HDSS site of the woman, the mean time taken to complete the maternity history section of the survey was 1 min longer (mean 1·23 min [95% CI 1·11–1·36]) in the FPH group than in the FBH+ group (p<0·0001; appendix 2 p 19).

Discussion

Stillbirths and neonatal deaths remain a major preventable burden. Data gaps, especially in LMICs, have masked the issue and reduced the attention given, despite its major effect on families and particularly women's mental health.[7,20] We did this direct randomised comparison to investigate whether using an FPH approach would result in higher estimates of SBRs than using the standard FBH+ approach that is used in surveys in most LMICs. This study, which recorded 66 649 births across five HDSS sites found that on average SBRs were 21% (95% CI –10 to 62) higher with FPH than with FBH+. We found no evidence of a difference in NMR estimates between the two approaches. On average, the FPH survey took 1·4 min longer per woman than did the FBH+ survey.

In both study groups, estimated NMRs were similar to national-level NMR estimates.[1,11] The FBH+ module used in DHS or MICS model questionnaires was developed for improving measurement of child and infant mortality (including neonatal mortality), and our results were not expected to differ from the nationally comparable estimates computed from DHS and MICS (appendix 2 p 20). The NMRs in each HDSS site were similar in magnitude to national NMR estimates, although these sites are not nationally representative, given that they have factors that might both increase NMR (eg, more rural settings) and decrease NMR (eg, research studies and health facility strengthening). By contrast, the SBRs in both study groups were lower than national-level SBR estimates for most sites. Simple tests of data quality suggest that, in both groups, estimated SBRs were lower than might be expected (eg, SBR:NMR was ≤1 compared with an expected ratio of ≥1·2; appendix 2 p 20).[2,21] Hence, even with increased reporting of stillbirths by the FPH method, the estimated SBR is still lower than expected.

Omission and misclassification are key challenges for accurate perinatal data collection across all data systems. Variation in the level of omission and misclassification of events by site might explain some of the lower SBRs, and also the observed differences between sites in SBR. We found substantial heterogeneity between sites. Four of the five sites found higher SBRs in the FPH group than in the FBH+ group. By contrast, in Matlab, the SBR in the FPH group was lower. IgangaMayuge was an outlier, reporting a much higher SBR in the FPH group than did other sites.

Omission involves non-reporting or removal (intentional or unintentional) of events that are asked about during a survey interview.[14,22] This might be especially pronounced for pregnancy outcomes, such as stillbirth, which are frequently associated with stigma and shame for women who experience them.[23,24] We found that the number of births differed between the summary history (subsection 2.1) and the full history (subsection 2.2) in the FPH. These differences were most noticeable in the Matlab site. Some of these differences could be accounted for by differences in training. In all sites except Matlab, the EN-INDEPTH survey team from the London School of Hygiene & Tropical Medicine and Makerere University School of Public Health participated in the training of the data collectors and supervisors. In Matlab, training was done in-house with shorter

	FBH+ group	FPH group	Mean difference (95% CI)	p value
Overall	9·13 (9·76)	10·52 (10·25)	1·39 (1·23 to 1·54)	<0·0001
Health and Demographic Surveillance System sites				
Bandim	8·41 (7·43)	10·16 (8·37)	1·74 (1·38 to 2·11)	<0·0001
Dabat	8·24 (10·10)	8·72 (9·99)	0·48 (0·11 to 0·83)	<0·0001
IgangaMayuge	8·94 (11·43)	11·42 (14·72)	2·48 (2·01 to 2·96)	<0·0001
Matlab	9·01 (7·26)	10·17 (8·09)	1·15 (0·95 to 1·36)	<0·0001
Kintampo	10·99 (9·38)	12·03 (9·44)	1·31 (0·96 to 1·65)	<0·0001
Parity				
0	2·07 (5·20)	2·08 (6·32)	0·01 (–0·26 to 0·27)	0·96
1	5·12 (5·23)	5·85 (5·29)	0·73 (0·54 to 0·91)	<0·0001
2	7·42 (6·48)	8·37 (6·30)	0·95 (0·73 to 1·17)	<0·0001
3	9·81 (7·61)	10·58 (6·58)	0·77 (0·49 to 1·05)	<0·0001
4	11·72 (7·86)	13·52 (10·47)	1·80 (1·34 to 2·25)	<0·0001
≥5	17·28 (11·41)	19·38 (12·97)	2·10 (1·69 to 2·52)	<0·0001

Data are mean min (SD) unless otherwise specified. We calculated mean difference as FPH minus FBH+. We excluded 1·6% of women with implausible values <0·5 min (222 in FBH+ and 740 in FPH) and 0·6% of women with implausible values >180 min (180 in FBH+ and 199 in FPH). We did linear tests for trend; evidence for a trend was found (p<0·0001; data not shown). FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module.

Table 5: **Unadjusted analysis of response times for FBH+ and FPH survey modules by site and parity for the Every Newborn-INDEPTH study (n=65 511)**

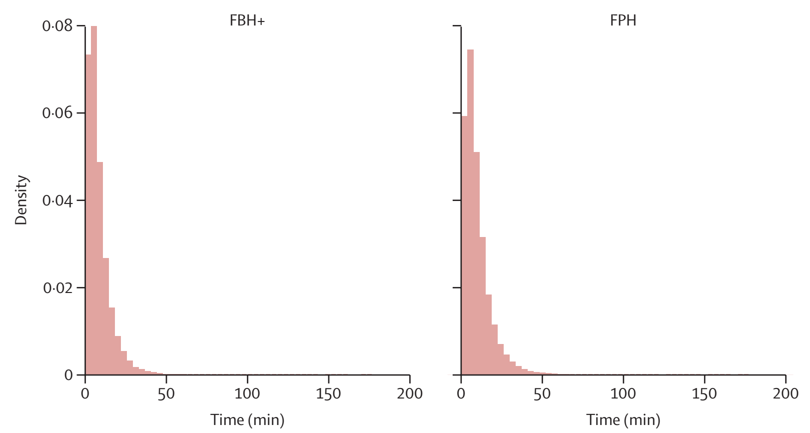

Figure 4: **Time taken to complete questions for the two survey modules (n=65 511) in the Every Newborn-INDEPTH study**
We restricted analyses of time taken to complete the maternity history section of the survey to interviews lasting 0·5–180 min to exclude implausible values. FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module.

overall training, with less training time on the FPH (around 2–3 h) compared with the standard training in other sites (>1 day), and less emphasis was placed on checking and correcting between the summary and FPH sections (appendix 2 p 20). The tablet-based app design included an automatic error message in the FBH+ module if the total number of reported births in subsection 2.1 and recorded births in subsection 2.2 did not match. This error message was not programmed in the FPH. Assuming that women reported the correct number of events in subsection 2.1, the total number of lifetime pregnancy losses in the FPH in Matlab could be underestimated by around 14·5%, but as no further details were collected on these losses it is not possible to determine how many were stillbirths (≥7 months) in the 5 years preceding the survey. A major focus of interviewer training for FPH was consistency checks between subsections 2.1 and 2.2. Hence, differences in training in Matlab could partly explain the differing performance of FPH there, highlighting the importance of consistent implementation and of incorporating automatic error checks and messages when programming electronic data collection devices. Our experience in Matlab shows the need for future surveys that use the FPH module to emphasise checking of consistency between the summary history (subsection 2.1) and the full history (subsection 2.2).

Misclassification between stillbirths and early neonatal deaths is likely to be common in household surveys. A Malawian study[25] reported that a fifth of neonatal deaths identified in a household survey were classified as stillbirths on verbal autopsy. In an Afghanistan survey,[26] in addition to neonatal deaths being misclassified as stillbirths, a small number of stillbirths were misclassified as either miscarriages or neonatal deaths. Misclassification of outcomes might explain part of the large positive effect of FPH in IgangaMayuge, where the SBR was 127% higher and the NMR 16% lower in the FPH group than in the FBH+ group, a much greater difference than seen in other sites. However, this cannot completely explain the results because perinatal mortality was higher in the FPH group than in the FBH+ group. This result is surprising because more misclassification of neonatal deaths as stillbirths might have been expected in the FBH+ group, in which women were required to decide themselves whether a baby was born alive and then died or was born dead. By contrast, in the FPH group, women who reported that a baby was born dead were asked, "Did the baby cry, move, or breathe when it was born?", in an effort to reduce misclassification of neonatal deaths as stillbirths. Further research is required to better understand misclassification and how it could be reduced, particularly in community contexts with no gold standard measure (such as heart rate at birth) to aid differentiation.

Gender-related interviewer dynamics could affect the reporting of stillbirths in an interview. We found that reported SBRs differed between FBH+ and FPH modules for male interviewers (RR 1·76, 95% CI 1·01–3·07) but not for female interviewers (1·06, 0·83–1·34). Interviewer gender did not affect the reporting of NMR.

A notable strength of this study was the randomisation to the different study groups and its sample size (69 176 women) in multiple and varied contexts across sub-Saharan Africa and South Asia.[15] Thus, our results are likely to be generalisable to LMICs, although because survey respondents were residents of the HDSS and accustomed to routine surveillance, this participation might have affected responses given. The randomisation of maternity modules allowed for direct comparison between FPH and FBH+ survey modules using robust methods. We accounted for clustering of pregnancy outcomes within individual women by using GEE models with exchangeable correlation matrices. Although it was not possible for the interviewers to be masked to the module type (FBH+ *vs* FPH), they were not aware of the study hypothesis and, thus, this knowledge is unlikely to have systematically biased the results. The definitions of stillbirth and neonatal deaths used were consistent with WHO definitions for international comparisons.[21] Additionally, we used standard DHS tools and, consistent with the DHS programme, analysed the results using the century month code dates so that the results would be directly applicable to DHS.[19] The tools and protocols we used were standardised and the survey was done using the same Survey Solutions platform across sites. Overall, the study was implemented with minimal deviations from the planned activities as per the study protocol[15] and with data monitoring every 14 days during data collection. We also collected data on duration of interviews, generating for the first time evidence on the mean time taken to complete the FBH+ and FPH modules.

This study has several limitations. First, as is the case for DHS, although we produced a standard interviewer manual for this study, interviewer and supervisor trainings in different sites were done by different trainers, and this disparity appears to have affected consistency. The interviewer manual was translated into each site's local language by local experts, which could have altered the original meaning and messages in the manuals. Second, we did not have a gold standard with which to compare the estimates obtained from FBH+ and FPH approaches. However, based on previous evidence that stillbirths are commonly under-reported in household surveys, it is more likely that the true population stillbirth rates are higher than those captured in either FBH+ or FPH, with FPH closer to true population levels. Third, the study was done in HDSS sites where women are accustomed to surveillance visits, which could heighten their awareness and recall of pregnancy outcomes. Lastly, the World Bank's Survey Solutions software we used produced various versions of the same software during the lifetime of this study, resulting in delays in data collection in some sites, and analytical challenges due to incompatibility of some of the attributes and variables between earlier versions and later versions—although these are unlikely to affect the overall result (Thysen SM, unpublished).

Additional analyses from EN-INDEPTH, including linking survey data to routine HDSS data, detailed qualitative work on barriers and enablers of reporting pregnancies and adverse pregnancy outcomes in household surveys, and a detailed analysis of experiences of implementation of this survey, will contribute to closing the knowledge gap in measurement of stillbirths and neonatal deaths, and will inform further survey improvements (Thysen SM, Akuze J, unpublished).[27]

In summary, FPH might improve the recording of stillbirths compared with FBH+, but FPH appears to have little, if any, effect on recording of neonatal deaths or on misclassification between stillbirths and early neonatal deaths.

The DHS programme has made the decision to make FPH the core module for the next phase of DHS.[28] However, switching from FBH+ to FPH will require retraining of interviewers who are used to FBH+. Further work is needed to review and develop improved guidance for the implementation of the FPH approach, including considering electronic checks between sections and reviewing interviewer prompts, translations, and training materials.

More investment is required to develop and implement better approaches for capturing information on stillbirths. However, improved measurement methods might fail without changes in social norms, including societal stigma around stillbirth. Most of these deaths are preventable, but ending preventable stillbirths will require altering prevailing perceptions that stillbirths are inevitable.[29] Wider investment in approaches to reduce stillbirths and neonatal deaths is needed to meet national targets for both by 2030 and to reduce this preventable burden on women, families, and society.[3]

Contributors

The Every Newborn-INDEPTH study was conceptualised by JEL. All site teams contributed to the design of the study protocol and undertook data collection. JA, together with HB, VSG, AB, SC, and JEL, developed the detailed research questions and overall analysis plan for this analysis. These were refined with inputs from the wider Every Newborn-INDEPTH study collaborator group at a multi-country workshop in Entebbe Uganda in February, 2019. JA analysed the data. SC provided statistical oversight. JA, HB, JEL, and SC drafted the manuscript. All authors reviewed and revised the manuscript. All authors agreed to the final version.

The Every Newborn-INDEPTH Study Collaborative Group

Tadesse Awoke Ayele, Telake Azale Bisetegn, Nafisa Delwar, Lemma Derseh Gezie, Collins Gyezaho, Judith Kaija, Kazuyo Machiyama, Grace Manu, Alexander A Manu, Justiniano S D Martins, Tesfahun Melese, Sayed S Alam, Tryphena Nareeba, Victoria Ponce Hardy, Charles Zandoh, Fred Arnold, Peter Byass, Trevor Croft, Kobus Herbst, Sunita Kishor, Florina Serbanescu, Turgay Unalan, Shane Khan, and Attila Hancioglu.

Declaration of interests

We declare no competing interests.

Data sharing

Data sharing and transfer agreements were jointly developed and signed by all collaborating partners. Other users can request our data through a written application.

To **make a written application** see https://doi.org/10.17037/DATA.00001556

Acknowledgments

This work was funded by the Children's Investment Fund Foundation by means of a grant to the London School of Hygiene & Tropical Medicine (Principal Investigator, JEL), and a sub-award to the INDEPTH Maternal Newborn Child Health working group managed in Accra and with technical leadership by the Makerere School of Public Health (Principal Investigator, PW). We thank the 117 interviewers and many HDSS staff participating in this study for their hard work and dedication to this project. Many thanks to Samuelina Arthur, Claudia DaSilva, Olivia Nakisita, and all site administrator staff for their administrative support. We also thank the EN-INDEPTH expert advisory group: Fred Arnold, Peter Byass, Trevor Croft, Kobus Herbst, Sunita Kishor, Florina Serbanescu, Stephane Helleringer, Turgay Unalan, Shane Khan, and Attila Hancioglu. Finally, and most importantly, we thank the women participating in the EN-INDEPTH study and their families, without whom this work would not have been possible.

Editorial note: the *Lancet* Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

References

1 UNICEF, WHO, World Bank, UN. Levels and trends in child mortality: report 2018. Geneva: World Health Organization, 2018.
2 Blencowe H, Cousens S, Jassir FB, et al. National, regional, and worldwide estimates of stillbirth rates in 2015, with trends from 2000: a systematic analysis. *Lancet Glob Health* 2016; **4:** e98–108.
3 WHO. Every Newborn: an action plan to end preventable deaths. Geneva: World Health Organization, 2014.
4 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. *Lancet* 2015; **386:** 1407–18.
5 Mikkelsen L, Phillips DE, AbouZahr C, et al. A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet* 2015; **386:** 1395–406.
6 Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *Lancet* 2014; **384:** 189–205.
7 Lawn JE, Blencowe H, Waiswa P, et al. Stillbirths: rates, risk factors, and acceleration towards 2030. *Lancet* 2016; **387:** 587–603.
8 Healthy Newborn Network. Roadmap to accelerate progress for every newborn in humanitarian settings 2020–2025. May 24, 2019. htttps://www.healthynewbornnetwork.org/resource/roadmap-to-accelerate-progress-for-every-newborn-in-humanitarian-settings-2020-2025/ (accessed Sept 3, 2019).
9 Cleland J. Demographic data collection in less developed countries 1946–1996. *Popul Stud* 1996; **50:** 433–50.
10 Demographic and Health Surveys. DHS model questionnaires 2018. https://dhsprogram.com/What-We-Do/Survey-Types/DHS-Questionnaires.cfm (accessed Dec 8, 2019).
11 Demographic and Health Surveys. DHS overview. 2017. https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm (accessed April 19, 2018).
12 Cleland J, Verma V. The world fertility survey: an appraisal of methodology. *J Am Stat Assoc* 1989; **84:** 756–74.
13 Espeut D, Becker S. The validity of birth and pregnancy histories in rural Bangladesh. *J Health Popul Nutr* 2015; **33:** 17
14 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, MD: ICF International, 2015.
15 Baschieri A, Gordeev VS, Akuze J, et al. "Every Newborn-INDEPTH" (EN-INDEPTH) study protocol for a randomised comparison of household survey modules for measuring stillbirths and neonatal deaths in five Health and Demographic Surveillance sites. *J Glob Health* 2019; **9:** 010901.
16 World Bank. Survey Solutions Computer-Assisted Personal Interviewing (CAPI)/Computer Assisted Web Interviewing (CAWI) platform. Versions 5.21, 5.22, 5.23, 5.24, 5.25, 5.26, 18.04 and 18.06. http://surveys.worldbank.org/capi (accessed March 12, 2019).

17 O'Donnell O, Van Doorslaer E, Wagstaff A, Lindelow M. Analyzing health equity using household survey data: a guide to techniques and their implementation. Washington, DC: World Bank, 2007.

18 Croft T. DHS data editing and imputation. 1991. https://dhsprogram.com/pubs/pdf/DHSG3/DHS_Data_Editing.pdf (accessed Feb 21, 2020).

19 Croft TN, Marshall AMJ, Allen CK, et al. Guide to DHS statistics. Rockville, MD: ICF, 2018.

20 de Bernis L, Kinney MV, Stones W, et al. Stillbirths: ending preventable deaths by 2030. *Lancet* 2016; **387:** 703–16.

21 WHO. Neonatal and perinatal mortality 2000: country, regional and global estimates. Geneva: World Health Organization, 2006.

22 Pullum TW, Becker S. Evidence of omission and displacement in DHS birth histories. DHS Methodological Reports No. 11. Rockville, MD: ICF International, 2014.

23 Heazell AE, Siassakos D, Blencowe H, et al. Stillbirths: economic and psychosocial consequences. *Lancet* 2016; **387:** 604–16.

24 Frøen JF, Cacciatore J, McClure EM, et al. Stillbirths: why they matter. *Lancet* 2011; **377:** 1353–66.

25 Liu L, Kalter HD, Chu Y, et al. Understanding misclassification between neonatal deaths and stillbirths: empirical evidence from Malawi. *PLoS One* 2016; **11:** e0168743.

26 Christou A, Dibley MJ, Rasooly MH, et al. Understanding country-specific determinants of stillbirth using household surveys: the case of Afghanistan. *Paediatr Perinat Epidemiol* 2019; **33:** 28–44.

27 Kwesiga D, Tawiah C, Imam MA, et al. Qualitative study of barriers and enablers to reporting pregnancy and adverse pregnancy outcomes in population-based surveys: EN-INDEPTH multi-country study. *BMC Popul Health Metr* (in press).

28 USAID. DHS-8 questionnaires: revision process and new content. https://dhsprogram.com/pubs/pdf/DHSM11/DHSM11.pdf (accessed Dec 10, 2019).

29 Flenady V, Wojcieszek AM, Middleton P, et al. Stillbirths: recall to action in high-income countries. *Lancet* 2016; **387:** 691–702.