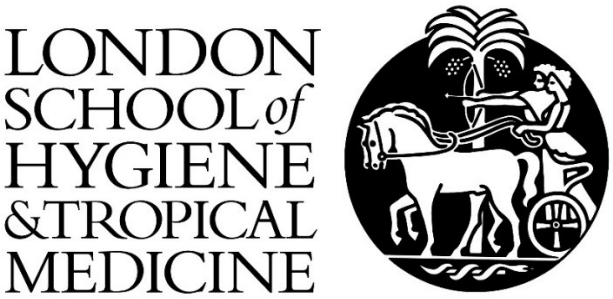


Measurement of stillbirths and neonatal deaths in standardised population-based surveys

Joseph Akuze Waiswa

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

of the

University of London

November 2021

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Work funded by Children's Investment Fund Foundation

Research group affiliation(s):

1. **Centre for Maternal Adolescent Reproductive & Child Health (MARCH), LSHTM**
2. **Maternal Newborn Health Group, LSHTM**
3. **Makerere University Centre of Excellence for Maternal Newborn and Child Health**

Declaration of own work

I, Joseph Akuze Waiswa, confirm that the work presented in this thesis is my own. Where information has been derived from other sources. I confirm that this has been indicated in the thesis

Signed:

Date: *26/05/2022*

Abstract

Background: Worldwide, 4.4 million stillbirths and neonatal deaths (SB&NND) are estimated to occur annually. Household surveys, notably the Demographic and Health Surveys (DHS), are an important source of SB&NND data. This PhD aimed to review the DHS's evolution for SB&NND data capture and compare the full birth history with additional questions on pregnancy losses (FBH+) and full pregnancy history (FPH) approaches in terms of data quality and potential measurement errors.

Methods: A literature and programmatic review of the DHS programme and HDSS pregnancy surveillance system was conducted. A randomised comparison of FBH+ versus FPH modules (the EveryNewborn-INDEPTH (EN-INDEPTH) survey) was undertaken in five health and demographic surveillance sites (HDSS). Reported stillbirth rates (SBR) and neonatal mortality rates (NMR); time for completion of survey modules; evidence for heterogeneity between sites; patterns of corrections between question types and structures, and modules were assessed. Survey data were compared to and HDSS data in four sites.

Results: Both FPH and FBH+ modules have been used within the DHS programme, but there is limited evidence concerning their accuracy for SB&NND. The FBH+ was the core model questionnaire between DHS-phases I-VII. A total of 69,176 women consented. 34,805 (50·3%) were randomised to FBH+ and 34,371 (49·7%) to FPH. There was little difference between the average time to administer questions in FBH+ (9.1 minutes) and FPH (10.5 minutes). The SBR was 15.2/1000 and 17.4/1000 total births for FBH+ and FPH, respectively. SBR was 21% (95% CI (-10% - 62%)) higher in FPH than in FBH+ with strong evidence of heterogeneity across the sites (I-squared=80·9% (p<0.001)). The NMR was similar in FPH (25.1/1000 births) and FBH+ (25.4/1000 births) with no evidence of heterogeneity between the sites (I-squared=0.0% (p=0.48)). Corrections were similar by survey module and occurred in 84% of survey interviews. Single corrections were the most common, multiple-select, and free-text questions increased response time by two minutes on average and had the most corrections. Pregnancy surveillance systems differed between sites (frequency of surveillance visits, main respondent reporting pregnancy outcomes, pregnancy testing, and data capture and surveillance modes). In three of the four sites, both arms of the EN-INDEPTH survey reported more pregnancies than the HDSS Matlab being the exception. Overall, the survey data produced higher estimated SBRs (FBH+: RR=1.13, 95%CI (0.79, 1.63), p=0.519; FPH: RR=1.20, 95%CI (0.76, 1.90), p=0.444) and NMRs (FBH+: RR=1.19, 95%CI (0.98, 1.43), p=0.071; FPH: RR=1.15, 95%CI (1.01, 1.30), p=0.030) than the HDSS data. In Matlab, the HDSS recorded more stillbirths and miscarriages than the survey in the FPH arm only. In Kintampo, the HDSS recorded more stillbirths than the survey, even though the HDSS recorded fewer pregnancies.

Conclusions: Evidence from this PhD of improved capture of stillbirths using the FPH influenced the DHS's switch to the FPH module in its eighth phase; however, limited evidence concerning SB&NND data accuracy and quality in standardised surveys is available. Therefore, additional efforts towards improved survey implementation of the FPH approach, including training, interviewer prompts, translations, and developing and testing standard data quality criteria for SB&NND in surveys and routine data, are warranted.

Acknowledgements

Above all, my appreciation is to the Lord for giving me life and the ability to accomplish this work.

Enrolling and studying in LSHTM's prestigious PhD programme has been a privilege as an African student. This journey was made possible with funding from the Children's Investment Fund Foundation (CIFF), United Kingdom, through the LSHTM. This PhD journey would have been impossible without my supervisors' continuous support and guidance (Hannah Blencowe, Simon Cousens, and Peter Waiswa) and the PhD advisory committee members (Joy Lawn, Angela Baschieri and Vladimir Gordeev). Thank you for guiding me to making my mark in measurement for stillbirth and neonatal deaths.

Peter, thank you for taking a chance on me ten years ago as a young novice undergraduate student seeking to learn how to do research and for your continuous encouragements towards career growth and academic training. You have played various roles in my life but above all, thank you for opening doors and opportunities for me and for introducing me to Joy Lawn, Hannah Blencowe and Simon Cousens.

Thank you, Hannah, and Simon, for being my friends, encouraging me throughout this PhD, and not getting tired to answer my unending questions and requests to review and comment on my work. You have both guided me into becoming a better scholar and writer, I look up to you both, and I hope I can measure up to you one day. Joy, your mix of work, mentoring and London based parties and luncheons made me have the most enjoyable time during my visits at LSHTM. Thank you for supporting your fellow Ugandan and African, you always gave me the kindest words or encouragement and wisdom. Hannah, Peter, and Joy, I am at a loss of words, but want to say that I am grateful for your financial support when my funding ended, for without you finding alternative sources of funding for me, this PhD would not have been achieved. Thank you, Angela, for being a friend and for your supervision/guidance during the first year and for helping me navigate the initial steps in the PhD especially when we were designing the study. Thank you, Vladimir, for being a friend and mentoring me above all, I still remember our late-night talks and walks in Gondor, Ethiopia.

I thank Lule Abubaker Swaib and Simon Kigozi, my LSHTM colleagues who always inspired me throughout this journey for without your continued words of encouragement and guidance on how to become a better doctoral student, I would have been lost. To my other LSHTM colleagues (Robert Kalyesubula, David Katende, Themba Mzende, Leslie Edwards, and Darshini Govindasamy), thank you for being a part of my PhD journey. Thank you, Claudia Da Silva, Dorothy Boggs, and Georgia Gore-Langton for always looking out for me and for making my visits to London very enjoyable. Special thanks go to my colleagues and collaborators on the EN-INDEPTH study, you are so many but special thanks go to Doris Kwesiga, Judith Yargawa, Vicki Ponce Hardy, Simon Kasasa, Moin Haider, Ane Fisker,

Francis Dzabeng, Davis Natukwatsa and others, your support with reviews and clarifying contextual issues was highly valued. Thank you, Kimberly Peven and Doris Kwesiga, for your unwavering support during this PhD work.

Thank you, Makerere University School of Public Health, Health Policy, Planning and Management department for providing me with a conducive atmosphere to conduct my PhD research while in Uganda. Thank you colleagues and friends your inspiration and words of encouragement, you are many, but special thanks go to Elizabeth Ekirapa, Philip Wanduru, Ayub Kakaire, Monica Okuga, Darius Kajjo, Gertrude Namazzi, Paul Mubiri, Rornald Kananura, Elizeus Rutebemberwa, Moses Tetui and Aloysious Ssenyonjo.

I am grateful to Kananura Rornald, Philip Mwondha, and Muzafaru Rubanga for being my brothers away from home and for making my stays in London enjoyable and memorable.

Finally, and most importantly, my family members; Phoebe Hilda Alitubeera my partner, thank you for being patient with me and for encouraging me throughout this PhD; our children Alystair Israel Akuze and Araminta Isadora Akuze, I am sorry for not giving you both your deserved attention during this time, I hope you will understand why one day. My brothers and sisters (Richard, Dennis, Henry Gorretie, Jackie, Winnie, and Immy), Mum, in-laws (Mr and Mrs Mwanja, Humphrey, Adrian, and Nigel) and nephews and nieces (Tamara, Bernice, Christian) thanks for the prayers and morale support.

Dedication

I dedicate this work to the special people in my life; Phoebe Hilda Alitubeera, Alystair Israel Akuze, Isadora Araminta Akuze, Mum Jane Francis Akuze, Richard Akuze, Dennis Akuze, Henry Akuze, Jackie Kalembe, Winnie Mukyala, Gorretie Nairuba, Immaculate Akuze, Joy Nakaibale, Dylan Louise Akuze, Declan Akuze, Dominic Akuze, Donald Gabula, Bella Nagawa, Ethan Ssegawa, Gabriella Upendo, Mitchell, Jesse Luzzi McArthur, and Ian Otim.

Supervisory and Advisory Committee members

Supervisors:

Associate Professor. Hannah Blencowe

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene and Tropical Medicine

Prof. Simon Cousens

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene and Tropical Medicine

Associate Professor. Peter Waiswa

Department of Health Policy Planning and Management

School of Public Health, College of Health Sciences

Makerere University

Advisory committee:

Prof. Joy Lawn

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene and Tropical Medicine

Associate Professor. Angela Baschieri

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene and Tropical Medicine

Dr. Vladimir Gordeev. S

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene and Tropical Medicine

Table of Contents

Declaration of own work i

Abstract ii

Acknowledgements iii

Dedication v

Supervisory and Advisory Committee members vi

Table of Contents viii

Table of Figures xi

Table of Tables xii

Abbreviations xiii

Definition of key terms as used in this thesis xiv

Chapter 1: Background 1

1.1. Introduction 1

1.2. Background 2

1.3. Data to estimate Stillbirth and Neonatal Mortality Rates 3

1.4. Estimation of Stillbirth Rate and Neonatal Mortality Rate from survey data 4

1.5. Review of potential biases and measurement errors in surveys 5

1.6. Conceptual framework for sources of biases and errors in surveys 6

1.7. Rationale 9

Chapter 2: Thesis aim, objectives, design and structure, candidate's contribution, and list of associated publications 12

2.1. Introduction 12

2.2. Thesis aim 12

2.3. Thesis objectives 12

2.4. Thesis structure 12

2.5. Candidate's contribution 14

2.5.1. EN-INDEPTH study 14

2.5.2. PhD work 21

2.6. Associated conferences and workshop presentations 23

2.7. Chapter summary 23

Chapter 3: DHS Historical Review of Demographic and Health Surveys 24

3.1. Introduction 24

3.2. List of Figures 24

3.3. List of Tables 24

3.4. Citation 25
3.5. Research paper 28
A: Four decades of measuring stillbirths and neonatal deaths in Demographic and Health Surveys: historical review 28
Chapter 4: Methods for the EN-INDEPTH study 43
4.1. Introduction 43
4.2. List of Figures 43
4.3. List of Tables 43
4.4. Citation 43
4.5. Research paper 47
B: "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 47
Chapter 5: Randomised comparison of the Full Birth History with additional questions on pregnancy losses and the Full Pregnancy History 63
5.1. Introduction 63
5.2. List of Figures 63
5.3. List of Tables 63
5.4. Citation 64
5.5. Research paper 67
C: Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in five countries: the Every Newborn-INDEPTH study 67
Chapter 6: Further assessments of effects of time taken for completion of FBH+ and FPH survey modules and data quality using Paradata 80
6.1. Introduction 80
6.2. List of Figures 80
6.3. List of Tables 81
6.4. Citation 81
6.5. Research paper 84
D: Paradata analyses to inform population-based survey capture of pregnancy outcomes: EN-INDEPTH study 85
Chapter 7: Population-level comparison of pregnancy outcomes between EN-INDEPTH survey and HDSS data 100
7.1. Introduction 100
7.2. List of Figures 100
7.3. List of Tables 100
7.4. Research paper 104

E: Comparing population-level pregnancy outcomes between population-based surveys and prospective demographic surveillance systems: an EN-INDEPTH study 104
Chapter 8: Discussion 137
8.1. Introduction 137
8.2. Overview of main findings 137
8.3. Generalisability of findings 144
8.4. Strengths of the PhD 145
8.5. Limitations of the PhD 147
8.6. Evidence synthesis - Measurement changes and implications now 155
8.7. Future research 157
8.8. Recommendations 159
8.9. Conclusion 161
References 162
Appendices 166
A1: Summary of the role of the candidate in the work presented in this thesis 166
A2: Supplementary methods, analyses and results to Chapter 7 171
A3: Ethical approvals 190
London School of Hygiene and Tropical Medicine: For Project and PhD study 190
London School of Hygiene and Tropical Medicine: Confirmation for PhD study 192
Bandim HDSS: 193
Dabat HDSS: 194
IgangaMayuge HDSS: 195
Kintampo HDSS 197
Matlab HDSS: 201
A6: Data sharing agreements 202
A2: HDSS women's Listing Manual 216
A3: Survey Solution's Headquarters Manual 236
A4: EN-INDEPTH Survey Implementation Guidelines 260
A7: Data management plan 292

Table of Figures

Figure 1.1: Definition of stillbirth and neonatal deaths 1
Figure 1.2: Conceptual framework showing potential measurement errors that affect the SBR and NMR estimates computed from retrospective survey data 8
Figure 7. 1: Linkage of pregnancy episodes captured by HDSS and EN-IDEPTH survey for four HDSS sites..... 122
Figure 7. 2: Distribution of pregnancy outcomes between EN-INDEPTH survey and HDSS by module 124
Figure 7. 3: Distribution of pregnancy outcomes between EN-INDEPTH survey and HDSS by module 125
Figure 7. 4: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to HDSS..... 126
Figure 7. 5: Ratio of Stillbirth Rate in EN-INDEPTH survey compared to HDSS..... 127
Figure 8.1: Conceptual framework showing potential measurement errors that affect the SBR and NMR estimates computed from retrospective survey data 139
Figure S 1a: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to HDSS by year 251
Figure S 1a: Ratio of Stillbirth Rate in EN-INDEPTH survey compared to HDSS by year 251
Figure S 2a: Percentage distribution of pregnancy outcomes between EN- INDEPTH survey and HDSS by module 253
Figure S 2b: Distribution of pregnancy outcomes between EN- INDEPTH survey and HDSS by module 253
Figure S 3a: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to HDSS..... 256
Figure S 3b: Ratio of stillbirth rate in EN-INDEPTH survey compared to HDSS.....254

Table of Tables

Table 1.1: Summary of available studies and data comparing FBH+ and FPH 11
Table 2.1: Overview of thesis chapters, research questions and methods 16
Table 7. 1: Overview of HDSS in five EN-INDEPTH study sites..... 118
Table 7. 2: Characteristics of continuously resident women included in the EN-INDEPTH survey, for four HDSS sites 121
Table 7. 3: Characteristics of interviewers who participated in the EN-INDEPTH survey, by sites, for four HDSS sites 121
Table 7. 4: Comparison of pregnancy outcome indicators between EN-INDEPTH survey and HDSS by module 123
Table 8. 1: EN-INDEPTH Data collection software/mobile data collection platform review 151
Table S 1: Definition of independent variables to be used in this analysis 240
Table S 2: Questions for assessing the HDSS system for pregnancy surveillance 241
Table S 3: Reasons for exclusion of data by year from the supplementary analysis..... 242
Table S 4: Overview of HDSS in five EN-INDEPTH study sites..... 243
Table S 5a: Comparing pregnancy outcome between EN-INDEPTH survey and Health Demographic Surveillance Sites by module and year 246
Table S 5b: Comparison of pregnancy outcome indicators between EN-INDEPTH survey and HDSS by module (Truncated)..... 247

Abbreviations

BA	Backward Approach
CPS	Contraceptive Prevalence Surveys
DHS	Demographic Health Surveys
ENAP	Every Newborn Action Plan
EN-INDEPTH	Every Newborn – INDEPTH study
FA	Forward Approach
FBH	Full Birth History
FBH+	Full Birth History with additional questions on pregnancy losses as used in DHS phase seven
FPH	Full Pregnancy History
FPH	Full Pregnancy History
GDHS	Ghana Demographic Health Survey
GMHS	Ghana Maternity Health Survey
HDSS	Health and Demographic Surveillance Sites
INDEPTH	International Network for the Demographic Evaluation of Populations and their Health
LMIC	Low and Middle-Income Countries
LSHTM	London School of Hygiene and Tropical Medicine
MakSPH	Makerere University School of Public Health
MDG	Millennium Development Goals
MICS	Multiple Indicator Cluster Surveys
MNCH-WG	Maternal Newborn Child Health Working Group
NMR	Neonatal Mortality Rate
PNMR	Perinatal Morality Rate
RHS	Reproductive Health Surveys
SBR	Stillbirth Rate
SDG	Sustainable Development Goals
SEA	South and South-Eastern Asia
SMS	Safe Motherhood Survey
SSA	Sub-Saharan Africa
WFS	World Fertility Surveys
WHS	World Health Surveys

Definition of key terms as used in this thesis

Stillbirth	A fetal death with reported gestational age of ≥7 months
Neonatal death	A death before 28 days after birth
Data linkage	Individual or population level linkage/merging of similar data from two different sources using either deterministic or probabilistic algorithms
Maternity histories	Full Birth history and Full Pregnancy history survey modules
Omission	Entire removal or not reporting of stillbirths and neonatal deaths in the survey
Displacement	Transferring of dates of birth beyond the threshold or period of inquiry
Response time	Time taken to complete questions in either the birth history or pregnancy history
Standardised survey	Surveys conducted in more than one site with similar tools, protocols for data collection
Pregnancy losses	Occurrence of stillbirth or miscarriage or abortions or pregnancy terminations

Chapter 1: Background

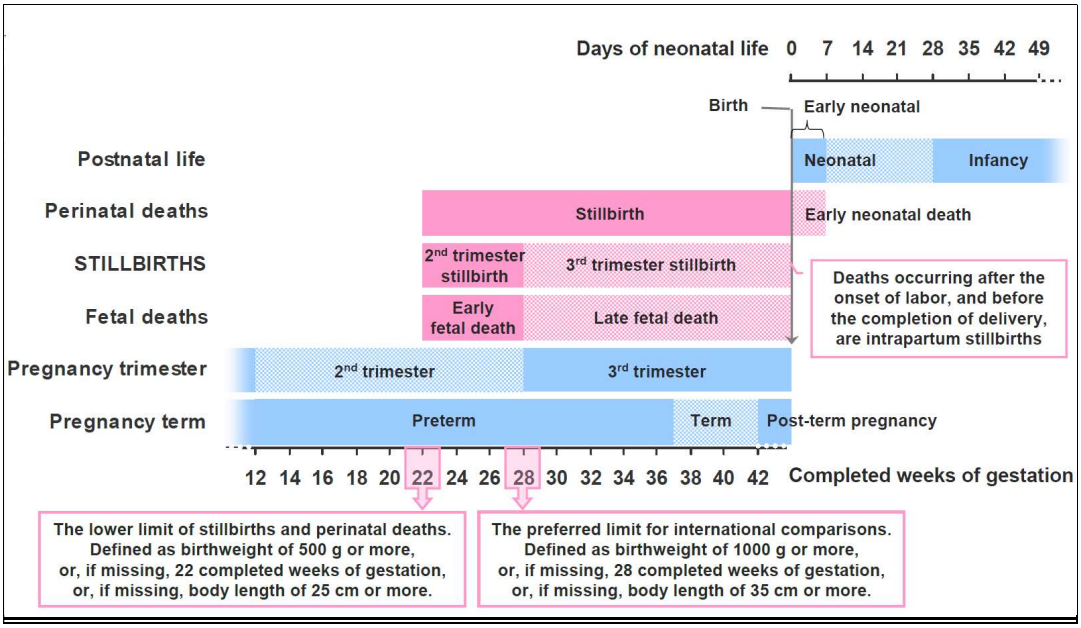
1.1. Introduction

This thesis examines the measurement of stillbirths and neonatal deaths in standardised surveys. The focus is on measurement, including comparisons of data collection approaches and methodologies, assessments of data quality and timing/length of interview, and comparisons of survey data with Health and Demographic Surveillance Sites' data. Neonatal mortality indicators have shown some progress in various low and middle-income countries (LMIC), but stillbirth data and indicators are often not collated or even collected in many LMICs data systems including surveys. Furthermore, varying definitions for stillbirths are used in many settings and contects.[1]

Defining stillbirths and neonatal deaths

For purposes of this PhD thesis, **stillbirths** are defined as "fetal deaths or pregnancy losses at or after 7 months of gestation", whereas **neonatal death** is defined as "the death of a live-born infant before the 28th day of life." The terms "neonate" and "newborn" are interchangeably used in many settings (see Figure 1.1); in this PhD, the term neonate is used. Other types of pregnancy outcome, including preterm births and congenital abnormalities, are beyond this PhD study's scope.

Figure 1.1: Definition of stillbirth and neonatal deaths



Source: J Frederik Frøen et al., Making stillbirths count, making numbers talk - Issues in data collection for stillbirths (reproduced with permission from the author)

1.2. Background

Every year, nearly 2.0 million babies are stillborn and 2.4 million die in the first 28 days of life.[2] LMIC in Sub-Saharan Africa (SSA), and South and South-Eastern Asia (SEA) account for approximately 98% of the global stillbirths and neonatal deaths.[3,4] Of the 166 million unregistered births, 87% are in SSA and SEA. Furthermore, less than 5% of all neonatal deaths that happen globally are estimated to have a death certificate.[5,6] Stillbirth information and data capture are worse than for neonatal deaths.[2]

Stillbirths and neonatal deaths may have lifelong psychological and social effects on the mothers who experience them. They share some causes, which differ from deaths that happen beyond the neonatal period.[7] Decreases in estimates of infant and child mortality rates reflect socioeconomic development including improved nutrition, housing and primary healthcare services of a country, while reduced stillbirth rates (SBR) and neonatal mortality rates (NMR) suggest evidence of improvements in the health status of women and quality of care during the perinatal period.[7]

During the Millennium Development Goals (MDG) period (1990-2015), global reductions in number of stillbirths and neonatal deaths were observed with an estimated annual reduction rate of 2.0% and 3.1% for stillbirths and neonatal deaths, respectively.[8] Despite progress in reducing these deaths, SBR and NMR remain high in many settings, hence their continued interests during the Sustainable Development Goals (SDGs) and adoption of The Every Newborn Action Plan (ENAP) to end preventable deaths.[3,4,9] The SDGs and ENAP goals target a NMR of less than or equal to 12 newborn deaths per 1,000 live births and a SBR of less than or equal to 12 stillbirths per 1,000 total births in every country by 2030.[4,9] Large data gaps and measurement issues need to be addressed to enable tracking of progress towards these goals.[3,10]

High-income countries have well-established civil and vital registration statistics (CRVS) systems where stillbirths and neonatal deaths are captured. Their LMIC counterparts often have unreliable data on these deaths captured by the CRVS systems or other data systems; therefore, data and knowledge gaps on these indicators exist. LMIC often rely on longitudinal demographic surveillance systems and population-level surveys like the Demographic Health Surveys (DHS) to collect data on these indicators.[11] Even with the existence of surveys, the data and knowledge

gaps exist due to the differing application of the standardised definitions, tools and different protocols used in surveys. For example, countries (for the DHS programme) may use different maternity histories and procedures (backward and forward or back truncated histories) when collecting data.[11-13] With these differences, the SBR and NMR indicators can still be reconstructed.[14] However, there is a need to check their comparability and the associated measurement errors.[15]

1.3. Data to estimate Stillbirth and Neonatal Mortality Rates

Estimation of SBR and NMR has remained problematic globally, despite the enormous improvements observed in longitudinal surveillance systems[16], survey design, data collection, and implementation witnessed over the last 50 years in LMICs.[11]

LMICs depend on several systems when computing their subnational and national estimates.[11] Examples include; Health and Demographic Surveillance Sites (HDSS), censuses and standardised retrospective surveys like the Demographic Health Surveys (DHS), World Health Surveys (WHS), Reproductive Health Survey (RHS), Multiple Indicator Cluster Surveys (MICS) and World Fertility Surveys (WFS). These surveys are valuable in providing estimates on the health and well-being of populations in low-resource settings. However, they face challenges, such as; sampling and non-sampling errors and incomplete information and data.[11,17] Data on stillbirths and neonatal deaths are gathered in surveys using maternity histories.[12] In the WFS and DHS surveys, complete maternity histories were introduced over three decades ago to gather retrospective data on women's reproduction.[11]

The initial efforts to obtain data for indicator estimates from surveys focused on data collection methods and techniques such as; comprehensive training and supervision of interviewers on how to conduct interviews, probing, and data entry.[11] It was proposed that survey measurement could be improved through introduction and rewording of questions in retrospective surveys.[18] These changes have led to improvements in information on neonatal deaths and stillbirths collected in surveys. However, challenges remain, including the reliance on women's reports about events that happened earlier on in their lives, leading to unconscious or deliberate omission, or misreporting of the date of events. Theoretically, some surveys on maternal and

reproductive health seek to capture all child deaths and stillbirths. However, in practice, neonatal deaths, stillbirths, miscarriages, and abortions are not fully captured because of failure to recall the events or timing when they occurred, or not understanding the difference between stillbirths and livebirths or miscarriages, or concerns regarding stigma in the community and associating adverse outcomes with evil spirits.[19,20]Additionally interviewer bias may be introduced if interviewers avoid probing or misreport/omit events as a means to finish the interview faster[21] introducing omission and displacement errors to surveys.

1.4. Estimation of Stillbirth Rate and Neonatal Mortality Rate from survey data

Direct and indirect estimation methods are used to compute mortality estimates in standardised surveys. Direct estimation methods utilise reported data from surveys on survival, for example, data on date of birth, pregnancy termination or death. Direct mortality estimation has been used within the WFS [22,23], and DHS [14] surveys and has been widely adopted by many other population-level surveys globally. Indirect methods, such as standardised mortality ratios (SMR) and indirectly standardised mortality ratios, have been used for under-five mortality, but not for stillbirths or neonatal deaths.[24]

Three alternative approaches for estimating SBR and NMR from survey data have been used under the direct methods[14]

a) **Vital statistics** is an approach where the number of neonatal deaths or stillbirths is divided by the number of births in the same period. This approach provides us with a rate of mortality and is the approach used by the DHS to compute stillbirth rates.
b) **True cohort life table** is an approach where the number of deaths of a specific cohort is divided by the number of births in that same cohort. Probabilities of death or stillbirths are computed using this approach. The main problem with this approach is that the true cohort rates do not relate to a particular period of occurrence of the deaths. For example, the effect of structural events like drought and famine, which can affect several cohorts over time.
c) **A synthetic cohort life table** is an approach where mortality probabilities for small age segments are combined into common age segments (such as 0, 1-2, 3-5, 6-11 months). This approach specifies the time period of inquiry (i.e., periods of three or five years) and allows

usage of all recent survey data, and is the approach used by the DHS to compute neonatal and child mortality rates.

1.5. Review of potential biases and measurement errors in surveys

Direct estimation methods face biases and errors that may compromise the mortality estimates. The cardinal biases and errors encountered in population-based surveys include structural bias, sample bias, sampling errors, and reporting or non-sampling errors.[25] This PhD assesses several aspects including structural, sampling biases and reporting errors.

Structural biases: Structural biases result from the design of the questionnaire and survey, and truncation of data: Bias can arise when the questionnaire items are unclear to the respondent or interviewer and lack appropriate instructions.[26-29]

Sample bias and sampling errors: Occur when the sample design systematically omits specific groups within the population from the sample, for example, restriction to residents of a particular area or people with specific characteristics. An important potential source of bias is introduced as most surveys exclude women who died, yet the probability of stillbirth or child death increases when the mother has died. Sample bias is minimised when using well-designed samples, correctly implemented surveys with minimal non-response and if few women aged 15-49 years die.

The population parameters may be either lower or higher than the survey estimates due to sampling variation. These errors are characteristic of surveys and indicate the sampling variability associated with a particular estimate. However, random sampling does not result in bias in the sense that the expected value is equal to the true value, but estimate may be above or below true value.

Reporting or Non-sampling errors: These errors arise in surveys when the quality of data collection is compromised during data collection, such as misreporting events or omission of events because of deliberate non-reporting, failure to remember, or if the interviewer does not record the event, either deliberately or otherwise.[30] These can result in displacement or complete omission of events. Examples of displacement of events may include; misreporting of age at death or birth; and heaping or digit preference in reporting of the age of death, birth dates,

death dates to particular hours, days and months; while omission occurs when data about events is missing entirely.[17,31]

Non-sampling or reporting errors affect the computed mortality rates, leading to underestimation or overestimation of the rates and distorted mortality trends. To evaluate the validity and reliability of maternity history data, an assessment for potential omission[1] or displacement[2] of births or deaths using three methods is recommended in the DHS methodological report 11 notably: a) examination of the process data such as household listing, household registers, number of interviews expected vs conducted or conducting re-interviews to check for consistency; b) examination of consistency between successive surveys or rates computed by other sources; and c) checking for internal consistency of recorded events after data has been collected but before computation of the mortality estimates.[31]

1.6. Conceptual framework for sources of biases and errors in surveys

The reliability and validity of estimates are dependent on the survey design and implementation processes, interview and respondent attributes, human errors when programming the data collection/entry software and questionnaire design, data management and analysis processes.[15] Low data quality may result in unreliable and invalid survey estimates. Figure 1.2 provides a visual overview of sources of bias and errors that are associated with computed estimates from standardised surveys.

Reporting or non-sampling errors bias the estimates, and their effects (underestimation or overestimation of the mortality rates and the distortion of time trends for mortality estimates) are caused by displacement or omission or lack of representativeness. These errors are common for surveys that include both demographic and health information collected for specified periods for example, 0-3 years or 0-5 years.[14,17,31,32]

Misclassification between stillbirths and early neonatal deaths may not bias the computed perinatal mortality rate but can have a large impact on the SBR and NMR because it moves these

[1] Removal or non-reporting of events that may occur during an interview in a survey. This can be deliberate removal, intentional or non-intentional.[32]

[2] Misreporting about the occurrence of events in a survey.[32]

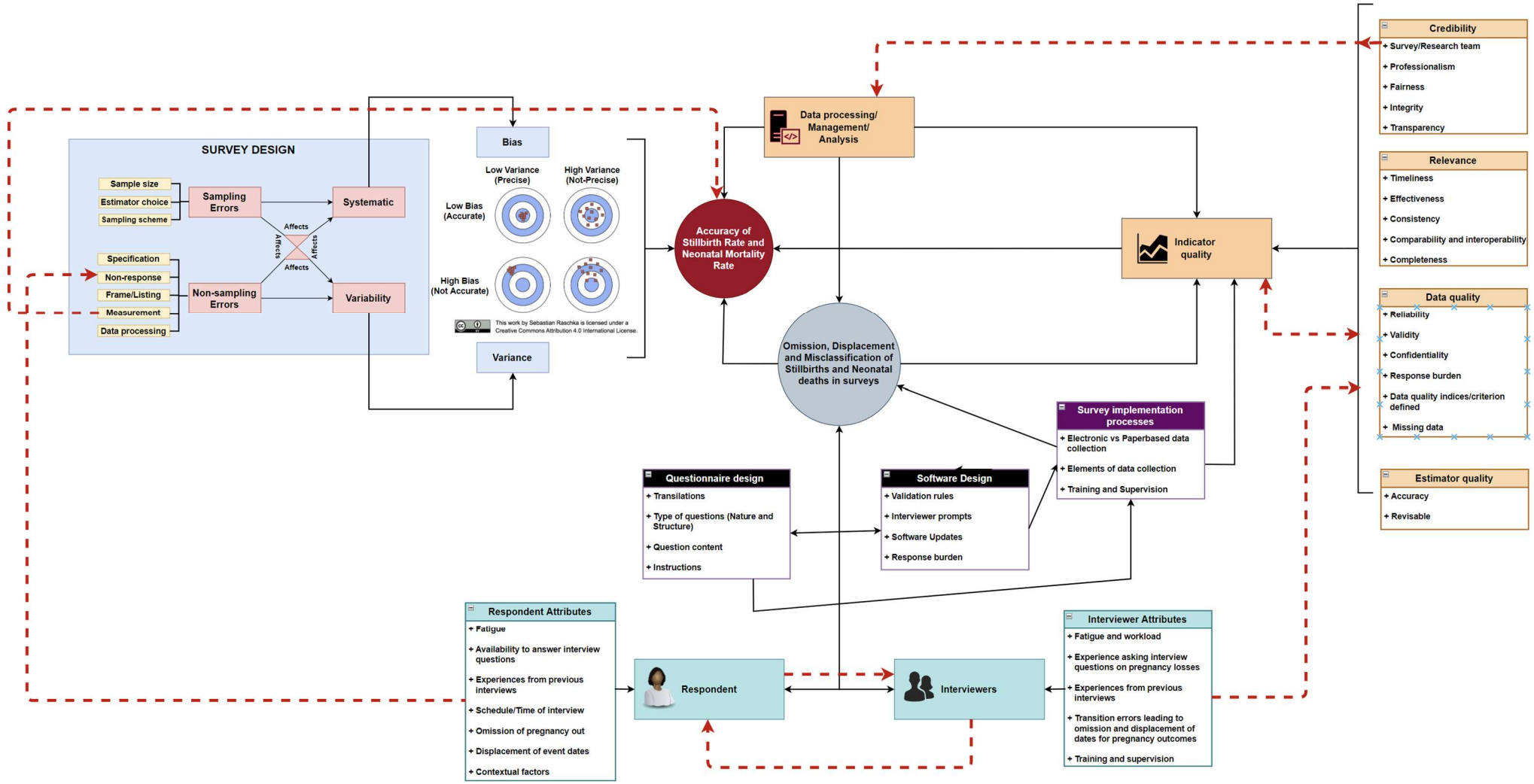
events at the inclusion and exclusion boundaries for stillbirths (pregnancy loss at ≥7 months of gestation) and newborn deaths (death within 0-27 days after birth) (See Figure 1).

In surveys, omission is more common for miscarriages, stillbirths and early neonatal deaths compared to other events, especially where the respondent has had many pregnancies or births, or as the time between the event and the survey increases.[14,17,31,32]

Software and questionnaire design may influence the survey estimates indirectly through survey implementation processes. Omission, displacement, and misclassification can occur in surveys implemented with either a paper-based or electronic system and directly influences the indicator and data quality.

Indicator quality defined as "any measure of the process, performance, or outcome of health care delivery." has several factors, including credibility (*believability of the source*), relevance (*usefulness or appropriateness of the indicator*), data quality (*data that fits its purpose*), and estimator quality (*an unbiased estimate of a population parameter*), influencing the survey computed estimates.[33] Finally, data processing, management and analysis may also directly or indirectly influence the quality and accuracy of the survey computed estimates (See Figure 2).

Figure 1.2: Conceptual framework showing potential measurement errors that affect the SBR and NMR estimates computed from retrospective survey data



Source: Author

1.7. Rationale

The DHS programme surveys are the commonest source of data on stillbirths and neonatal deaths in LMICs. Therefore, this PhD's focus is on data capture and measurement of stillbirths and neonatal deaths in the DHS.

The DHS programme has developed standardised methodologies and approaches over the last four decades; however, challenges to collecting data on stillbirths and neonatal deaths remain due to the various methods and approaches used, including maternity histories and reproductive calendars.[11] Therefore, we need to review the measurement procedures, test and standardise questionnaires in surveys. Furthermore, no previous studies have systematically reviewed the many changes in the DHS programme concerning the measurement of stillbirths and neonatal deaths. The DHS women's questionnaires have used either a full birth history (FBH) or full pregnancy history (FPH) module as one of the primary sources of data on stillbirths and neonatal deaths in LMIC,[11,12] although the FBH module has been the standard core DHS questionnaire. The choice of implementing either FBH or FPH is left to the countries/researchers to make.[12,13] Several country-level DHS implementers have preferred the FBH module because it is assumed to be shorter than the FPH, although no studies have been conducted to investigate the time taken to complete either module.

Four previous studies report comparisons of estimates produced from both FPH and FBH, and data suggests that the FPH produces higher estimates of NMR and Perinatal Morality Rate (PNMR) to the FBH (Table 1).[7,17,32,34]. Only one of these studies conducted a direct comparison and did not include stillbirths. This was conducted in 1994 in Bangladesh, and the study found a 3% (95%CI (1% – 5%)) higher early neonatal mortality rate (ENMR) in FPH compared to FBH. This study had a "small" sample size (n=3039), did not randomly allocate questionnaires with the different modules to respondents, questionnaires with different modules were administered to different women, both individual and aggregate level analysis was conducted, and no comparison for stillbirths was done.[17] A 35% (RR=0.65 (0.46 – 0.93)) higher SBR and no difference in NMR (RR=1.08 (0.82 – 1.42)) was observed between the Ghana DHS and the Ghana maternity health survey (GMHS). The Ghana DHS (GDHS) had the standard FBH, while the Ghana maternity survey had the FPH. The GMHS (n=6960) and GDHS (n=2949) were conducted in 2007 and 2008, respectively with nationally representative samples.[32,34] Another study by Stanton compared the same maternity history module, a full pregnancy

(PH) module in the Philippine's 1993 DHS and the Safe Motherhood survey (SMS) with a shorter pregnancy history module. The SMS was a follow-up study for the women identified in the DHS who had at least one pregnancy outcome. Stanton found 34% and 26% higher stillbirths and early neonatal deaths respectively in the SMS than in the 1993 DHS. The SMS study had a considerably shorter questionnaire compared to the full 1993 DHS.[7] The differences observed in the Ghana maternity surveys and in Stanton's PhD could have arisen due to fatigue from the longer administered standard DHS questionnaires compared to the shorter focused questionnaires used in the SMS and Ghana maternity survey. However, this hypothesis has not been tested.

The impact of omission and displacement on underestimation or overestimation and distortion of time trends for estimates of births, neonatal deaths, and age is reported in the guide to DHS statistics,[14] DHS methodological reports 5 and 11 [31,35] and several studies.[17,30,31,35] The DHS methodological report 11 focused on measurement errors in FBH conducted in 192 surveys between the years 1990-2013, while a study in Bangladesh looked at both the FBH and FPH; however, it used data from more than thirty years ago.[17,35] Another study by Pullum conducted in 1997 using Pakistan DHS 1991 (n=8418 births) assessed recall errors using both individualised and aggregated models focusing on fertility. This study concluded that these models could assess potential omission and displacement of events in any retrospective surveys.[30] None of these sources included the impact on stillbirths.

Due to several factors including, the inconclusiveness of findings about comparison of estimates from FBH and FPH, no previous assessments for potential omission and displacement around the perinatal period and no previous assessment of time taken to complete either a FBH or FPH survey module. There is a need to investigate the unanswered questions left by previous studies. This PhD study will aim to bridge the knowledge gap and will adopt similar approaches used by Espeut[17], DHS methodological report 11[31] and the Pullum studies[30,35] but will furthermore include stillbirths; assess for the combined and separate measurement errors in FPH and FBH and the determinants of response times to questions in the FBH and FPH.

Table 1.1: Summary of available studies and data comparing FBH+ and FPH [36]

Source	*Reference*	*Study site*	*Dates of Study*	*Comparator 1*	*Comparator 2*	*Number of observations*	*Rates in comparator 1*	*Rates in comparator 2*	*Differences*
DHS program	Ghana Statistical Agency [13]	Ghana (national)	2008 DHS (birth history) and 2007 Maternal Health Survey (FPH)	Standard DHS with FBH+	Maternal Health Survey (MHS) with FPH	DHS – 2,949 MHS – 6,960	Stillbirths 40/2949 13.6 per 1000	Stillbirths 146/6960 21.0 per 1000	Rates were 35% lower in birth history compared to pregnancy history RR=0.65 (0.46 – 0.93) There was no difference in Early Neonatal Death (ENND) between the two surveys RR=1.08 (0.82 – 1.42)
DHS program	Ghana Statistical Agency [14]			107 surveys using birth history	13 surveys using pregnancy history		Mean SBR: ENMR ratio = 0.75 Mean SBR=12.7	Mean SBR: ENMR ratio = 0.9 Mean SBR=13.3	
Older studies which were not direct comparison of the birth/pregnancy history approach, or did not assess SBR difference									
Stanton C.	PhD JHU [15]	Philippines, National	1993 DHS and SMS surveys	Standard DHS with FPH replacing birth history	Safe Motherhood Survey (SMS) – pregnancy history plus sex of stillbirth, additional probe if pregnancy interval ≥4years	In SMS 6329 births in three years prior to survey			# Stillbirths 34% higher in SMS # Early losses 26% higher in SMS [note comparison not direct as different household, same pregnancy history module but DHS many more questions – findings may be more about survey fatigue than the module]
Espeut D., Becker S.	Journal of Health, Population and Nutrition[16]	Bangladesh, Matlab HDSS	1994 Matlab Demographic survey	50% questionnaire FBH+ forward (as per Bangladesh DHS)	50% questionnaire FPH forward	Matlab DHS – 3225 households sampled, 3009 household interviews, 3480 women reproductive age, 3039 women interviews	ENND – 46/58 (79%) (HDSS baseline) captured in survey Late Neonatal Death (LNND) – 22/26 captured	ENND – 63/77 (82%) (HDSS baseline) captured in survey	Did not report on stillbirths For ENNDs – 2 – 3% higher capture in pregnancy history [note also very small sample size]

Source: EN-INDEPTH Protocol paper

Chapter 2: Thesis aim, objectives, design and structure, candidate's contribution, and list of associated publications

2.1. Introduction

This chapter presents the thesis aim, objectives, thesis design and structure, and the candidate's roles and contributions to the thesis.

2.2. Thesis aim

The overall aim of this thesis is to improve our understanding of data capture for stillbirth and neonatal deaths in standardised DHS surveys through reviewing previous practices and comparing two alternative approaches to inform future population survey methodology.
The overall aim will be addressed through the following objectives.

2.3. Thesis objectives

Objective 1: To review the Demographic and Health Survey programme's evolution with respect to measurement of stillbirths and neonatal deaths and the performance of the different variants used over the course of DHS against potential markers of data quality.
Objective 2: To undertake a randomised comparison of the reproductive module used in the latest version of DHS-7 "full birth history" (FBH+) versus a "full pregnancy history" (FPH) module for estimation of stillbirth rate and neonatal mortality rate
Objective 3: To examine the variation in response time to completion for questions in the FBH+ and FPH survey modules
Objective 4: To use the data collected as part of the EN-INDEPTH study across 5 well-phenotyped HDSS sites in Africa and Asia to identify and make recommendations concerning ways in which measurement of pregnancy outcomes in population-based surveys and HDSS data systems can be improved.

2.4. Thesis structure

This PhD thesis follows a 'research paper style' and presents five research papers associated with the PhD project. It is composed of eight chapters. All research papers were written as standalone manuscripts and therefore, there are some inescapable repetitions such as study setting, population, definitions, and randomisation among others in the research papers. The

presentation of the papers is not based on the timing of the publication but on how cohesive they are in telling a complete storyline. Details about the thesis chapters, related objectives, research questions and themes and methods are provided in Table 1.

- Chapter 1: Provides a background on stillbirths and neonatal deaths, including their definitions, global burden and targets, data gaps, measurement procedures, potential bias and measurement errors and study rationale.
- Chapter 2: Summarises the thesis aim, objectives, thesis structure, and the candidate's roles and contributions.
- Chapter 3: Addresses thesis objective 1 "To review the Demographic and Health Survey programme's evolution with respect to measurement of stillbirths and neonatal deaths and the performance of the different variants used over the course of DHS against potential markers of data quality". Literature review and systematic review of DHS core module questionnaire with respect to the evolution of stillbirth and neonatal death measurement within the programme and stillbirth and neonatal death data quality assessment within the DHS was conducted
- Chapter 4: Describes the EN-INDEPTH study protocol including study design, population, setting, sampling, randomisation, data collection and ethical clearance. It also briefly describes the methods for objectives 2, 3 and 4
- Chapter 5: Addresses thesis objective 2 "To undertake a randomised comparison of the reproductive module used in the latest version of DHS-7 "full birth history" versus a "full pregnancy history" module for estimation of stillbirth rate and neonatal mortality rate" and objective 3 "To examine the variation in response time to completion for questions in the FBH+ and FPH survey modules"
- Chapter 6: Addresses objective 3 through further assessments of factors affecting the time taken for completion of FBH+ and FPH survey modules. This chapter also assesses the data quality of the two modules by examining interviewer corrections during the interview and missingness patterns. It further highlights future considerations for improving stillbirth and neonatal data quality through analysing survey Paradata.
- Chapter 7: Addresses thesis objective 4 "To use the data collected as part of the EN-INDEPTH study across 5 well-phenotyped HDSS sites in Africa and Asia to identify and make recommendations concerning ways in which measurement of pregnancy outcomes in population-based surveys and HDSS data systems can be improved" and

its methods. It describes the five-HDSS systems for pregnancy and outcome surveillance and compares population-level estimates computed from the HDSS data and survey data by study grouping in four sites.

- Chapter 8: Summarises and discusses the key PhD findings in relation to the already existing literature. It presents the implications of the findings for measurement of stillbirths now, generalisability strengths, limitations, and future research. Chapter 8 also summarises the conclusions and recommendations from the lessons learnt through implementing the FBH+ and FPH for collecting data on stillbirths and neonatal deaths.
- Appendices contain additional relevant documents, including ethical approvals, data management plans, study manuals, supplementary findings, and further manuscripts.

2.5. Candidate's contribution

2.5.1. EN-INDEPTH study

I have been part of the EN-INDEPTH study team since its inception. As the technical coordinator I have made substantial contributions to the overall study through design, implementation, analysis and writing up of results. Some of my contributions include.

1) **Site selection**
   - In October 2015, I led the development of the request for applications (RFA) that was advertised within the INDEPTH Network asking for sites to compete and submit their applications to participate in the EN-INDEPTH study.
   - Working with the INDEPTH Network administrative secretariat, I received the applications on behalf of the EN-INDEPTH study team and coordinated the official feedback process about the statuses of the sites' applications. Fourteen applications were received.
   - Working with the principal investigators at London School of Hygiene & Tropical Medicine (LSHTM) and Makerere University School of Public Health (MakSPH), I coordinated and led the development criteria for scoring the applications received from the INDEPTH sites, I also participated in the scoring process as one of the reviewers for the fourteen applicants.
   - I coordinated the drafting the letters to the sites with feedback about the status of their applications (success or fail) and sent the feedback to the sites.

2) **Study design and ethics**

- Together with the study team, I provided input to the study protocol development for institutional ethical review submission, and protocol paper.
- I supported the questionnaire review and adaptation processes.
- I led the development of the data sharing agreement and coordinated its signing by all members of the consortium.
- I led the development of some of the EN-INDEPTH manuals and standard operating procedures; for example (Tablet management, Survey Solutions implementation manual specific to the study).
- I coordinated the HDSS sites' ethical submissions and feedback processes and supported the LSHTM ethical approval applications.

3) **Workshops**

- As part of the EN-INDEPTH study team, I coordinated/supported the organisation of all three EN-INDEPTH workshops:

Table 2.1: Overview of thesis chapters, research questions and methods

Chapters	PhD Objectives	Research themes and questions	Methods
Chapter 1	Background	○ What are stillbirths and neonatal deaths and why are they important? Why do we need data on these adverse pregnancy outcomes? What are the global goals and targets? ○ Introduction to measurement procedures, protocols, and platforms. Data for global estimates ○ Definitions and indicators for measuring stillbirths and neonatal deaths ○ Potential measurement challenges including biases and measurement errors in measurement. ○ Data sources and platforms for measuring and collecting data on stillbirths and neonatal deaths estimates. ○ The rationale for focusing on stillbirths and neonatal deaths measurement	Targeted literature review
Chapter 2	Thesis aim, objectives, structure, candidate's contribution, and list of associated publications	○ Summary of the PhD thesis aim, objectives	
Chapter 3	**Objective 1:** To review the Demographic and Health Survey programme's evolution with respect to measurement of stillbirths and neonatal deaths and the performance of the different variants used over the course of DHS against potential markers of data quality.	○ What approaches have been used by household surveys to capture birth outcomes (stillbirths and neonatal deaths)? ○ How has the DHS program implemented these approaches; how have these evolved? ○ How has the evolution of approaches used in the DHS program impacted stillbirth estimates over time? ○ Overview of the measurement of stillbirths and neonatal deaths and how this evolved	Targeted literature review Systematic reviews of DHS core questionnaires Descriptive analyses Spatial analyses

Chapters	PhD Objectives	Research themes and questions	Methods
		○ Which countries have implemented either a FBH+ and FPH survey module in their national DHS survey? ○ Review DHS data on stillbirths and neonatal deaths and data quality assessment against indicators of data quality for stillbirths and neonatal deaths	
Chapter 4	Methods (published protocol paper)	○ The rationale for the EN-INDEPTH study and need for further research around adverse pregnancy outcomes. ○ Aim and objectives of the EN-INDEPTH study ○ Methods including study design, settings, sampling and sample size, data collection approaches, randomisation approaches and analytical plan	Targeted literature review
Chapter 5	**Objective 2:** To undertake a randomised comparison of the reproductive module used in the latest version of DHS-7 "full birth history" versus a "full pregnancy history" module for estimation of stillbirth rate and neonatal mortality rate	○ Does the FPH module result into a different SBR estimate compared to FBH+ module in a 5-country HDSS sites survey? ○ Does the FPH module result into a different NMR estimate compared to a FBH+ module in a 5-country HDSS sites survey? ○ Does the FPH module result into a different PNMR estimate compared to FBH+ module in a 5-country HDSS sites survey? ○ Computing and comparing population-level stillbirth and neonatal deaths estimates ○ Conducting meta-analysis with random effects to compute Crude risk ratios (CRR) (adjusting for clustering of stillbirth or neonatal deaths within women) with their 95% CI	Targeted literature review Descriptive analysis Generalised Estimation Equations (GEE) regression Meta-analysis

Chapters	PhD Objectives	Research themes and questions	Methods
	Objective 3: To examine the variation in response time to completion for questions in the FBH+ and FPH survey modules.	○ How long does it take to collect data using the pregnancy history questionnaire compared to a birth history questionnaire? ○ Does the time taken to completion of birth history and pregnancy history questionnaire vary by women and interviewer characteristics, and HDSS site? ○ What are the determinants of the response time to questions in a pregnancy history and birth history randomised survey? ○ Determining the time taken to administer a questionnaire module with either a FPH or FBH+ maternity history ○ Discussing the implications for DHS's switch from FBH+ to FPH for collecting data on pregnancy losses	Descriptive analysis Linear regression with random effects
Chapter 6	**Objective 4:** To examine the variation in response time to completion for questions in the FBH+ and FPH survey and how question characteristics (type, nature, and structure) affect the response time and correction rate, and to classify correction types.	○ How long does it take to collect data using the pregnancy history questionnaire compared to a birth history questionnaire? ○ Does the time taken to completion of birth history and pregnancy history questionnaire vary by question characteristics (types, nature, and structure)? ○ How are answer corrections rates affected by question types and characteristics? ○ What are the determinants of the response time to questions in a pregnancy history and birth history randomised survey? ○ What is the classification of answer correction types and how do these differ between FBH+ and FPH?	Descriptive analysis Independent sample *t* test Generalised Linear model (GLM) regression Sequence analysis

Chapters	PhD Objectives	Research themes and questions	Methods
		○ How can survey process monitoring improve data capture and quality of stillbirths and neonatal deaths?	
Chapter 7	**Objective 5:** To use the data collected as part of the EN-INDEPTH study across 5 well-phenotyped HDSS sites in Africa and Asia to identify and make recommendations concerning ways in which measurement of pregnancy outcomes in population-based surveys and HDSS data systems can be improved.	○ What are the characteristics of the data collecting systems of the 5 HDSS sites participating in the EN-INDEPTH survey? ○ How do rates of key pregnancy outcome events between the two arms of EN-INDEPTH survey (FBH+/FPH) and routine HDSS data for women participating in survey and residing continuously since 1st Jan 2012 compare? ○ Conduct a formative review of the pregnancy and outcome surveillance system in the five HDSS sites ○ Conduct a population-level comparison of estimates computed from HDSS and survey data by study grouping in four sites	Formative review Targeted literature review Country/site consultations Descriptive analysis Meta-analysis
Chapter 8	**Discussion** and study conclusions and recommendations from lessons learnt throughout this study	○ What are the implications of this research? ○ What are the remaining unanswered questions and knowledge gaps? ○ Summarising the lessons and practical implications for using FPH or FBH+ maternity history module, HDSS data, electronic data collection system for collecting data on stillbirths and neonatal deaths. Strengths and limitations of this research	

- Protocol development workshop (in Kampala-Uganda June 2016)
- Implementation workshop (in Matlab-Bangladesh April 2017)
- Scientific writing workshop (in Entebbe-Uganda February 2019).

- I coordinated the monthly study meetings throughout the EN-INDEPTH study's lifetime and during the manuscript writing phase.

4) **Study setup**

- I provided support in the programming and testing of the questionnaire in Survey Solutions and on Android tablets respectively.
- I coordinated the tablet set-up process in all sites during the study set-up phase
- I led the training of the HDSS study teams in three of the five sites (Iganga-Mayuge, Bandim and Dabat)
- I coordinated the field pilot testing of the study tools and Survey Solutions App in the three sites where I led the training of the HDSS study teams. I also provided support to the other two sites where I did not lead training. After the pilot-study, I led the review process of the pilot data.

5) **Study monitoring**

- I led the internal reporting process about the study progress to the principal investigators and research team.
- I contributed to and supported the development of the data monitoring syntaxes and do-files. After the syntax was fully developed and functional, I led the study's data monitoring process and meetings which were held bi-weekly throughout the entire data collection phase of the EN-INDEPTH study (July 2017 to August 2018).
- I led the data review and data collection supervisory processes in one site - IgangaMayuge

6) **Data management and analysis**

- With support from the LSHTM advisors, I led the development of the syntax files/do files that were used for data anonymisation, cleaning and standard analyses across the sites.
- With support from the LSHTM investigators, I led the development of the main results data analysis plans and analysis.
- I led the development of standard datasets for third-party sharing and all data dictionaries. I also provided statistical editorial support to all country teams

- I provided data management, statistics, and data analysis support to the HDSS site teams in relation to the EN-INDEPTH study
- I developed the data dictionaries and meta-data documentation from the EN-INDEPTH study's quantitative datasets
- I led several analyses including meta-analysis for the main results paper and the survey and HDSS paper, systematic reviews synthesis of the DHS questionnaires, descriptive analysis and generation of study flow diagrams and visualisations, spatial analysis, Paradata management and analysis.

7) **Manuscript writing and dissemination**

- I supported the preparation process of the preliminary results dissemination to MONITOR[37] to DHS[38] as part of the consultation on DHS-8.[39]
- I led the writing of the EN-INDEPTH study's main results paper that was published in the Lancet Global Health and supported the writing of several manuscripts in the EN-INDEPTH supplement published in BMC Population Metrics
- I participated in the EN-INDEPTH supplement launch preparatory meetings, and dissemination meeting
- I coordinate the data dissemination and sharing process with third-party users.

2.5.2. PhD work

On 16th January 2017, I was registered and enrolled on the PhD program at LSHTM

This thesis contains some work undertaken collaboratively with a wider research team; therefore, in addition to the cover sheets included in the main thesis, Supplementary Table 1 shows details of chapters, associated publications, activity and roles played by both candidate and others. Summaries of the role of the candidate in the work are presented below.

- Chapters 1 and 2: I conceived and wrote these chapters with input from Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa and Vladimir Gordeev.
- Chapter 3: I conceptualised the paper jointly with Drs. Hannah Blencowe, Angela Baschieri, Vladimir Gordeev, and Professors. Peter Waiswa, Simon Cousens and Joy Lawn. I developed the detailed research questions and analysis plan with feedback from co-authors. I conducted the literature review and undertook a systematic review of the Demographic and Health Survey's programme questionnaire modules. I led the development of the analysis plan and conducted the data analysis of all survey data

from 1984 to 2019 on the Demographic and Health Surveys' programme focusing on stillbirth and neonatal deaths. I developed all the geospatial maps and I wrote the first draft of the manuscript and prepared the subsequent revisions with feedback from co-authors. All authors reviewed and gave approvals of the final manuscript.

- Chapter 4: The study was conceptualised by Professor Joy Lawn, and the protocol developed and refined by the team including Joy Lawn, Peter Waiswa, Hannah Blencowe, Angela Baschieri, Vladimir Gordeev and myself. I led the development of the questionnaires and supported Vladimir Gordeev with the development and programming of a section of the EN-INDEPTH survey questionnaire in the World Bank's Survey Solutions data collection platform and application. I led the development of various implementation manuals and standard operating procedure documents including study manuals, ethical clearance application and data sharing agreement. I led the training of interviewers in three out of five sites and coordinated the rollout and monitoring of the data collection processes in all five sites.

I supported Angela and Vladimir in conducting the background literature reviews and in drafting and revising the manuscript.

- Chapter 5: I developed the detailed research questions jointly with Drs. Hannah Blencowe, Angela Baschieri, Vladimir Gordeev, and Professors. Peter Waiswa, Simon Cousens and Joy Lawn. I conducted the literature reviews and with guidance from Prof. Simon Cousens and Dr. Hannah Blencowe, I developed the detailed analysis plan and led all the data management, development of statistical code and undertook all analyses. I drafted the initial version of the paper and coordinated the revisions and submission of this manuscript until it was published in the Lancet Global Health. All co-authors reviewed and gave approvals of the final manuscript
- Chapter 6: I led the overall data management of the EN-INDEPTH paradata which was analysed for this paper. Jointly with Drs. Vladimir Gordeev, and Hannah Blencowe, Professor Joy Lawn and I contributed the conceptualisation of this paper and to the development and refining of detailed research questions. I met regularly with Vladimir Gordeev and Hannah Blencowe to review the preliminary results from the data analysis. I supported Dr. Vladimir Gordeev in drafting and revising of the manuscript. I developed some of the graphics in the manuscript and supported the peer-review

process providing responses to reviewers' comments on the manuscript after submission.

- Chapter 7: I led the conceptualisation and writing of the manuscript. I developed and revised the detailed research questions with feedback from co-authors. I conducted the literature reviews, and data management. With guidance from Prof. Simon Cousens and Dr. Hannah Blencowe, I developed the detailed analysis plan, and conducted the data analysis. I drafted the initial draft of the paper. I led the writing and preparation of manuscript for submission and coordinated the reviews process by all authors.
- Chapter 8: I led the conceptualisation and writing of this chapter. I received guidance and feedback from Simon Cousens, Hannah Blencowe and Peter Waiswa.

2.6. Associated conferences and workshop presentations

Over the course of this PhD, I have participated in several conferences and workshops at which I have presented some of my preliminary findings.

- International Union of African Population Studies November 2017
- Global Women's Conference September 2019
- Union of African Population conference November 2019
- Grand Challenges 2019 Annual Meeting (poster)
- EveryNewborn-INDEPTH supplement launch December 2020
- EN-INDEPTH Paradata consultative meeting, August 2021
- Sub Saharan Congenital Anormalities Network and The Global Health Network – Birth Defects Surveillance Seminar, August 2021
- Grand Challenges 2021 Annual Meeting – Knowledge Integration session, November 2021

2.7. Chapter summary

In this chapter, I present an overview of the PhD thesis, aim, research questions, themes and specific objectives, thesis design and structure, and the overall candidate's roles and contributions.

Chapter 3: DHS Historical Review of Demographic and Health Surveys

3.1. Introduction

This chapter provides a review of the Demographic and Health Survey programme's evolution for measurement of stillbirths and neonatal deaths and their performance against potential markers of data quality in population-based surveys (Objective 1)

This chapter was published February 2021 in BMC Population Health Metrics.[40] The manuscript was published under a Creative Commons Attribution 4.0 International License (CC BY 4.0) - http://creativecommons.org/licenses/by/4.0/ and the published manuscript is included in full below. The web appendix referenced in the paper is available at https://pophealthmetrics.biomedcentral.com/articles/10.1186/s12963-020-00225-0

3.2. List of Figures

Figure 1 – DHS model questionnaire overview of content
Figure 2 – Overview of reproduction section used in Demographic and Health Surveys (DHS) questionnaire
Figure 3 – Comparing the model Demographic and Health Surveys' full birth history (FBH) questionnaire across phases I–VII
Figure 4 – Demographic and Health Surveys (DHS) programme number of surveys in phases I–VII by country
Figure 5 – Demographic and Health Surveys phases I–VII (1985 to 2018) by type of maternity history
Figure 6 – Location of Demographic and Health Surveys programme countries with FBH and FPH (phases I–VII)
Figure 7 – Distribution of stillbirth rate to neonatal mortality rate ratio by DHS module by phases II–VII
Figure 8a – Comparison of DHS surveys (1990–2018) for SBR:NMR ratios a data quality marker by module
Figure 8b – Comparison of all DHS surveys (1990–2018) for SBR:NMR ratios by module by phase

3.3. List of Tables

Table 1 – DHS phases' summary of neonatal mortality and stillbirth rates by maternity history type

3.4. Citation

Joseph Akuze, Simon Cousens, Joy E. Lawn, Peter Waiswa, Vladimir Sergeevich Gordeev, Fred Arnold, Trevor Croft, Angela Baschieri & Hannah Blencowe. **Four decades of measuring stillbirths and neonatal deaths in Demographic and Health Surveys: historical review**. BMC Population Health Metrics 2020, 19(Suppl 1):8 https://doi.org/10.1186/s12963-020-00225-0

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1512635	**Title**	Mr
First Name(s)	Joseph		
Surname/Family Name	Akuze Waiswa		
Thesis Title	Measurement of stillbirths and neonatal deaths in standardised population-based surveys		
Primary Supervisor	Hannah Blencowe		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	BMC population metrics as: Joseph Akuze, Simon Cousens, Joy E. Lawn, Peter Waiswa, Vladimir Sergeevich Gordeev, Fred Arnold, Trevor Croft, Angela Baschieri, Hannah Blencowe. Four decades of measuring stillbirths and neonatal deaths in Demographic and Health Surveys: historical review. London School of Hygiene and Tropical Medicine. BMC Population Health Metrics https://doi.org/10.1186/s12963-020-00225-0		
When was the work published?	08 February 2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	Not Applicable		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Not Applicable
Please list the paper's authors in the intended authorship order:	Not Applicable
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceptualised the paper jointly with Drs. Hannah Blencowe, Angela Baschieri, Vladimir Gordeev, and Professors. Peter Waiswa, Simon Cousens and Joy Lawn. I developed the detailed research questions and analysis plan with feedback from co-authors. I conducted the literature review and undertook a systematic review of the Demographic and Health Survey's programme questionnaire modules. I led the development of the analysis plan and conducted the data analysis of all survey data from 1984 to 2019 on the Demographic and Health Surveys' programme focusing on stillbirth and neonatal deaths. I developed all the geospatial maps and I wrote the first draft of the manuscript and prepared the subsequent revisions with feedback from co-authors. All authors reviewed and gave approvals of the final manuscript.

SECTION E

Student Signature	Joseph Akuze Waiswa
Date	27th September 2021

Supervisor Signature	Hannah Blencowe
Date	28th September 2021

3.5. Research paper

A: Four decades of measuring stillbirths and neonatal deaths in Demographic and Health Surveys: historical review

Akuze et al. Population Health Metrics 2020, 19(Suppl 1):8
https://doi.org/10.1186/s12963-020-00225-0

Population Health Metrics

From Every Newborn-INDEPTH study: Improving the measurement of pregnancy outcomes in population-based surveys

RESEARCH Open Access

Four decades of measuring stillbirths and neonatal deaths in Demographic and Health Surveys: historical review

Joseph Akuze[1,2*†], Simon Cousens[1], Joy E. Lawn[1], Peter Waiswa[2,3], Vladimir Sergeevich Gordeev[1,4], Fred Arnold[5], Trevor Croft[5], Angela Baschieri[1] and Hannah Blencowe[1†]

Abstract

Background: Worldwide, an estimated 5.1 million stillbirths and neonatal deaths occur annually, 98% in low- and middle-income countries. Limited coverage of civil and vital registration systems necessitates reliance on women's retrospective reporting in household surveys for data on these deaths. The predominant platform, Demographic and Health Surveys (DHS), has evolved over the last 35 years and differs by country, yet no previous study has described these differences and the effects of these changes on stillbirth and neonatal death measurement.

Methods: We undertook a review of DHS model questionnaires, protocols and methodological reports from DHS-I to DHS-VII, focusing on the collection of information on stillbirth and neonatal deaths describing differences in approaches, questionnaires and geographic reach up to December 9, 2019. We analysed the resultant data, applied previously used data quality criteria including ratios of stillbirth rate (SBR) to neonatal mortality rate (NMR) and early NMR (ENMR) to NMR, comparing by country, over time and by DHS module.

Results: DHS has conducted >320 surveys in 90 countries since 1984. Two types of maternity history have been used: full birth history (FBH) and full pregnancy history (FPH). A FBH collecting information only on live births has been included in all model questionnaires to date, with data on stillbirths collected through a reproductive calendar (DHS II-VI) or using additional questions on non-live births (DHS-VII). FPH collecting information on all pregnancies including live births, miscarriages, abortions and stillbirths has been used in 17 countries. We found no evidence of variation in stillbirth data quality assessed by SBR:NMR over time for FBH surveys with reproductive calendar, some variation for surveys with FBH in DHS-VII and most variation among the surveys conducted with a FPH. ENMR:NMR ratio increased over time, which may reflect changes in data quality or real epidemiological change.

(Continued on next page)

* Correspondence: jakuze@musph.ac.ug
Joseph Akuze is the first author and Hannah Blencowe is the senior author.
[1]Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene & Tropical Medicine, London, UK
[2]Centre of Excellence for Maternal Newborn and Child Health Research, Dept. of Health Policy, Planning and Management, Makerere University School of Public Health, Kampala, Uganda
Full list of author information is available at the end of the article

© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Conclusion: DHS remains the major data source for pregnancy outcomes worldwide. Although the DHS model questionnaire has evolved over the last three and half decades, more robust evidence is required concerning optimal methods to obtain accurate data on stillbirths and neonatal deaths through household surveys and also to develop and test standardised data quality criteria.

Keywords: Stillbirths, Neonatal deaths, Demographic and Health Surveys, Questionnaires

Key findings

What is new?

- **What was known already**: Demographic and Health Surveys (DHS) have been the main source of information on child mortality in most low- and middle-income countries over the past three and a half decades, and the major data input for two thirds of the world's estimated 5.1 million stillbirths and neonatal deaths.
- **What was done:** Survey tools have evolved over time, but these changes and the potential effects on national stillbirth and neonatal mortality data have not been systematically assessed before. Our study addresses this gap.

What changed in DHS over time?

- From 1984 to date, the DHS programme had seven phases (DHS-I to DHS-VII) collecting data from more than 400 surveys in more than 90 countries. The model questionnaires are revised for each phase with two main approaches for capturing information on births:

  Full birth history (FBH), capturing a woman's lifetime live births and survival status, is used to calculate neonatal and child mortality. Throughout all DHS phases, the model questionnaire included an FBH and most countries have implemented this approach. Minor changes to the FBH have been made during the last three decades, including adding and then refining a question to capture omitted child deaths (DHS-III to DHS-V) and introducing a question of the day of death in DHS-VII.

  Full pregnancy histories (FPH) capture miscarriages, terminations of pregnancy and stillbirths, as well as live births. FPH has been used by DHS in 17 countries (five in Central Asia, two in Southeast Asia, two in Western Asia, two in Africa, two in Eastern Europe and one in Latin America).
- Stillbirths were initially not captured or reported in DHS-I. In DHS-II to DHS VI, reproductive calendars were used to generate stillbirth data. Since DHS-III, stillbirth data have been shown in the standard national DHS tabulation. DHS-VII introduced a reverse truncated history for non-live births in the last 5 years.

What changed in the data over time?

- **DHS data quality assessment criteria**: Neonatal deaths in the DHS programme include sex ratios at birth and of neonatal deaths; heaping of neonatal deaths on day 7; and the proportion of infant deaths that are in the neonatal period. These are all problematic as may be due to true epidemiological change, not just data quality.
- **Data quality for stillbirths**: Data quality for stillbirths is often assessed by SBR:NMR ratio. Our assessment of SBR:NMR suggested that the ratio did not change across DHS-II to DHS-VI, and stillbirth data seem mostly low quality for surveys conducted in DHS-VII. Using FPH, stillbirth data quality are more variable, with some apparently higher quality, which may be related to the use of differing pregnancy history tools and varied implementation between surveys. Contextual societal barriers to reporting pregnancy loss may also play a role.

What next in measurement and research?

- **Measurement improvement now:** From 2020, the DHS programme (DHS-VIII) has changed its model questionnaire to be based on FPH. This change was influenced by the EN-INDEPTH study's randomised comparison of the two approaches, showing higher reporting of stillbirths but not neonatal deaths with FPH, compared with FBH.

Key findings *(Continued)*

However, whilst FPH may improve capture of stillbirths, optimising data quality is also dependent on survey implementation including training and supervision of data collectors, optimal use of electronic platforms, plus addressing contextual barriers to women reporting pregnancy losses.

- **Research needed:** More research is required to develop robust measures of data quality for stillbirths and neonatal deaths.

Background

There were an estimated 5.1 million stillbirths and neonatal deaths worldwide in 2018. 98% of these deaths occurred in low- and middle-income countries (LMICs), with over 75% in sub-Saharan Africa and South Asia [1–5]. These deaths have an impact on women, families, health-workers and wider society [6], yet the majority are preventable through high quality antenatal, childbirth and newborn care [1, 7, 8]. Measuring and monitoring trends in stillbirth and neonatal mortality, therefore, provides an important indicator of maternal health and access to high-quality care [9, 10].

However, whilst high-income countries have national civil and vital registration statistics (CRVS) systems that record these outcomes in a timely and reliable way, CRVS systems in most LMICs are limited in coverage and quality. Even when such CRVS systems capture adult outcomes, there is known to be selective under-reporting of neonatal deaths (especially preterm neonates) and even more so of stillbirths [11]. At the global level, fewer than 5% of all stillbirths and neonatal deaths are captured in CRVS; this is not much higher for under-5 child deaths [2]. LMICs therefore rely on population-level household surveys for data on these indicators [12]. Indeed, as LMICs account for the majority of the world's births and an even higher proportion of child deaths and stillbirths, such survey data are the main input for estimating over two thirds of the burden worldwide.

The largest survey platforms used for estimating child mortality include Demographic Health Surveys (DHS), Reproductive Health Surveys (RHS), Pan Arab Project for Family Health (PAPFAM) surveys and UNICEF's Multiple Indicator Cluster Surveys (MICS). However, PAPFAM and most MICS do not include stillbirths, and RHS are mainly conducted in middle-income countries. Only DHS has systematically captured stillbirths and

neonatal deaths in LMICs throughout most of its history, and hence, this paper focuses on the approaches taken over time by DHS to capture these outcomes.

The DHS programme, primarily funded by the United States Agency for International Development (USAID), is a follow-on to the World Fertility Surveys and Contraceptive Prevalence Surveys that were conducted between 1972 and 1984 to collect data on fertility, mortality and contraceptive use [13–15]. A large focus of the DHS programme remained on analysis of fertility patterns and trends, and child mortality; however, in addition, it also collects information on maternal and child health, nutrition, human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS), malaria, domestic violence and other country-specific indicators of interest [13, 16]. The DHS programme uses a basic approach of collecting comparable data across countries using a model questionnaire which is revised every 5 years through a consultation process (see Fig. 1) [17]. DHS is currently in the 8th phase of its programme [15].

Omission of stillbirths and neonatal deaths in surveys is known to affect the data quality of these indicators; however, assessing data quality for stillbirths and neonatal deaths in surveys is challenging in the absence of high-quality population-based data with which to compare survey estimates. The ratio of stillbirths to overall neonatal mortality rates (SBR:NMR), which detects where stillbirths are under-reported compared with neonatal deaths, has been used as a stillbirth data quality criterion [18, 19]. High-quality historical data from high-income countries report ratios of at least 1 for countries with NMRs of 10–35 per 1000, and similar ratios may be expected from LMICs [20, 21]. DHS data quality assessments for neonatal deaths have used several criteria including sex ratios at birth and of neonatal deaths and heaping of neonatal deaths on day 7 and the proportion of infant deaths that are in the neonatal period [22]. However, the latter has limited utility due to well-documented epidemiological variations with ratios varying by mortality contexts [23, 24]. As early (days 0–6) neonatal deaths are the most frequently omitted deaths, the proportion of neonatal deaths that occurred on days 0–6 (or the early to overall neonatal mortality (ENMR:NMR) ratio) is another potential marker of data quality [25].

Challenges to collecting data on stillbirths and neonatal deaths in surveys have led to a variety of approaches being used over time. No previous studies have systematically described these, how they have evolved over time or the effect of these changes on indicator comparability over time. The objectives of this paper are for DHS phases I–VII are as follows:

1) ***Provide an overview of the measurement of stillbirths and neonatal deaths*** and how this has changed over time.
2) ***Review DHS data on stillbirths and neonatal deaths, and their performance*** against potential markers of data quality.

Methods

We conducted a review between the 17th of November 2017 and 9th of December 2019. We searched POPLINE and PubMed databases and the DHS website using combinations of key words including "*Birth History*", "*Pregnancy History*", "*Questionnaire*", "*World Fertility Surveys*", "*Demographic and Health Surveys*", "*stillbirth*", "perinatal death", "neonatal death", "child death", "perinatal mortality", "neonatal mortality" and "child mortality" for reports and journal articles published since 1982 with a focus on the implementation of DHS for capturing stillbirths and neonatal deaths.

We obtained the DHS's model woman's questionnaires for all DHS phases from the DHS website and reviewed all

The first DHS model questionnaire was developed during two years of consultations with the health and demographic communities, taking into account lessons learned from World Fertility Surveys and Contraceptive Prevalence Surveys.

The model questionnaire includes four components:

1. Household questionnaire: usually responded to by the head of the household or by any adult member who provided information about the characteristics of the households, its members and visitors and to identify household members eligible for the individual interview.

2. Individual woman's questionnaire: responded to by eligible women who provided information on background characteristics; reproduction behaviour and intentions; births and child deaths; contraception; antenatal, delivery and postnatal care; breastfeeding and nutrition; children's health; status of women; HIV and other sexually transmitted infections; husband's background and other topics related to health insurance and tobacco use.

3. Individual man's questionnaire: responded to by eligible men, covers similar topics to the woman's questionnaire plus questions on employment and gender roles, but is shorter as fewer details of maternal and child health and care are included.

4. Biomarker questionnaire: used to collect biomarker data on women, men and children.

Fig. 1 DHS model questionnaire overview of content

the eleven previous model questionnaires (DHS model questionnaires: I-A, I-B, II-A, II-B, III-A, III-B, IV-A, IV-B, V, VI and VII), reports and journal articles for information relevant to survey implementation relating to the measurement of stillbirths and neonatal deaths in the questionnaire's reproduction section.

We extracted summary data from the DHS website's STATcompiler on stillbirths and neonatal deaths from all surveys from 1984 to December 2019. The extracted data were exported first to Microsoft Excel spreadsheets and then imported into Stata 16.0 for further data management and analysis. These data were analysed using descriptive and geospatial techniques using choropleth maps in Stata version 16.0.

We summarised data quality for neonatal mortality and stillbirth rate data over time and by data collection method using a single available measure for each outcome focused on the detection of omission: SBR:NMR ratio for stillbirths and ENMR:NMR ratio for neonatal mortality. Results are presented using descriptive statistics and graphical summaries such as two-way scatter plots. We compared mean SBR, mean NMR, SBR:NMR and ENMR:NMR ratios by DHS phase and type of module implemented using statistical tests for trend and differences in proportions, and for SBR:NMR using box plots.

Results

Overview of the measurement of stillbirths and neonatal deaths in DHS

The predominant method to collect information on neonatal deaths (deaths in the first 28 days of life) in DHS has been through the use of full maternity histories. To collect information on stillbirths, both full maternity histories and reproductive calendars have been used. As neither of these approaches allows capture of pregnancy length in weeks or days, it is not possible to apply standard ICD-11 stillbirth definitions, and a pregnancy loss at seven or more months of gestation is used to approximate late fetal deaths or stillbirths [26].

Maternity history approach

In the DHS, full maternity histories were introduced more than three decades ago to gather retrospective data on women's fertility, births and infant and child deaths [12]. This is in contrast to many other surveys which used summary birth histories, collecting information only on the number of children ever born and the number that survived, and then using indirect methods to estimate overall child mortality rates only [27–29]. From the mortality estimation perspective, these questions were initially predominantly used to estimate infant and overall under-5 mortality rates; however, as information on the precise age at death was included in the full maternity history, it is possible to also estimate neonatal mortality from these questions.

Two types of full maternity histories have been implemented in DHS: full birth history (FBH) and full pregnancy history (FPH) [30]. Both the FBH and FPH are administered in the woman's questionnaire to women aged 15–49 years who consent to participate in the survey. These modules are implemented in the reproduction section of the woman's questionnaire (Additional file 1). The FBH collects information on all pregnancies that resulted in a live birth, survival status of the child and where relevant the age at death. Data on stillbirths are not collected directly in an FBH. The FPH collects information on all pregnancy outcomes including miscarriages, terminations of pregnancy, stillbirths and live births. As with the FBH, the survival status of all live births and where relevant the age at death is included [30].

Both full maternity histories have predominantly been implemented in DHS using a forward approach, starting with the earliest events and detailing each birth or pregnancy in time order [12, 30]. Whilst there is some evidence from the World Fertility Survey that a backward approach starting with the latest event is associated with more detailed probing of later events and slightly fewer missed or time-displaced events, the advantages were not considered sufficient to change the standard DHS approach [12, 31, 32].

To seek to reduce the length of the survey tool, early surveys in Peru and the Dominican Republic experimented with using a back truncated approach, collecting data on events only in the last 5 or 6 years [33, 34]. Overall, they found similar data quality compared with the full maternity history approach. Whilst there was some improvement in the quality of reporting of dates of very recent events, in Peru displacement of events prior to the 5-year period was found, and therefore, this approach has not been adopted in standard DHS.

Reproductive calendar approach

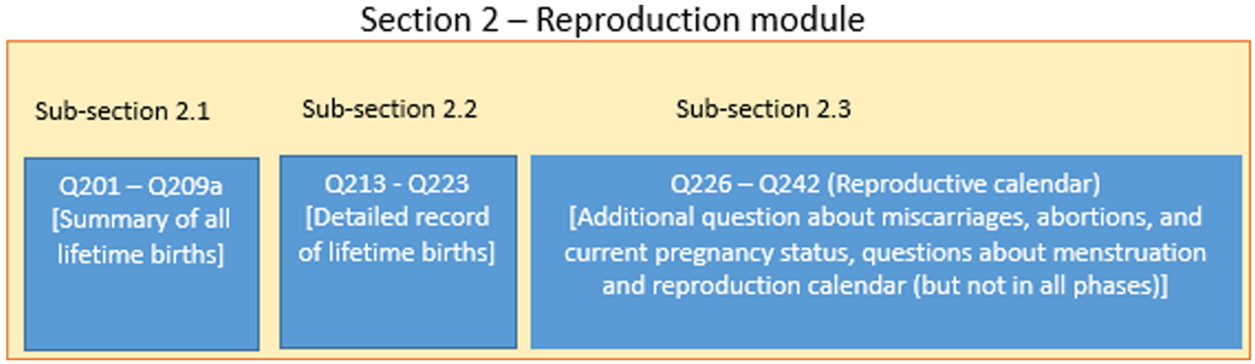
Reproductive calendars collect information on pregnancies, births and contraceptive use by month for the 5-year period preceding the survey. The use of a reproductive calendar first developed in experimental studies in Peru and Dominican Republic in 1986 was found to improve the accuracy of the capture of contraceptive use within surveys [33, 35, 36]. As reproductive calendars record for each month of the preceding 5 years whether a woman was using contraception, was pregnant or gave birth/had a pregnancy end, these data can be used to estimate stillbirth rates [37].

Evolution of DHS's FBH reproduction section in the woman's questionnaire for phases I to VII

The DHS's woman's questionnaire contains a reproduction section which comprises three sub-sections (Fig. 2). This section has evolved through the DHS phases with additions of new questions and removal or modifications of questions. In phases I to IV, two separate model woman's questionnaires were in use— Questionnaire A for high contraceptive prevalence countries and Questionnaire B for countries with low prevalence of contraceptive use. Sub-section 1 did not change across phases. More changes that are substantial were made to sub-sections 2.2 and 2.3, particularly in later phases (Fig. 3 and Additional file 2).

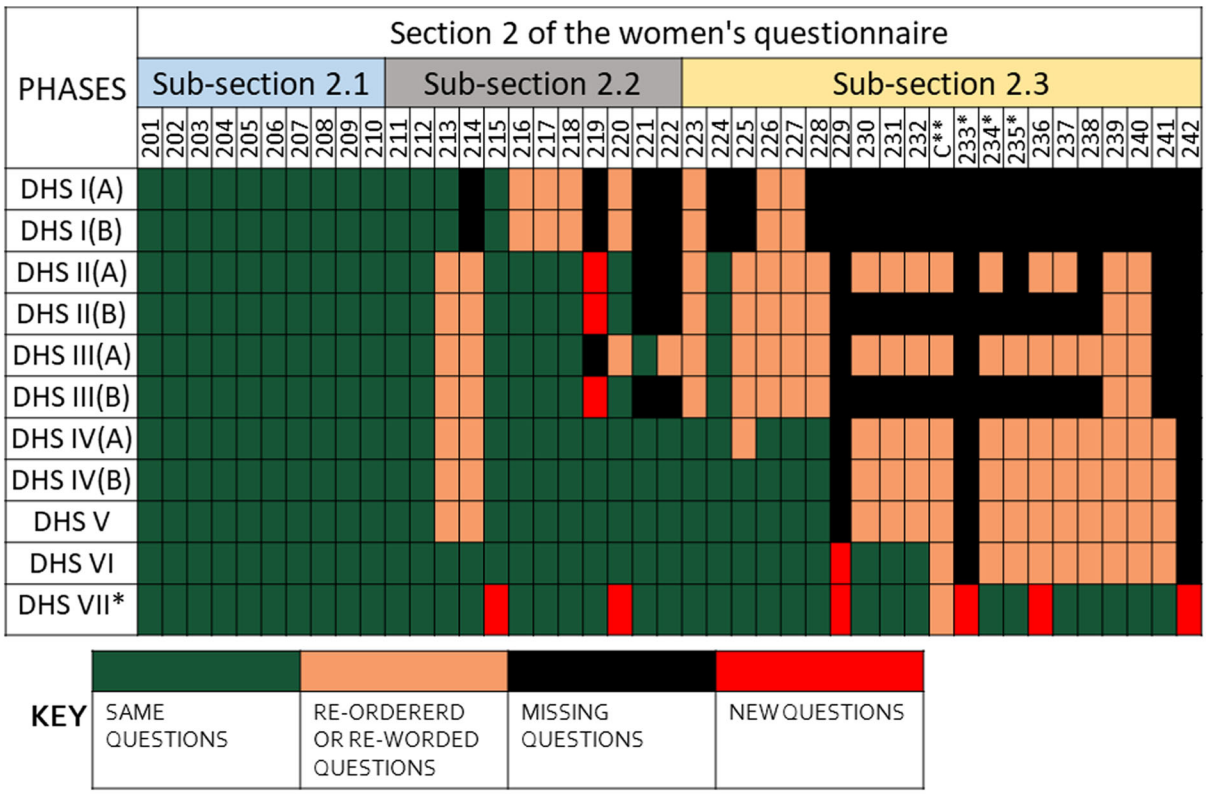
Methods to estimate child mortality, including neonatal mortality, have not changed substantially throughout the DHS phases and do not vary between FBH and FPH. Questions around the precise age at death for any deceased children have been included in all DHS phases. These include information in days if less than 1 month, in months if less than 2 years, otherwise in years. In DHS-III, an interviewer calculation and probe where added to investigate the potential omission of children who had died in all cases with a reported birth interval of more than 4 years. In DHS-IV onwards, this was simplified to probe simply if there were any other live births between each reported birth; from DHS-V onwards, this was expanded to specify "*Were there any other live births between (NAME OF PREVIOUS LIVEBIRTH) and (NAME) including any children who died after birth*?" Reported child mortality has been disaggregated to show neonatal mortality rates in standard DHS reports since DHS-II.

There has been much greater variation in the collection of data to inform estimates of stillbirths in DHS over time. A FBH alone does not capture any information on non-live births. No questions enabling stillbirth rate estimates were included in the model woman's questionnaire prior to 1993 (DHS-I). In DHS-II, a reproductive calendar to capture pregnancy, birth and contraceptive use was included in the model woman's questionnaire A for high contraceptive prevalence countries and in DHS-IV the reproductive calendar was also added to the model woman's questionnaire B for low contraceptive prevalence countries. However, in both cases, the reason stated for including these questions was to avoid misclassifying months in which the woman is pregnant as months of exposure to the risk of pregnancy when calculating rates of contraceptive failure and discontinuation rather than to be able to estimate perinatal mortality. Therefore, despite this information being collected, estimates of stillbirth rates were infrequently included in the final DHS reports for countries using the standard approach of FBH and a reproductive calendar. Hence, whilst stillbirth rates can be calculated retrospectively from these surveys, and are included in the DHS platform's STATcompiler, these have not generally been available to users of these reports such as policy-makers and programme managers in the country. Model questionnaires in DHS-V and DHS-VI also used a similar approach with a reproductive calendar. However, a few countries modified the questionnaire to add additional questions or prompts to seek to improve the capture of stillbirths. Most standard reports from DHS-V onwards contain an estimate of perinatal mortality. In view of ongoing data quality concerns about the information collected regarding non-live births using the reproductive calendar in previous phases, DHS-VII introduced a new table in sections 2–3 to record the details of all non-live births in the last 5 years including the month and year of the event and the length of gestation in months, and a prompt "*since January 2010* (YEAR VARIES DEPENDING ON YEAR OF INTERVIEW*)*



Footnote: summaries are obtained through a series of questions asked about whether the woman has had any births, the numbers of those alive and living or not living with her, the numbers of those born alive and who died later, all stratified by gender. (17, 22). For the FPH two additional questions are asked about any non-live birth pregnancies she may have had. See Additional file 1, for examples of FPH and FBH generic questions.

Fig. 2 Overview of reproduction section used in Demographic and Health Surveys (DHS) questionnaire



A denotes model questionnaire for high contraceptive prevalence settings; B denotes model questionnaire for low contraceptive prevalence settings; C** - Calendar approach; *SBR calculated as a reverse truncated non-live birth history (233 -5) instead of calendar in DHS-VII; DHS – Demographic and Health Survey

Fig. 3 Comparing the model Demographic and Health Surveys' full birth history (FBH) questionnaire across phases I–VII

have you had any other pregnancies that did not result in a live birth?" (Fig. 3 and Additional file 2). This is effectively a back-truncated history for non-live births.

Unlike previous phases that asked for only the month and year of birth, the DHS-VII model questionnaire has a modification in wording to include asking for the day of birth [17]. This change also led to the introduction of the Century Day Code, an algorithm used to standardise dates for DHS survey across all surveys. Previously, the DHS has used the Century Month Code [38].

Data for stillbirths and neonatal deaths in DHS phases I–VII

More than 350 surveys have been conducted in over 90 countries since 1984 (Fig. 4, Additional file 3 A and B) in the DHS programme. A FBH has been included in the standard model questionnaire and has been the predominant maternity history used in all phases of the programme so far (Figs. 5 and 3 and Additional files 4 and 5). Whilst country-specific adaptations can be undertaken, most countries have implemented the model questionnaire, although a report on DHS surveys in 27 countries in phase I found some changes in the question numbering and ordering [39].

An FPH module has been used in only 19% of countries that have conducted DHS. In 1986, Peru conducted the first DHS with an FPH module in an experiment to introduce questions around termination of pregnancies [40]. The other 16 countries started implementing an FPH from phase III (Additional file 4). A total of 35 and 323 FPH and FBH surveys, respectively, have been conducted since 1984 (see Fig. 6 and Additional files 4 and 5). Overall, some Asian and European countries use an FPH whilst only Ghana, South Africa and Peru from Africa and Latin America have (Fig. 6).

Initially, the DHS data collection was conducted using a paper-based system with data entry and processing done using the Census and Survey Processing System (CSPro). In the most recent phase of the DHS, both paper and computer-assisted personal interviewing (CAPI) were used. The DHS programme is planning to use the CSPro computer-assisted interviewing system for all future surveys [38, 41].

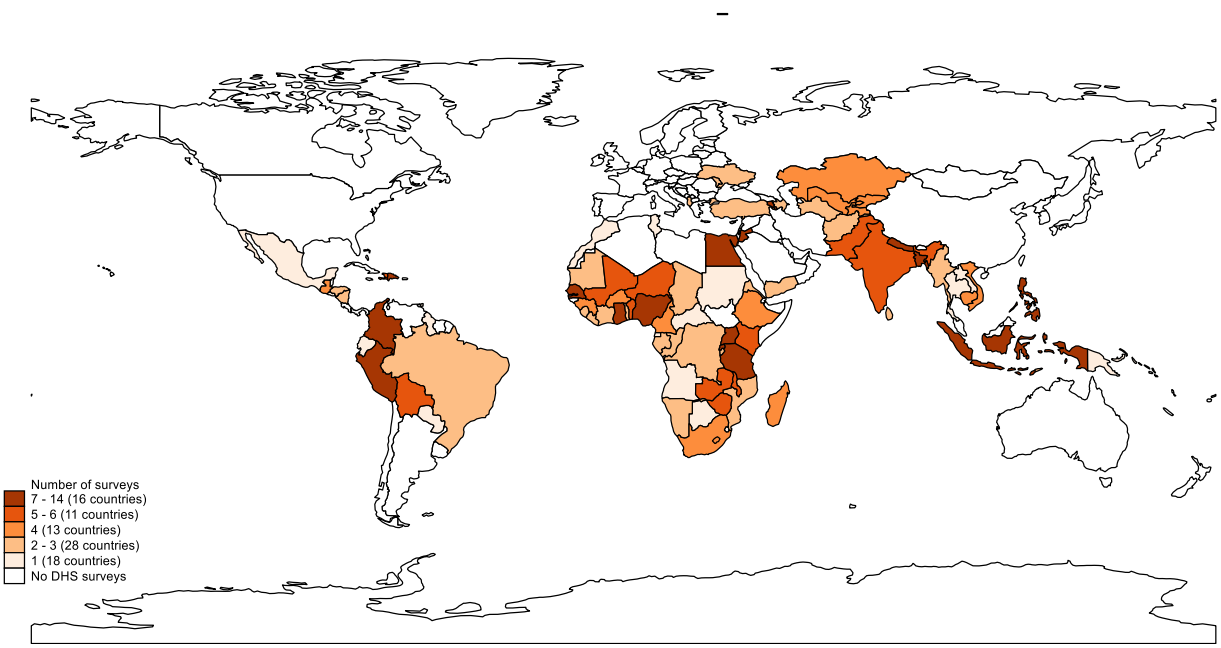


Fig. 4 Demographic and Health Surveys (DHS) programme number of surveys in phases I–VII by country

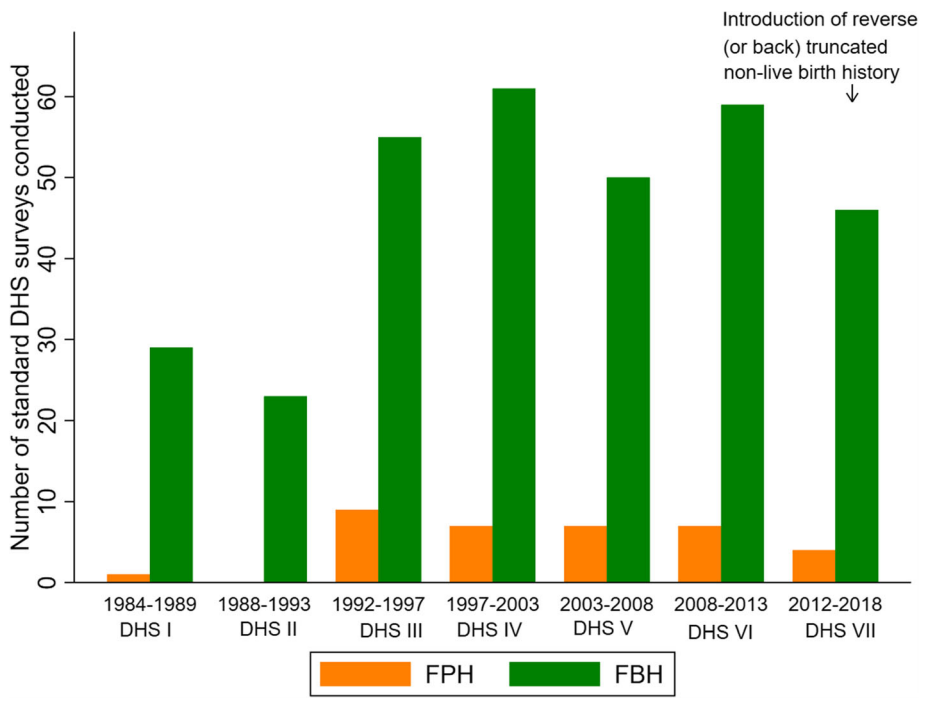


Fig. 5 Demographic and Health Surveys phases I–VII (1985 to 2018) by type of maternity history

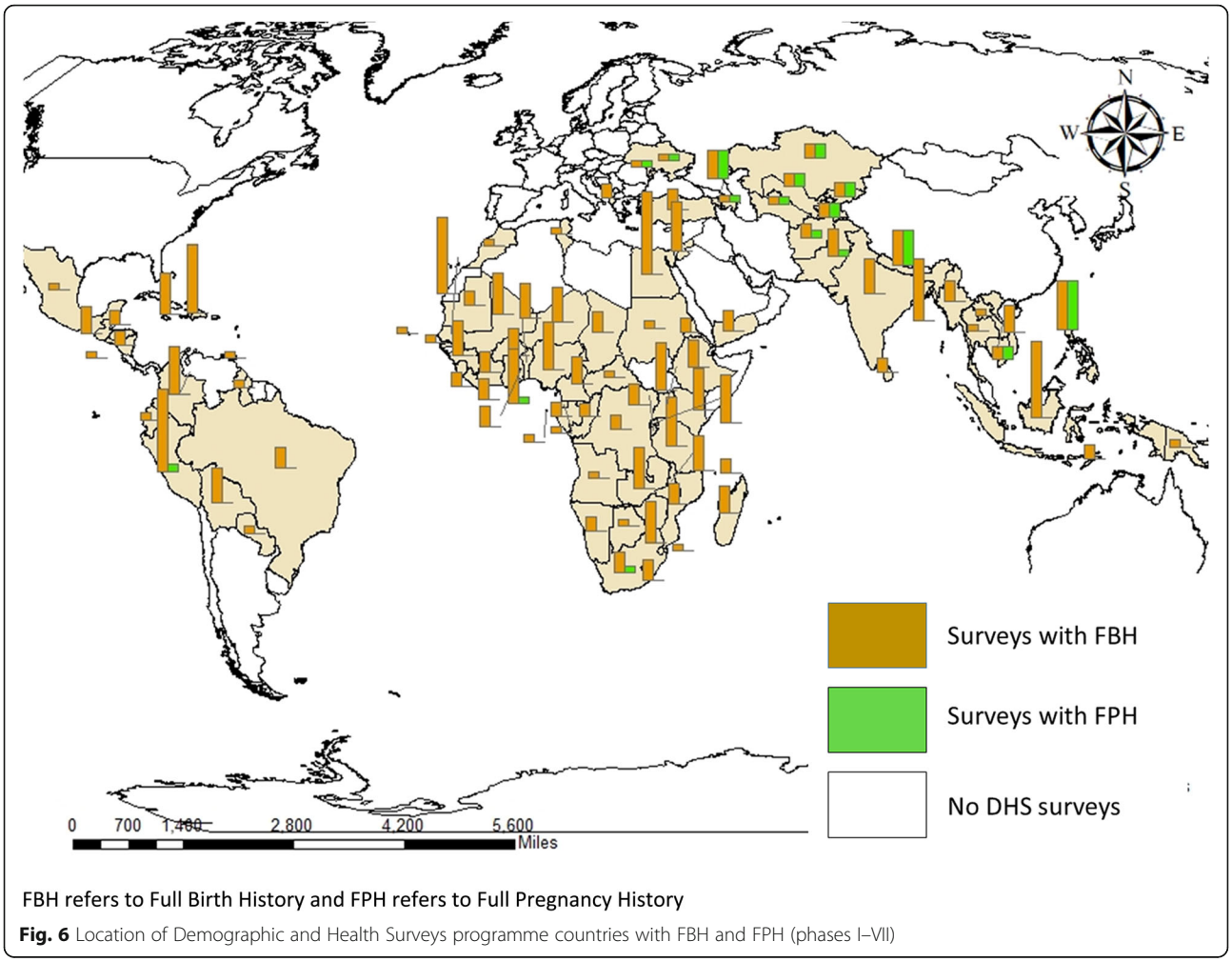


Fig. 6 Location of Demographic and Health Surveys programme countries with FBH and FPH (phases I–VII)

Data quality for stillbirths and neonatal mortality for DHS phases I–VII

Data quality may vary by maternity history approach [42]. Changes to the model DHS questionnaire over time designed to improve stillbirth or neonatal death data quality include the changes in prompts to seek to reduce omission of live births who died from the FBH over phases III–V and including additional questions for non-live births in DHS-VII (Fig. 3).

Overall, there was little variation in the SBR to NMR ratio in surveys with an FBH and reproductive calendar (RC) (DHS II-VI) and a slightly higher variation in surveys with an FBH with additional questions on pregnancy losses (DHS-VII), and all mean SBR:NMR ratios were below one (Table 1, Fig. 7). Only three surveys with an FBH had an SBR:NMR ratio > 1, 0.7% of those with a RC, and 5.9% of those with additional questions (Figs. 8a, b). The SBR to NMR ratio in the FPH was variable, implying that the FPH's data quality was highly variable across DHS-III-VII, and all median SBR:NMR ratios were below one (Fig. 7). Four (11.7%) of the surveys had an SBR:NMR ratio >1 (Table 1 and Fig. 8a, b). We found no evidence of a trend in mean SBR:NMR in FBH or FPH across the ordered DHS (III–VII) (Table 1).

Comparing the mean SBR between FBH and FPH by phase, we found evidence of a difference for estimates obtained in DHS-VII (p = 0.016). Similarly, evidence of a difference in the mean NMR estimates in DHS-V (p = 0.026) and DHS-VII (p = 0.020) was found (Table 1). For neonatal mortality, the ENMR:NMR ratio increased over time for both surveys using an FBH (p = 0.025) and an FPH (p = 0.070) (Table 1).

Discussion

We reviewed the 35-year evolution of Demographic and Health Surveys' model questionnaires since inception of the programme, covering over 90 countries and with more than 350 surveys, and synthesised implications for data on stillbirths and neonatal deaths. To the best of our knowledge, this is the first detailed overview of the measurement of stillbirths and neonatal deaths in DHS

Table 1 DHS phases' summary of neonatal mortality and stillbirth rates by maternity history type

DHS phase	Number of surveys (conducted)	Mean stillbirth rate (95%CI)[3]	Mean neonatal mortality rate (95%CI)[3]	Median SBR:NMR[4]	Median proportion of NND that are day 0-6 ENMR:NMR[5]
Full birth history					
I[1]	29				
II	23	14.08 (10.59–17.56)	24.38 (17.87–30.88)	0.61	0.69
III	54	13.89 (10.09–17.68)	28.13 (22.82–33.42)	0.47	0.69
IV	61	15.81 (13.51–18.11)	30.78 (27.27–34.29)	0.55	0.73
V	50	11.83 (10.21–13.45)	26.21 (22.84–29.57)^[2]	0.43	0.75
VI	59	11.90 (10.23–13.56)	24.63 (21.87–27.39)	0.51	0.78
VII	45	13.16 (10.81–15.53)^[1]	23.79 (20.70–26.89)^[3]	0.61	0.78
Full pregnancy history					
I[1]	1				
II	0				
III	9	15.10 (7.74 – 22.46)	24.94 (15.63–34.25)	0.51	0.74
IV	7	13.6 (8.00 – 19.20)	27.38 (19.09–35.66)	0.49	0.76
V	7	14.16 (7.96 – 20.36)	19.57 (9.70–29.44)^[2]	0.67	0.81
VI	7	12.10 (0.81 – 23.39)	24.66 (6.70–42.58)	0.53	0.77
VII	5	9.18 (2.51 – 15.84)^[1]	18.60 (0.51–36.69)^[3]	0.66	0.80

[1]Data collected, but neonatal mortality rate not disaggregated in the reports
[2]Stillbirth rates only available from surveys in countries with high contraceptive prevalence using model questionnaires 2A and 3A
[3]Proportion test comparing the mean stillbirth rate or mean neonatal mortality rate by survey module in each DHS phase: $p = 0.016$ in ^[1]; $p = 0.026$ in ^[3]; $p = 0.020$ in ^[3]
[4]Trend test comparing SBR:NMR for each maternity history type found "no trend" across the ordered DHS phases (full birth history: $p = 0.949$ and full pregnancy history: $p = 0.317$)
[5]DHS refers to the Demographic and Health Surveys; SB refers to stillbirths; NND refers to neonatal deaths; NMR refers to neonatal mortality rate; SBR refers to stillbirth rate; and ENMR refers to early neonatal mortality rate
*Survey SBR and NMR estimates extracted from DHS STATcompiler

surveys over time examining the impact of different maternity histories modules and reproductive calendar approaches, including the resultant data quality. We focus on stillbirths and neonatal deaths because these are important indicators of maternal health, and universal healthcare access, utilisation and quality.

Most of the surveys used an FBH as provided in the DHS model questionnaire in each of the phases DHS-I to DHS-VII. Some changes have been made to the model questionnaire to improve data on stillbirths and neonatal deaths, most notably the inclusion of a reverse-truncated history for non-live births to capture information on stillbirths in the last 5 years and inclusion of a question on the day of birth in DHS-VII. We also found that the implementation of questions for collection of data on stillbirths in the woman's questionnaire varies, with information captured as part of the main maternity history with an FPH approach, and in sub-section 2.3 for the reproductive calendar or additional non-live birth questions approaches used with an FBH [17].

Both FPH and FBH approaches seek to capture all child deaths including neonatal deaths which comprise around half of under-5 child deaths globally. Whilst FPH also captures information on stillbirths, the FBH has been supplemented with a reproductive calendar or additional questions on non-live births to capture such data in surveys since DHS-II. In practice, stillbirths, miscarriages and terminations of pregnancy are not fully captured, especially in FBH plus reproductive calendar or additional questions in the current DHS-VII [43]. Based on recent qualitative research, we hypothesise that this may be because the women do not report the event or timing, or do not perceive stillbirths as "valued", or avoid reporting due to stigma in the community associating adverse outcomes with evil spirits [44]. Alternatively, the interviewers may omit or mis-record the event deliberately or unconsciously. Early neonatal deaths are live births, but may be misclassified as stillbirths [45] or also not reported, as these are also a source of stigma [46].

Our study found an increasing proportion of neonatal deaths that are in the early neonatal period. Whilst previous studies have reported that this proportion is stable across different levels of NMR, within survey data they also found that earlier surveys had on average lower proportions of deaths in the first week [25]. The changes found in the DHS model questionnaire regarding the capture of neonatal

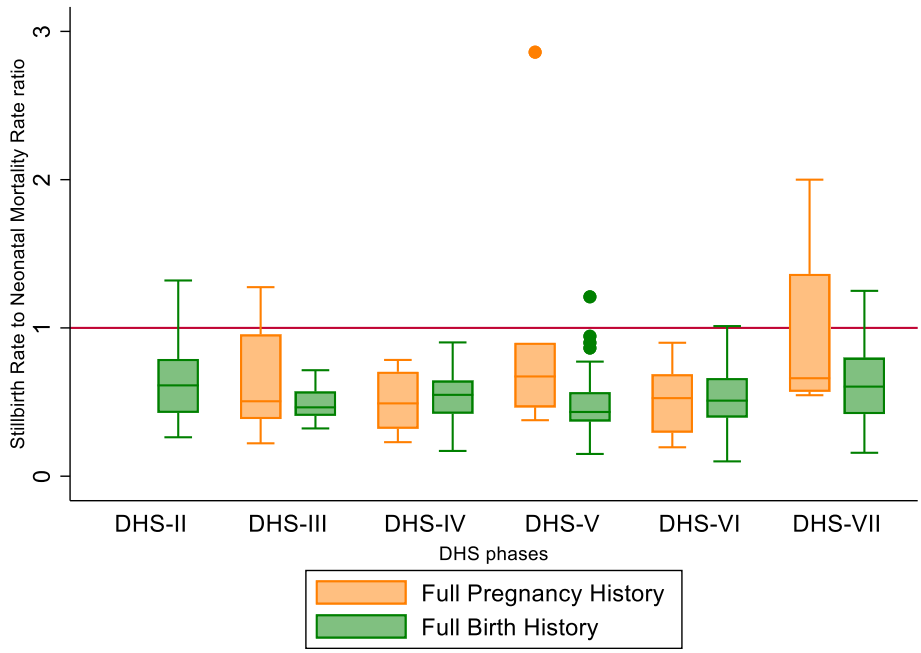


Fig. 7 Distribution of stillbirth tate to neonatal mortality rate ratio by DHS module by phases II–VII

deaths were minor and are unlikely to have influenced the changes in the mortality estimates over time. Hence, observed changes may be due to true epidemiology variation, or changes in measurement such as reductions in the heaping on day 7, reductions in the omission of deaths due to increased societal recognition or changes in misclassification of stillbirths and neonatal deaths.

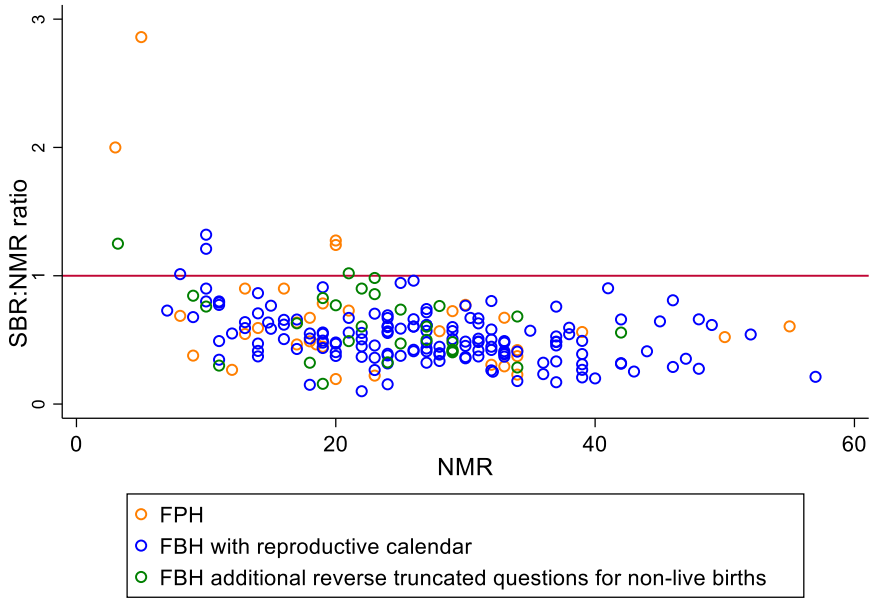
The high levels of variability in the SBR to NMR ratio observed across surveys with an FPH may reflect variation in data quality due to variations in the FPH tool and training between surveys. Whilst the standard FBH from the DHS model questionnaire was used in most FBH surveys, and as shown this varied little over time, no standard tool was available for FPH, and these varied both in questions asked and implementation across different countries.

A strength of our study is that it presents the first overall overview of the capture of stillbirths and neonatal deaths in DHS throughout, reviewing DHS model questionnaires, programme implementation and data on these outcomes throughout the lifetime of the programme from inception in 1984 to 2019. An important limitation of this study is that it only examines question similarities or differences in the questionnaires in terms of numbering, wording and interviewer instructions. In analysing data quality, only compiled data available via the STATcompiler were used; future studies could use survey microdata to further assess data quality aspects, such as heaping on day 7.

The most important limitation is the lack of robust measures for data quality for stillbirth and neonatal mortality rates. For stillbirths, we applied a simple test comparing the SBR to NMR ratio by survey modules, but this is simplistic and any potential variation with the level of NMR and SBR in LMICs is as yet poorly understood [7, 18]. SBR were lower than expected in almost all the surveys conducted by DHS phase, including both FBH and FPH, and this was consistent with findings from the EN-INDEPTH study [42].

A crucial research gap left unanswered by this study is how to better assess data quality for stillbirths. Unlike stillbirths, neonatal deaths have several data quality indicators developed, recommended and used to establish the quality of the data. To be able to improve the data over time requires measures that may be simple and generic, such as completeness, percentage of don't knows, and if relevant (e.g. for continuous variables such as dates or birthweight), measures of heaping. However, for stillbirths, the main quality criterion SBR:NMR remains as yet poorly understood, and work is required to develop further methods to assess data quality in both survey and routine data. These measures would need to be validated against "gold standard" data with accurate measures of stillbirths, which require first trimester gestational age assessment or accurate birthweight and accurate assessment of signs of life at birth.

The DHS programme announced a major shift for DHS phase VIII by including for the first time an FPH



Note: SBR-Stillbirth Rate; NMR-Neonatal Mortality Rate; FBH refers to Full Birth History and FPH refers to Full Pregnancy History

Fig. 8A Comparison of DHS surveys (1990–2018) for SBR:NMR ratios a data quality marker by module

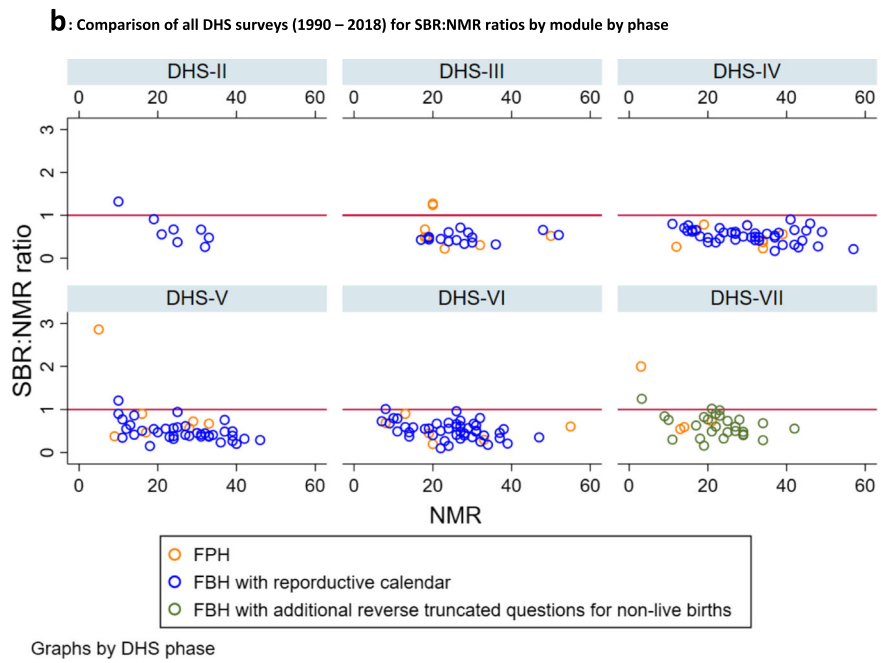


Fig. 8B Comparison of all DHS surveys (1990–2018) for SBR:NMR ratios by module by phase

in its model questionnaire [47]. The change followed findings from the EN-INDEPTH study, a randomised comparison of the DHS-VII model questionnaire's FBH plus additional questions on recent non-live births to the most recent FPH implemented in Nepal. Evidence from that study suggested that FPH improved reporting of stillbirths, but had no effect for neonatal deaths [42]. However, whilst the FPH approach resulted in higher SBR estimates overall, this finding varied across sites underlining that changes to the questionnaire tools may be necessary but not sufficient, with other factors such as survey training and implementation likely playing an important role. In addition, the SBR estimates using an FPH were still lower than expected, assuming that, consistent with previous studies, an SBR:NMR of around 1 would be seen in these populations [18, 20, 21]. Implementation of the survey and also contextual barriers to reporting adverse pregnancy events such as stillbirths must also be considered and addressed [46].

Conclusions

DHS remains the major data source for pregnancy outcomes worldwide. Although the DHS model questionnaire has evolved over the last three and a half decades, more robust evidence is required concerning optimal methods to obtain accurate data on stillbirths and neonatal deaths through household surveys.

The change in DHS-VIII from FBH to FPH is expected to improve the capture of data on pregnancy losses. However, whilst FPH may improve the capture of stillbirths, optimising data quality is also dependent on implementation including training and supervision of data collectors, and addressing contextual barriers to women reporting pregnancy losses. More research is required to develop and test standardised robust data quality measures for stillbirths and neonatal deaths for use in both survey and routine data. Investing in these will ensure that by the end of the Sustainable Development Goals era, countries have more data of higher quality to use for tracking their national targets and reducing these five million preventable neonatal deaths and stillbirths.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12963-020-00225-0.

Additional file 1: Maternal history questions in the FBH and FPH modules (Section II)

Additional file 2: Comparing the DHS's model FBH questionnaire across Phases I-VII

Additional file 3: A. Location of Demographic and Health Survey program countries and number of surveys, Phases I-IV. **B**. Location of Demographic and Health Survey program countries and number of surveys, Phases V-VII

Additional file 4: DHS surveys with Full Pregnancy History modules by DHS phase

Additional file 5: DHS surveys with Full Birth History modules by DHS phase

Abbreviations

DHS: Demographic and Health Surveys; EN-INDEPTH: Every Newborn - International Network for the Demographic Evaluation of Populations and their Health; FBH+: Full birth history (+ denotes additional questions on

pregnancy losses); FPH: Full pregnancy history; LMICs: Low- and middle-income countries; MICS: Multiple Indicator Cluster Survey; NMR(s): Neonatal mortality rate(s); SB: Stillbirths; SBR(s): Stillbirth rate(s); NND: Neonatal deaths

About this supplement

This article has been published as part of *Population Health Metrics Volume 19 Supplement 1, 2021: Every Newborn-INDEPTH study: Improving the measurement of pregnancy outcomes in population-based surveys*. The full contents of the supplement are available online at https://pophealthmetrics.biomedcentral.com/articles/supplements/volume-19-supplement-1.

Authors' contributions

JA together with HB, JEL, SC, PW and AB developed the detailed research questions and overall analysis plan for this paper. Analysis was undertaken by JA. HB, SC and JEL provided statistical oversight. The manuscript was drafted by JA, HB, JEL and SC. HB, JEL, SC, PW, VSG, FA and TC participated in reviewing and helped to revise the manuscript. All authors reviewed and agreed on the final version.

Funding

The EN-INDEPTH study (including publication costs) was funded by the Children's Investment Fund Foundation (CIFF) by means of a grant to LSHTM (PI *Joy E. Lawn*) and a sub-award to the INDEPTH MNCH working group with technical leadership by the Makerere School of Public Health (PI *Peter Waiswa*).

Availability of data and materials

This review used DHS data available from the DHS programme website on request (https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm). All studies included in the review have been cited in the reference section.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare no competing interests.

Author details

[1]Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene & Tropical Medicine, London, UK. [2]Centre of Excellence for Maternal Newborn and Child Health Research, Dept. of Health Policy, Planning and Management, Makerere University School of Public Health, Kampala, Uganda. [3]Global Health Division, Karolinska Institutet, Stockholm, Sweden. [4]Institute of Population Health Sciences, Queen Mary University of London, London, UK. [5]The Demographic and Health Surveys (DHS) Program, ICF, Rockville, USA.

Published: 8 February 2021

References

1. de Bernis L, Kinney MV, Stones W, ten Hoope-Bender P, Vivio D, Leisher SH, Bhutta ZA, Gülmezoglu M, Mathai M, Belizán JM. Stillbirths: ending preventable deaths by 2030. Lancet. 2016;387:703–16.
2. Lawn JE, Blencowe H, Oza S, You DZ, Lee ACC, Waiswa P, Lalli M, Bhutta Z, Barros AJD, Christian P, et al. Every newborn: progress, priorities, and potential beyond survival. Lancet. 2014;384:189–205.
3. World Health Organization: Every newborn: an action plan to end preventable deaths. 2014. http://www.healthynewbornnetwork.org/hnn-content/uploads/Every_Newborn_Action_Plan-ENGLISH_updated_July2014.pdf. [Accessed November 2019].
4. Cousens S, Blencowe H, Stanton C, Chou D, Ahmed S, Steinhardt L, Creanga AA, Tuncalp O, Balsara ZP, Gupta S, et al. National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. Lancet. 2011;377:1319–30.
5. UN Inter-agency Group for Child Mortality Estimation: Levels and trends in child mortality. pp. 8; 2017:8.
6. Heazell AEP, Siassakos D, Blencowe H, Burden C, Bhutta ZA, Cacciatore J, Dang N, Das J, Flenady V, Gold KJ, et al. Stillbirths: economic and psychosocial consequences. Lancet. 2016;387:604–16.
7. Lawn JE, Blencowe H, Waiswa P, Amouzou A, Mathers C, Hogan D, Flenady V, Froen JF, Qureshi ZU, Calderwood C, et al. Stillbirths: rates, risk factors, and acceleration towards 2030. Lancet. 2016;387:587–603.
8. Bhutta Z, Das J, Bahl R, Lawn JE, Salam RA, Paul VK, Sankar MJ, Blencowe H, Rizvi A, Chou V. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? Lancet. 2014;384:347–70.
9. Edouard L. The epidemiology of perinatal mortality. World Health Statistics Quarterly. 1985;38:289–301.
10. Walsh JA, Feifer CM, Measham AR, Gertler PJ: Maternal and perinatal health. 1993.
11. AbouZahr C, De Savigny D, Mikkelsen L, Setel PW, Lozano R, Nichols E, Notzon F, Lopez AD. Civil registration and vital statistics: progress in the data revolution for counting and accountability. Lancet. 2015;386:1373–85.
12. Cleland J. Demographic data collection in less developed countries 1946–1996. Popul Stud. 1996;50:433–50.
13. Boerma JT, Sommerfelt AE. Demographic and Health Surveys (DHS): contributions and limitations. World Health Statistics Quarterly. 1993;46:222–6.
14. Fabic MS, Choi Y, Bird S. A systematic review of Demographic and Health Surveys: data availability and utilization for research. Bull World Health Organ. 2012;90:604–12.
15. DHS Program: Demographic and Health Surveys. 2019. https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm. [Accessed January 2020].
16. Corsi DJ, Neuman M, Finlay JE, Subramanian SV. Demographic and health surveys: a profile. Int J Epidemiol. 2012;41:1602–13.
17. Demographic and Health Surveys: DHS model questionnaires. 2018. https://dhsprogram.com/What-We-Do/Survey-Types/DHS-Questionnaires.cfm. [Accessed November 2019].
18. Blencowe H, Cousens S, Jassir FB, Say L, Chou D, Mathers C, Hogan D, Shiekh S, Qureshi ZU, You D, 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–e108.
19. 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. In DHS Methodological Reports, vol. 17. Rockville: ICF International; 2015.
20. Butler N. Perinatal mortality survey under the auspices of the National Birthday Trust Fund (preliminary communication). Organization and returns. Proc R Soc Med. 1961;54:1089.
21. World Health Organization: Neonatal and perinatal mortality for the year 2000: country, regional and global estimates. 2006. http://apps.who.int/iris/bitstream/handle/10665/43444/9241563206_eng.pdf;jsessionid=0712A8D434A36DF7FD37667960C32E89?sequence=1 [Accessed November 2019].
22. Pullum T, Becker S. Evidence of omission and displacement in DHS birth histories: methodological reports vol. 11. Rockville: ICF International; 2014.
23. UN Inter-agency Group for Child Mortality Estimation: Levels and trends in child mortality. United Nations Inter-agency Group for Child Mortality Estimation (UN IGME); 2019.
24. Oestergaard MZ, Inoue M, Yoshida S, Mahanani WR, Gore FM, Cousens S, Lawn JE, Mathers CD, Grp UNI-A, Epidemiology CH. Neonatal mortality levels for 193 countries in 2009 with trends since 1990: a systematic analysis of progress, projections, and priorities. PLoS Med. 2011;8:e1001080.
25. Oza S, Cousens SN, Lawn JE. Estimation of daily risk of neonatal death, including the day of birth, in 186 countries in 2013: a vital-registration and modelling-based study. Lancet Glob Health. 2014;2:e635–44.
26. World Health Organization: International classification of diseases for mortality and morbidity statistics. 2019. https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1947342847. [Accessed December 2019].
27. Hill K: Indirect estimation of child mortality. Tools for demographic estimation 2013:148-164.
28. Coale AJ, Brass W: The demography of tropical Africa. In Methods of analysis and estimation. Princeton University Press, Princeton, NJ; 1968: 88-139.
29. Rajaratnam JK, Tran LN, Lopez AD, Murray CJL. Measuring under-five mortality: validation of new low-cost methods. PLoS Med. 2010;7:e1000253.
30. Espeut D, Becker S. The validity of birth and pregnancy histories in rural Bangladesh. J Health Popul Nutr. 2015;33:17.
31. Becker S, Mahmud S: A validation study of backward and forward pregnancy histories in Matlab Bangladesh. International Statistical Institute; 1984.
32. Thompson L, Ali MN, Casterline JB: Collecting demographic data in Bangladesh: evidence from tape-recorded interviews. In. International Statistical Institute; 1982.

33. Goldman N, Moreno L, Westoff CF. Peru experimental study: an evaluation of fertility and child health information; 1989.
34. Westoff CF, Goldman N, Moreno L. Dominican republic experimental study an evaluation of fertility and child health information; 1990.
35. Goldman N, Moreno L, Westoff CF. Collection of survey data on contraception - an evaluation of an experiment in Peru. Stud Fam Plan. 1989;20:147–57.
36. Moreno L, Goldman N, Babakol O. Use of a monthly calendar for collecting retrospective data on contraception: an evaluation of the experimental field studies of the Demographic and Health Surveys (DHS). Notas Poblacion. 1991;18:11–37.
37. MacQuarrie KLD, Winfrey W, Meijer-Irons J, Morse A: Consistency of reporting of terminated pregnancies in DHS calendar. In DHS Methodological Reports No 25. pp. 6-7. Rockville, Maryland, USA: ICF; 2018:6-7.
38. Croft T, Marshall AMJ, Allen CK. Guide to DHS statistics. Rockville: ICF International; 2018.
39. Landers A, McNiff M: Comparability of questionnaires. In DHS Methodological Reports, vol. 4. Calverton: Macro International Inc; 1994.
40. Huntington D, Mensch B, Toubia N. A new approach to eliciting information about induced-abortion. Stud Fam Plan. 1993;24:120–4.
41. United States Census Bureau: Census and Survey Processing System (CSPro). 2019. https://www.census.gov/data/software/cspro.html. [Accessed December 2019].
42. Akuze J, Blencowe H, Waiswa P, Baschieri A, Gordeev VS, Kwesiga D, Fisker AB, Thysen SM, Rodrigues A, Biks GA, et al: Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in 69,176 pregnancies in five countries: the Every Newborn-INDEPTH study. Lancet Global Health 2020;8:E555–66.
43. Christou A, Alam A, Hofiani SMS, Rasooly MH, Mubasher A, Rashidi MK, Dibley MJ, Raynes-Greenow C. How community and healthcare provider perceptions, practices and experiences influence reporting, disclosure and data collection on stillbirth: Findings of a qualitative study in Afghanistan. Soc Sci Med. 2019;236:112413.
44. 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;8:24011.
45. Liu L, Kalter HD, Chu Y, Kazmi N, Koffi AK, Amouzou A, Joos O, Munos M, Black RE. Understanding misclassification between neonatal deaths and stillbirths: empirical evidence from Malawi. PLoS One. 2016;11:e0168743.
46. Kwesiga D, Tawiah C, Imam A, Kebede A, Nareeba T, Enuameh YA, Manu G, Beedle A, Fisker A, Waiswa P, et al.: Barriers and enablers to reporting of pregnancy and adverse pregnancy outcomes in population-based surveys: EN-INDEPTH multi-country study [in press]. BMC Population Health Metrics 2020.
47. Demographic and Health Surveys: DHS8 Questionnaire Summary: revision process and new content. DHS; 2019.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

Chapter 4: Methods for the EN-INDEPTH study

4.1. Introduction

This chapter provides a detailed overview of the EN-INDEPTH study, including selection of sites, study design, sampling, and analysis plans.

This chapter was published June 2019 in the Journal of Global Health [36]. The manuscript was published under a Creative Commons Attribution 4.0 International License (CC BY 4.0) - http://creativecommons.org/licenses/by/4.0/ and the published manuscript is included in full below. The web appendix referenced in the paper is available at http://www.jogh.org/documents/issue201901/jogh-09-010901.htm

4.2. List of Figures

Figure 1 - Pregnancy outcomes and neonatal deaths with definitions for international comparison

Figure 2 - Two DHS alternative approaches to estimate neonatal mortality rates and stillbirth rates

Figure 3 – Map showing the location of the EN-INDEPTH study HDSS sites.

Figure 4 – ENDEPTH Data Collection and Flow.

4.3. List of Tables

Table 1 – EN-INDEPTH study summary of research questions and data analysis approach, according to the four study objectives.

Table 2 – Expected sample size across the five INDEPTH sites

Table 3 – Required sample size by stillbirth rate (SBR) for the household survey randomised comparison, assuming alpha = 0.05 and an expected 15% difference in SBR

Table 4 – Current standard DHS Phase 7 questionnaire sections and adaptations for this study

4.4. Citation

Baschieri A., Gordeev, V. S., Akuze J., Kwesiga D., Blencowe H., Cousens S., Waiswa P., Fisker A. B., Thysen S. M., Rodrigues A., Biks G. A., Abebe S. M., Gelaye K. A., Mengistu M. Y., Geremew B. M., Delele T. G., Tesega A. K., Yitayew T. A., Kasasa S., Galiwango E., Natukwatsa, D., Kajungu, D., Enuameh, Y. A., Nettey, O. E., Dzabeng, F., Amenga-Etego, S., Newton, S. K., Manu, A. A., Tawiah, C., Asante, K. P., Owusu-Agyei, S., Alam, N., Haider, M. M., Alam, S. S.,

Arnold, F., Byass, P., Croft, T. N., Herbst, K., Kishor, S., Serbanescu, F., Lawn, J. E. "**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(1): p. 010901.
www.jogh.org doi: 10.7189/jogh.09.010901

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1512635	**Title**	Mr
First Name(s)	Joseph		
Surname/Family Name	Akuze Waiswa		
Thesis Title	Measurement of stillbirths and neonatal deaths in standardised population-based surveys		
Primary Supervisor	Hannah Blencowe		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Journal of global health as: Baschieri Angela, Gordeev Vladimir S, Akuze Joseph, Kwesiga Doris, Blencowe Hannah, Waiswa Peter, & Lawn Joy. (2019). Every Newborn-INDEPTH study protocol for a randomized comparison of household survey modules for measuring pregnancy outcomes in five Health and Demographic Surveillance sites. London School of Hygiene and Tropical Medicine. Journal of Global Health www.jogh.org. doi: 10.7189/jogh.09.0109011; June 2019. Vol. 9 No. 1		
When was the work published?	June 2019		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	Not Applicable		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Not Applicable
Please list the paper's authors in the intended authorship order:	Not Applicable
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The study was conceptualised by Professor Joy Lawn, and the protocol developed and refined by the team including Joy Lawn, Peter Waiswa, Hannah Blencowe, Angela Baschieri, Vladimir Gordeev and myself. I led the development of the questionnaires and supported Vladimir Gordeev with the development and programming of a section of the EN-INDEPTH survey questionnaire in the World Bank's Survey Solutions data collection platform and application. I led the development of various implementation manuals and standard operating procedure documents including study manuals, ethical clearance application and data sharing agreement. I led the training of interviewers in three out of five sites and coordinated the rollout and monitoring of the data collection processes in all five sites. I supported Angela and Vladimir in conducting the background literature reviews and in drafting and revising the manuscript.

SECTION E

Student Signature	Joseph Akuze Waiswa
Date	27th September 2021

Supervisor Signature	Hannah Blencowe
Date	28th September 2021

4.5. Research paper

B: "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

Electronic supplementary material:
The online version of this article contains supplementary material.

journal of global health

© 2019 The Author(s)
JoGH © 2019 EUGHS

"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

Angela Baschieri[1], Vladimir S Gordeev[1], Joseph Akuze[1,2,3], Doris Kwesiga[2,3], Hannah Blencowe[1], Simon Cousens[1], Peter Waiswa[2,3], Ane B Fisker[4,5,6], Sanne M Thysen[4,5,7], Amabelia Rodrigues[4], Gashaw A Biks[8], Solomon M Abebe[8], Kassahun A Gelaye[8], Mezgebu Y Mengistu[8], Bisrat M Geremew[8], Tadesse G Delele[8], Adane K Tesega[8], Temesgen A Yitayew[8], Simon Kasasa[2,9], Edward Galiwango[2,9], Davis Natukwatsa[2,9], Dan Kajungu[2,9], Yeetey AK Enuameh[10,11], Obed E Nettey[11], Francis Dzabeng[11], Seeba Amenga-Etego[11], Sam K Newton[10,11], Alexander A Manu[11], Charlotte Tawiah[11], Kwaku P Asante[11], Seth Owusu-Agyei[1,12,13], Nurul Alam[14], M M Haider[14], Sayed S Alam[14], Fred Arnold[15], Peter Byass[16], Trevor N Croft[15], Kobus Herbst[17], Sunita Kishor[18], Florina Serbanescu[19], Joy E Lawn[1]

[1] Maternal, Adolescent, Reproductive and Child Health (MARCH) Centre, London School of Hygiene &Tropical Medicine, London, United Kingdom
[2] School of Public Health, Makerere University, Kampala, Uganda
[3] INDEPTH Network Maternal, Newborn and Child Health Working Group Technical Secretariat
[4] Bandim Health Project, Bissau, Guinea-Bissau
[5] Research Center for Vitamins and Vaccines, Bandim Health Project, Statens Serum Institut, Copenhagen, Denmark
[6] OPEN, Odense Patient data Explorative Network, Odense University Hospital/Institute of Clinical Research, University of Southern Denmark, Odense, Denmark
[7] Center for Global Health, Department of Public Health, Aarhus University Denmark, Aarhus, Denmark
[8] Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia
[9] IgangaMayuge HDSS, Uganda
[10] Kwame Nkrumah University of Science and Technology, Kumasi, Ashanti, Ghana
[11] Kintampo Health Research Centre, Kintampo, Ghana
[12] University of Health and Allied Sciences, Kintampo Health Research Centre, Kintampo, Ghana
[13] Malaria Centre, London School of Hygiene &Tropical Medicine, London, United Kingdom
[14] Health Systems and Population Studies Division, icddr,b, Dhaka, Bangladesh
[15] ICF, Rockville, Maryland, USA
[16] Department of Epidemiology & Global Health, Umeå University, Umeå, Sweden
[17] Africa Health Research Institute, South Africa
[18] The DHS Program, ICF Rockville, Maryland, USA
[19] Centers for Disease Control and Prevention, Division of reproductive Health, USA

Correspondence to:

Prof Joy E Lawn
Maternal, Adolescent, Reproductive and Child Health (MARCH) Centre
London School of Hygiene & Tropical Medicine
Keppel Street
London, WC1E 7HT
United Kingdom
Joy.Lawn@lshtm.ac.uk

Background Under-five and maternal mortality were halved in the Millennium Development Goals (MDG) era, with slower reductions for 2.6 million neonatal deaths and 2.6 million stillbirths. The Every Newborn Action Plan aims to accelerate progress towards national targets, and includes an ambitious Measurement Improvement Roadmap. Population-based household surveys, notably Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys, are major sources of population-level data on child mortality in countries with weaker civil registration and vital statistics systems, where over two-thirds of global child deaths occur. To estimate neonatal/child mortality and pregnancy outcomes (stillbirths, miscarriages, birthweight, gestational age) the most common direct methods are: (1) the standard DHS-7 with Full Birth History with additional questions on pregnancy losses in the past 5 years (FBH+) or (2) a Full Pregnancy History (FPH). No direct comparison of these two methods has been undertaken, although descriptive analyses suggest that the FBH+ may underestimate mortality rates particularly for stillbirths.

Methods This is the protocol paper for the Every Newborn-INDEPTH study (INDEPTH Network, International Network for the Demographic Evaluation of Populations and their Health Every Newborn, Every Newborn Action Plan), aiming to undertake a randomised comparison of FBH+ and FPH to measure pregnancy outcomes in a household survey in five selected INDEPTH Network sites in Africa and South Asia (Bandim in urban and rural Guinea-Bissau; Dabat in Ethiopia; IgangaMayuge in Uganda; Kintampo in Ghana; Matlab in Bangladesh). The survey will reach >68 000 pregnancies to assess if there is ≥15% difference in stillbirth rates. Additional questions will capture birthweight, gestational age, birth/death certification, termination of pregnancy and fertility intentions. The World Bank's Survey Solutions platform will be tailored for data collection, including recording paradata to evaluate timing. A mixed methods assessment of barriers and enablers to reporting of pregnancy and adverse pregnancy outcomes will be undertaken.

Conclusions This large-scale study is the first randomised comparison of these two methods to capture pregnancy outcomes. Results are expected to inform the evidence base for survey methodology, especially in DHS, regarding capture of stillbirths and other outcomes, notably neonatal deaths, abortions (spontaneous and induced), birthweight and gestational age. In addition, this study will inform strategies to improve health and demographic surveillance capture of neonatal/child mortality and pregnancy outcomes.

Almost nine million women and children die each year, two-thirds during pregnancy and around the time of birth [1]. An estimated 2.6 million babies are stillborn (die in the last three months of pregnancy or during childbirth) [2], 2.6 million liveborn babies die within the first 28 days of life (neonatal deaths) [1], and 303 000 women die of pregnancy complications per year [3]. Whilst child and maternal mortality rates halved during the Millennium Development Goal era, slower progress has been made for preventing stillbirths and neonatal deaths [4]. To accelerate progress, the Every Newborn Action Plan (Every Newborn) was launched in June 2014 [5], including national targets of 12 or fewer neonatal deaths per 1000 live births and 12 or fewer stillbirths per 1000 total births by 2030 [5,6]. Neonatal mortality is also a sub-target under the third Sustainable Development Goal (SDG 3). Both neonatal and stillbirth rates (SBR) are tracked in the United Nation's Global Strategy for Women's, Children's and Adolescent's Health 2016-2030 [6,7]. The countries needing the greatest acceleration to meet these targets are mainly in sub-Saharan Africa and South Asia, with both the highest risk of mortality and the lowest availability of data. To track SDG progress and inform investments towards the Every Newborn targets, data are essential on both coverage of interventions and impact.

In response, the World Health Organization (WHO) and the London School of Hygiene & Tropical Medicine (LSHTM) published an ambitious Every Newborn Measurement Improvement Roadmap. This Roadmap prioritises specific measurement gaps and provides a multi-year, multi-partner pathway to test validity of selected coverage indicators [8], develop tools (eg, improved birth and death registration, audit, minimum perinatal data set, gestational age and birthweight), and promote use of data by 2020 [5,9,10]. The roadmap includes improved measurement and classification of pregnancy outcomes including stillbirths, miscarriage, or termination of pregnancy (TOP) and neonatal deaths. Data on birthweight, gestational age and vital status at birth are critical for correct classification of these outcomes (**Figure 1**) [12].

Population-based household surveys are the major source of population-level data on child and neonatal mortality rates (NMR) in settings without high coverage Civil Registration and Vital Statistics (CRVS) systems, and will continue to be an important source of data in the SDG-era [13,14]. Surveys are also

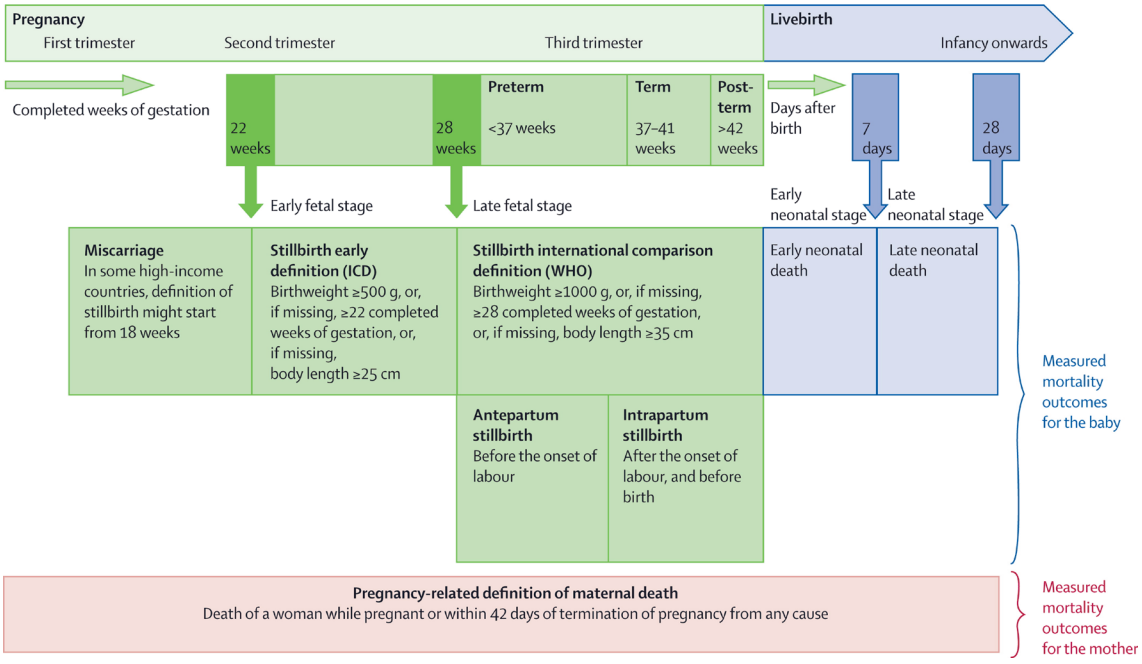


Figure 1. Pregnancy outcomes and neonatal deaths with definitions for international comparison. From [11]. Pregnancy outcomes include miscarriage, stillbirth, termination of pregnancy, gestational age at birth and birthweight. This figure does not include induced termination of pregnancy which are defined as an induced termination of pregnancy by medical or surgical means and this definition be different in countries depending on their law and regulations.

sources for adverse pregnancy outcome data including stillbirths, miscarriages, TOPs, and birthweight, and gestation age.

Birthweight data are collected in surveys from either health cards or maternal recall. Where this information is not available, surveys ask about 'maternal perceived size at birth', which has been used in the estimation of low birthweight rates from surveys [15]. Surveys have a high proportion of missing birthweight data and heaping of reported birthweights [15-17]. Gestational age data are not usually presented in the survey reports, although these data are collected in months in the reproductive calendar for live births and stillbirths, and in additional questions for non-live births (miscarriages and abortions), eg, "How many months pregnant were you when that pregnancy ended?". This answer relies on mother's recall of the length of her pregnancy. The usefulness and validity of these survey data on gestational age in months are not known. More information on gestational age is now increasingly available from health facilities, where the last menstrual period can be recorded by a clinician and may be supplemented with an early ultrasound scan, fundal height during pregnancy, or clinical assessment of the newborn. Birth registration coverage is also assessed through household surveys, and these questions have not been assessed for feasibility or acceptability.

Although DHS collects data on abortions, most of these questions do not distinguish between induced abortions (TOP) or spontaneous abortions (miscarriages) and early/late fetal deaths or stillbirths [18]. This may contribute to undercounting of these different pregnancy outcomes. In some countries like Armenia and Nepal, the reproductive and health surveys go into more detail in collecting data on abortions, including the number of abortions and reasons why. A challenge with accurate capture of induced abortions is a reluctance to report, especially where they are illegal.

Currently, DHS use two alternative approaches to estimate NMR and SBR (**Figure 2**, panel A and panel B):

(1) ***Standard DHS-7 core questionnaire Full Birth History (FBH⁺)***, with additional questions on pregnancies in the last five years resulting in a non-live birth used by a large majority of countries running DHS surveys;

(2) ***Full Pregnancy History (FPH)***, used only in some countries in Eastern Europe and Central Asia, and more recently Afghanistan, Ghana, Nepal, the Philippines, South Africa, Vietnam.

A

Survey Questions	Components
FBH+	Full history of all live births and questions on pregnancies in the last five years resulting in non-live births (including miscarriages, termination of pregnancy and stillbirths). Details of all pregnancies and pregnancy outcomes in the last five years entered in the reproductive calendar (calendar history). Used in reproduction module in DHS-7 model questionnaire.
FPH	Full history of all pregnancies and their outcomes (including live births, miscarriage, termination of pregnancy and stillbirths). Details of all pregnancies and outcomes in the last five years entered in the reproductive calendar (calendar history). Used in the reproduction module in some nationally adapted DHS questionnaires.

B

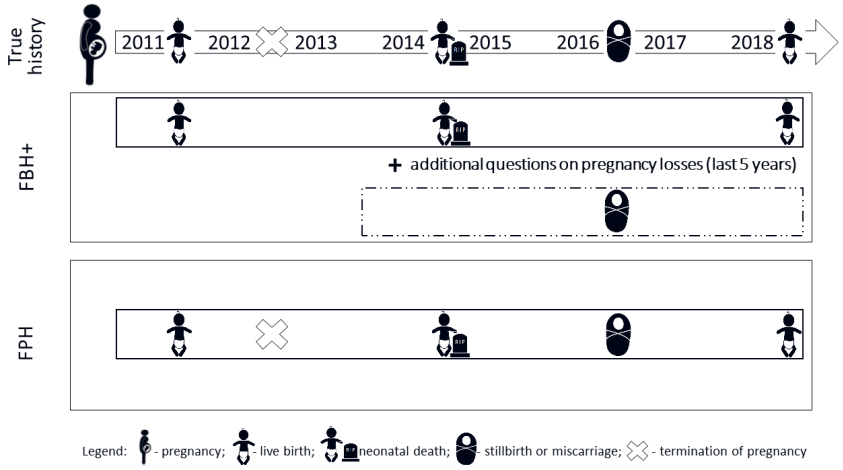


Figure 2. Two DHS alternative approaches to estimate neonatal mortality rates and stillbirth rates. **Panel A.** Full Birth History (FBH⁺) and Full Pregnancy History (FPH) approaches used to collect pregnancy outcomes including stillbirth and neonatal death. DHS-6 and DHS-5 also collected similar information with a Full Birth History, but information on pregnancies not resulting in a live birth were collected in the reproductive calendar only. The new design of DHS-7 questionnaire has additional questions inserted in the questionnaire after the Full Birth History to capture this information. **Panel B.** Data capture by FBH⁺ and FPH methodologies.

The main difference between these two is that in FPH data are collected on all pregnancies in a woman's lifetime; whereas in the FBH⁺ data are collected on all live births in a woman's lifetime and on any pregnancies not resulting in a live birth only for the five years preceding the survey.

Few comparisons of these two methods are available and no rigorous evaluation has been undertaken. No studies have directly compared the performance of the two methods in estimating SBR, but an indirect comparison between two contemporaneous surveys in Ghana: (1) 2008 DHS (using the standard FBH⁺) and (2) 2007 Maternal Health Survey (using FPH) found SBR were 35% lower using the FBH⁺ but there was no difference in Early Neonatal Mortality Rates (ENMR) [19,20] (Appendix S1 in **Online Supplementary Document**). Another study in the Philippines compared two surveys, both using a FPH, found 34% higher SBR when the FPH was part of a short maternal health survey (SBR 12.5 per 1000 total births), compared to when it was administered as part of a full DHS (SBR 9.6 per 1000 total births) [21]. The difference observed may have been due to a shorter, more focused questionnaire in the maternal health survey and not by the question structure itself. Only one published study has made a direct (non-randomised) comparison between the two methods for ENMR, and found 2%-3% higher rates using the FPH [22].

Whilst it is plausible that a FPH may yield improved capture of adverse pregnancy outcomes, it may be more time consuming and hence, data are required to show evidence of better capture before recommending this as standard in the already long DHS core questionnaire. In addition, it is important to understand interview procedures and barriers and enablers to reporting pregnancies and adverse outcomes, which have been understudied. Data from Tanzania suggests that socio-economic and cultural factors affect the quality of information collected on adverse pregnancy outcomes [23]. An analysis of 39 DHS suggested that fieldwork procedures affect data quality, notably including sex of the interviewer; whether or not a translator was used; the timing of the interview; and how many call-backs an interviewer had to make [24]. Some studies have reported on the stigma for women after stillbirths, newborn deaths and abortions in developing country contexts [25-27].

Aim

This is the protocol paper for the **E**very **N**ewborn-**INDEPTH** Network (EN-INDEPTH) study which is part of the Every Newborn Measurement Improvement Roadmap, and aims to improve household sur-

vey capture of stillbirths and neonatal deaths by assessing whether FPH leads to increased capture of selected pregnancy outcomes compared to the standard DHS-7 FBH+ (**Figure 2**, panel A and panel B). The study will investigate the performance of existing or modified survey questions regarding other important measures related to pregnancy related outcomes, including fertility intentions, TOP, birthweight, gestational age, and coverage of birth and death certification. In addition, the study will examine barriers and enablers to reporting of pregnancy and adverse pregnancy outcomes in surveys and through Health and Demographic Surveillance Systems (HDSS).

Research objectives

Research objectives, research questions and data analysis methods are summarised in **Table 1.**

Objective 1. FBH+ vs FPH approach: To undertake a randomised comparison of the reproductive module used in the latest version of FBH+ vs a FPH module to examine the variation in capture of stillbirths and neonatal deaths.

Table 1. EN-INDEPTH study summary of research questions and data analysis approach, according to the four study objectives.

Research objective	Research question	Data analysis approach
Objective 1: Full Birth History (FBH+) approach vs Full Pregnancy History (FPH) approach		
To undertake a randomised comparison of the reproductive module used in the latest version of FBH+ vs a FPH module to examine the variation in capture of stillbirths and neonatal deaths.	Is the FPH method better at capturing stillbirths and neonatal deaths in the last five years than the FBH+?	Descriptive and bivariate analyses comparing the two methods including meta-analysis: SBR; and NMR.
	How long does it take to collect data using the FPH questionnaire? Does the length of data collection vary by context and/or fertility level?	Bivariate analyses of the FPH and FBH+ by the time spent answering the questionnaires, variation by context and maternal characteristics.
Objective 2: Pregnancy outcomes		
To evaluate the use of existing/modified survey questions to capture the fertility intentions and selected pregnancy outcomes (top, miscarriage, birthweight, gestational age), as well as birth and death certification.	What is the answerability and data quality by indicator?	Descriptive analyses of selected indicators, and assessment of data quality per indicator (eg, non-response, heaping, missingness).
	How long does it take to collect data regarding these indicators? Does the length of data collection vary by data collector context and/or fertility level?	Analyses of survey paradata to assess variation by data collector (eg, gender, education level and training), time of day, rural/urban location, and time needed to complete survey questions and sections, frequency of repeated corrections of answers to questions.
Objective 3: Survey vs HDSS data collection platforms		
To compare the capture of pregnancy outcomes in the survey to that in the routine HDSS data collection	How do outcomes reported in the EN-INDEPTH survey compare with HDSS data?	Assess level of agreement at population-level between survey and routine HDSS data over the same time period for several indicators: SBR, NMR, miscarriage, TOP, birthweight, GA.
		For individually linked data, compare capture of pregnancy outcomes between survey and HDSS and assess predictors of capture.
Objective 4: Barriers and enablers to reporting (adverse) pregnancy outcomes		
To identify barriers and enablers to the reporting of pregnancy and adverse pregnancy outcomes during the survey and HDSS data collection, and particularly if these differ for the two survey questionnaire methods (FBH+ and FPH).	What are barriers and enablers to reporting of pregnancies and pregnancy outcomes (geographic, socioeconomic, cultural, data collection methodologies, etc.) in HDSS and survey data collection?	Quantitative analyses.
		Qualitative analyses of FGDs or IDIs for:
	What are interviewers' perceptions (both HDSS and survey interviewers) of barriers and enablers to collect data on pregnancy losses in survey setting?	- survey interviewers
		- HDSS interviewers
		- supervisors
		- mothers who had a pregnancy in the past five years
	What are women's perceptions and barriers for reporting pregnancy losses?	A priori coding. Use of the grounded theory and identify emerging themes and outliers; relationships and theories.
	How can data collection process be improved to obtain better data on adverse pregnancy outcomes?	

FBH+ – Full Birth History+; FPH – Full Pregnancy History; SBR – stillbirth rates ; NMR – neonatal mortality rates; TOP – termination of pregnancy; HDSS – Health and Demographic Surveillance Systems; GA – gestation age; FGDs – focus group discussions ; IDIs – in-depth interviews

Objective 2. **Pregnancy outcomes**: To evaluate the use of existing/modified survey questions to capture the fertility intentions and selected pregnancy outcomes (TOP, miscarriage, birthweight, gestational age), as well as birth and death certification.

Objective 3. Survey vs HDSS data collection platforms: To compare the capture of pregnancy outcomes in the survey to that in the routine HDSS data collection.

Objective 4. Barriers and enablers to reporting (adverse) pregnancy outcomes: To identify barriers and enablers to the reporting of pregnancy and adverse pregnancy outcomes during the survey and HDSS data collection, and particularly if these differ for the two survey questionnaire methods (FBH+ and FPH).

METHODS

Study design

This multi-site study will use a retrospective survey to compare two methods of recording pregnancy outcomes (FBH+ vs FPH methods), with random allocation at the individual woman level. Quantitative and qualitative data will be collected to answer four research objectives (**Table 1**).

Research protocol development was informed by wide consultation, including review by the research site teams and an Expert Advisory Group. 23 participants took part in a study design workshop in Kampala in mid-2016 [28]. In addition, in April 2017, a multi-site workshop was organised to agree on the data collection protocol in the five sites.

Study settings

The INDEPTH Network's HDSS is a network of research sites in 53 countries, established in 1998. Each site tracks vital events in a defined population on a continuous basis, but methods used vary [29]. Since some sites undertake pregnancy surveillance, the INDEPTH network provides an ideal platform for this multi-site retrospective population-based survey with a potential for linking with prospective HDSS data. The INDEPTH Network operates through Working Groups, one of which, the Maternal Newborn and Child Health (MNCH) working group [30], hosted by the Makerere University School of Public Health (MakSPH), Uganda, will coordinate this study, partnering with LSHTM.

A Request For Applications (RFA) was sent to all 53 HDSS sites in December 2015 by the INDEPTH Network secretariat. Fourteen proposals were received and reviewed by LSHTM, MakSPH and the INDEPTH Network secretariat [31]. The selection criteria were: HDSS total population of more than 30 000 people; annual SBR and NMR greater than 15 per 1000 total births; acceptable quality surveillance for pregnancy outcomes, including neonatal deaths and stillbirths; expertise on maternal and newborn health, and stillbirths among the HDSS team members and evidence of co-funding in the submitted estimated budgets. Five sites (Bandim, Dabat, IgangaMayuge, Kintampo, and Matlab) were selected (**Table 2** and **Figure 3**), all of which, as well as LSHTM, received ethical approval from local institutional review boards (Appendix S2 in **Online Supplementary Document**).

While all the selected HDSS sites undertake pregnancy surveillance, the quality of which is influenced by various factors including: frequency of surveillance rounds; the key informants ratio to population; gender; the method used to ask about pregnancy and pregnancy outcomes; the proportion of facility births; and linkage between HDSS and facility data. Cultural norms around pregnancy disclosure may also vary and affect the data collected (**Table 2**).

Sample size and data collection approach

Limited evidence from previous descriptive studies discussed above suggests a difference in SBR from 3 to 35% between a FBH+ and a FPH method [19,20]. A difference of at least 15% between the two methods could be sufficient to consider a major change in the DHS core questionnaire. Since this is the first direct comparisons of the two methodologies, we powered the randomized comparison to capture a 15% difference assuming that a difference less than 15% will not justify a change in survey methodology (see Appendix S1 in **Online Supplementary Document**).

Based on the recorded SBR and total births per HDSS over the last three years, the overall SBR for the final sample of births captured in the survey is expected to be around 28.4 per 1000 total births, if all women

Table 2. Expected sample size across the five INDEPTH sites

Characteristics	Bandim	Matlab	Kintampo	Dabat	Igangamayuge	Across the five sites
Estimated number of total births over the five years captured within the HDSS	29 173	25 799	24 008	7031	11 489	97 499
Surveillance system	Bi-annual update rounds in the rural area and monthly updates in urban area. Update rounds includes registration of pregnancies	Two monthly update rounds including pregnancy testing and registration	Bi-annual update, from 2017 they have shifted to an annual update rounds	Bi-annual update rounds. Monthly updates of births and deaths from local guides.	Bi-annual update rounds. Monthly updates of births and deaths from local scouts.	
Sampling frame*	Women in HDSS site with recorded birth outcome in last 5 y. (all in urban site and 80% in rural site)	Women in HDSS site with recorded birth outcome in last 5 y. (all)	Women in HDSS site with recorded birth outcome in last 5 y. (random sample)	Women of reproductive age in HDSS site	Women of reproductive age in HDSS site	
Expected number of total births to be captured in survey	17 000	21 000	14 500	5700	9800	68 000

HDSS – Health and Demographic Surveillance Systems

*See Appendix S3 in **Online Supplementary Document** for details.

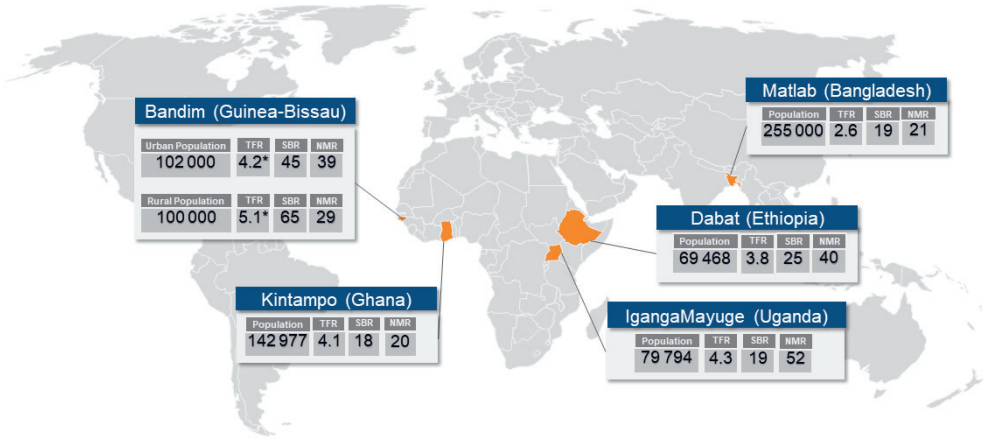


Figure 3. Map showing the location of the EN-INDEPTH study HDSS sites. Total fertility rate (TFR) for women ages 15-49; neonatal mortality rate (NMR) per 1000 live births; stillbirth rate (SBR) per 1000 live births. More detailed information on study HDSS sites: Bandim (http://www.indepth-network.org/member-centres/bandim-hdss); Dabat (http://www.indepth-network.org/member-centres/dabat-hdss); Iganga-Mayuge (http://www.indepth-network.org/member-centres/igangamayuge); Kintampo (http://www.indepth-network.org/member-centres/kintampo-hdss); Matlab (http://www.indepth-network.org/member-centres/matlab-hdss). Asterisk: Bandim - children/pregnancies only followed prospectively; TFR estimated by cumulative birth hazards (Nelson Ahlen) as observed for specific age bands between 2012-16 extrapolated to age span 15-50 years; SBR, NMR estimated among registered pregnancies ending in 2012-16.

are surveyed. However, we needed to account for possible outmigration, unavailability of women at any visit, or non-consent. After adjusting for these factors, we expect the SBR for the FPH arm to be around 26.7 per 1000 total births (see Appendix S3, Table S2.2 in **Online Supplementary Document** for details). Assuming a SBR of 26.7 per 1000 total births in the FPH arm, a total sample size of at least 68 000 births would be required across the five sites to have 80% power to detect a difference of 15% or more between the proportion of total births that are stillbirths in the FBH$^+$ and the FPH at the 5% significance level (alpha = 0.05), including a small design effect (DEFF = 1.1), as stillbirths may be clustered in individual women. The lower the SBR captured in the FBH$^+$ arm, the higher the sample size will be required to detect 15% difference between the two arms (see **Table 3**).

Table 3. Required sample size by stillbirth rate (SBR) for the household survey randomised comparison, assuming alpha = 0.05 and an expected 15% difference in SBR*

Assumed SBR in birth history group/ 1000 total births	Predicted SBR in pregnancy history group/ 1000 total births	Number of total births to achieve 80% power	Sample size - number of births required including design effect and non-response (15%)
23.00	26.45	63604	80459
23.20	26.70	62348	68583
24.00	27.60	60886	77021
25.00	28.75	58386	73858
26.00	29.90	56078	70939
27.00	31.05	53942	68237
28.00	32.20	51958	65727

*The Design effect (DEFF) is calculated as DEFF = 1 + (r – 1) × rho, where r = average number of observations in a cluster and rho = correlation between pairs of observations selected at random from the same cluster). Assuming in a 5-year period women will experience on average a maximum of 2 births, and that as stillbirth is a rare outcome rho<0.1, a design effect of 1.1 is included. For Bandim, due to the challenge of reaching women in rural area, we can only account for a maximum of two visits to reach the interview, for this reason we have assumed a higher rate of non-response rate (30%).

For each site the approach to reach the maximum number of births was agreed based on site-specific information on the level of migration, geographical accessibility, and assumed levels of non-response (either due to non-consent or inability to locate) and other factors. The numbers per site are shown in **Table 2** and details of the estimation are in Appendix S3, Table S2.2 in **Online Supplementary Document**. Three sites (ie, Matlab, Dabat and IgangaMayuge) will survey all women who have given birth in the last five years. Matlab has bi-monthly data collection with pregnancy testing for women who report having missed a menstrual period, and hence has reliable listings of total births, and will administer the EN-INDEPTH survey to eligible women (age 15-49) known to have had at least one birth (live birth or stillbirth) in the last five years; Dabat and IgangaMayuge have bi-annual updates and do not have a specific system of pregnancy tracking and therefore, all women of reproductive age (15-49 years) in the HDSS site will be included. Bandim and Kintampo, will select a random sample of women residing in the HDSS known to have at least one birth in the last five years (**Table 2**).

Data collection software application

Currently, the DHS and Multiple Indicator Cluster Survey (MICS) surveys are administered either using a paper version of the questionnaire (PAPI) or computer assisted personal interviewing (CAPI) and the CSPro (Census and Survey Processing System) platform for entering the data entry or data capture. We elected to use Android-based tablet data collection as most of the selected HDSS sites had some experience with these. We compared several existing data collection platforms (ODK, CSPro, Qualtrics, Red Hat and Survey Solutions). We selected the Survey Solutions Computer Assisted Personal Interview and Computer Assisted Web Interviewing platform (Survey Solutions) developed by the World Bank [32], given its ability of linking questions to specific household members listed in the roster file; its flexible and user-friendly online interface for the questionnaire design; the ability to integrate validation rules; lack of user fees; the online tool for centralised survey administration and data management; data aggregation and reporting features; end-user technical support provided by the World Bank technical support team; as well as the Interviewer application available for Android devices.

The data will be entered on Android-based tablets using the Survey Solution platform, stored locally on the tablets, and synchronised regularly to the dedicated country's physical or virtual server. The platform has different user roles with varying level of permissions and functions: Interviewer (function – data collection), Supervisor (assigning and monitoring data collection by interviewers), Headquarters (overall survey and data management), and Observer (monitoring) (**Figure 4**). The platform's survey and data management component ('Survey Solutions Headquarters', HQ) provides a dashboard of the progress of data collection, including duration and speed of the interviews, Global Positioning System coordinates of the interviewer, as well as paradata. The paradata contains information about the process of collecting survey data and records all events with timestamps on a tablet during data collection (data entry, data correction, responsibility changes, etc.). This type of data can be used for analysis of time per interview and time per question and section, as well as changes in productivity over time for different interviewers and teams. These data can be used for live data quality control, data monitoring and interview progress evaluation.

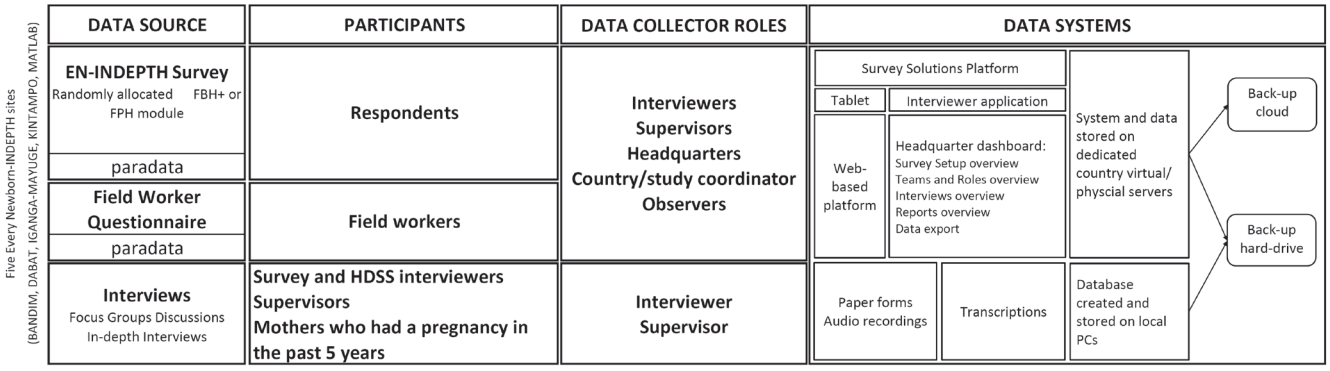


Figure 4. EN-INDEPTH Data Collection and Flow.

Procedures

Informed consent and respondent identification

All participants will receive a verbal explanation of the study objective by a trained field assistant and an information sheet. An adapted version of the standard DHS consent form will be used (Appendix S4 in **Online Supplementary Document**). Both will be translated into local languages and written consent will be obtained from all study participants. Study participants will be informed of their right to refuse and/or withdraw at any point of time from the survey interviews, and at liberty to answer those questions they want to and will not be coerced in case they refuse. Information obtained will be viewed only by the researchers and anonymity will be guaranteed by using identification numbers. Separate written consent will be sought for women and interviewers participating in the qualitative aspects. We will follow fieldwork procedure as outlined in the DHS interviewer manual. Interviewers will make up to three attempts to find respondents.

Randomisation and EN-INDEPTH survey

Women selected for the survey from all five sites will be randomised to receive either the reproduction questions from the DHS-7 women's questionnaire with a standard FBH$^+$ or a FPH. Randomisation will be done automatically within the application-based questionnaire using an in-built random number generator.

Table 4 provides a summary of adaptations to the DHS-7 questionnaire sections to meet the study objectives including questions on gestational age, birth certification, characteristics of stillbirths and fertility intentions. The latest DHS questionnaire (DHS-7) will be used, with a shortened introduction section 1 of the DHS-7 women's questionnaire; section 2 on reproduction; section 4 on pregnancy, delivery and post-

Table 4. Current standard DHS Phase 7 questionnaire sections and adaptations for this study

Survey questionnaire with FBH$^+$ or FPH detailing adaptations from standard DHS phase 7 section, where applicable	
Section 1. Consent form and background of interview	The content of section 1 will be reduced to focus on key maternal background characteristics only.
Section 2. Reproduction	Standard FBH$^+$ questions with pregnancy loss questions which include information on stillbirths, miscarriages and abortions. No adaptations made. Or randomly allocated to FPH from Nepal 2016 for the FPH and pregnancy losses questions – the detailed questions on abortion will only be administered in selected sites.
Section 4. Pregnancy and postnatal care	Section 4 will be administered with minor adaptations for all stillbirths and neonatal deaths, as well as for a sample of live births. Additional questions on gestational age (weeks), birth and death certification, and timing and characteristics of stillbirths will be added to test the feasibility of these questions in household surveys.
Section 8. Fertility preferences	Some questions on fertility intention to refine the measurement on unwanted pregnancy will be added. These questions have been developed and tested in a multi-country research study [33].
Section 9. Household characteristics	Questions on household socio-economic characteristics including household dwelling structure, flooring material, sanitation and toilet facility. These questions are adapted from the DHS household survey questionnaire.

FBH$^+$ – full birth history +; FPH – full pregnancy history; DHS – Demographic and Health Surveys.

natal care; and an additional section 9 adapted on household characteristics. Section 8 on fertility intentions was optional and the Matlab site opted not to administer this as they have just recently completed the collection of similar data. Some specific country add-on modifications will be made to the questionnaire (eg, Bandim site will collect questions on vaccinations, and the IgangaMayuge site will collect data using the Dietary Diversity questionnaire in section 10). Apart from section 2, the questionnaires for the FBH$^+$ and FPH arms will be identical. Section 2 in the FBH$^+$ arm will contain the standard DHS-7 core reproduction questions. In the FPH arm, section 2 will contain the whole reproduction section from the Nepal 2016 DHS for sites willing to include detailed questions on abortion, and a reduced version without the detailed abortion questions for other sites. Appendix S5 in **Online Supplementary Document** gives full details of the questionnaires.

Local language translations of the questions already used by DHS will be used whenever available. Where not available, translations to local languages will be made by the site teams and checked using back translation.

Assessment of barriers and enablers to reporting of pregnancy and adverse pregnancy outcomes

For objective 4, a self-administered questionnaire adapted from DHS fieldworker questionnaire, will assess demographic and other characteristics of the interviewers (Appendix S6 in **Online Supplementary Document**). A series of Focus Group Discussions (FGDs) will be held with HDSS and survey interviewers, as well as supervisors, to assess the barriers or enablers to collecting data on pregnancy and adverse pregnancy outcomes in the survey and HDSS. FGDs will be conducted with women who had at least one pregnancy in the past five years, focusing on perceptions, practices, barriers and enablers in the community. A minimum of six FGDs will be conducted in each site, each with approximately eight to ten participants. In-Depth Interviews (IDIs) will be undertaken in some sites with women who have experienced these adverse pregnancy outcomes, allowing for deeper exploration and triangulation of data.

Training of data collectors and supervisors

The EN-INDEPTH site teams with LSHTM and MakSPH jointly developed a training manual on the data collection procedures, adapting the standard DHS Interviewer's Manual [34], and tailored it to the specific country context and the HDSS site. Four additional manuals (data collection setup, Survey Solutions data management procedures, listing process, Survey Solutions Tester/Interviewer application) were developed adapting the World Bank Survey Solutions Manuals [35]. The training of data collectors and supervisors was led by the HDSS team with initial support from the core team for a minimum of two weeks in all HDSS sites. The training included at least one-week on the paper-based questionnaire and one to two weeks on tablet use, data collection using the Survey Solutions Tester application, as well as the Survey Solutions Interviewer application and interviewer field practice. Prior to data collection, additional trainings on survey management using Survey Solutions HQ were provided to supervisors and data analysists.

For the qualitative work, training manuals will be developed to guide the interviewers during the FGDs and IDIs to ensure comparability of interviews across sites. Additionally, a protocol will be developed to guide the interviewers on how to react in situations where the respondent gets distressed. After the initial two to three weeks training, all sites will initiate the pilot phase of data collection. The length of the data collection will vary by site depending on the fieldwork schedule and allocated sample size ranging between six to 12 months.

Data quality monitoring

Validations for value ranges were defined and programmed into the tablet application to avert predictable human errors. The skipping rules were programmed and additional rules were set to perform consistency checks. Warning messages were programmed to prompt to correct the input when values are outside the defined range, and to provide guidance as per the DHS manual. Data quality will be monitored using Survey Solutions platform's online data dashboard, providing real-time cumulative and detailed summary of ongoing surveys across teams and individual interviewers in each country. The platform allows Supervisors and Headquarters to validate collected data by Interviewer online and, if necessary, incomplete or erroneous questionnaires can be returned to the Interviewer for timely re-assignment and correction. In addition, bi-weekly reports will be sent to the LSHTM and MakSPH core teams by data analysts from all sites summarising the overall data collection progress. Regular all-site data monitoring calls will provide an opportunity for country teams to review and discuss progress in addition to promoting collaborative quality improvement initiatives between countries and sites.

Data management

Following synchronisation, data from tablets will be uploaded to the country's dedicated virtual or physical server with regular automatic back-up, with additional back-up on a separate server or external hard drive. The raw data will be stored in an encrypted format, accessed only by the country's data manager. The anonymization of the quantitative and qualitative data (removing any direct or indirect identifiers, including enumeration identifiers, geo-referenced data, transcripts and audio recording) will take place in-country before data sets are pooled into one multi-site data set (**Figure 4**).

Analysis by objective

The overview of research objectives, main research questions, and analytical approaches are summarised in **Table 1**. For all study objectives, the primarily analyses will be performed overall across countries (as pooled analyses), and comparative secondary analyses will be performed by site separately, whenever possible. Data will be cleaned according to an agreed protocol, including logical and completeness checks. Quantitative analyses will be undertaken with Stata 15SE (Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC). Qualitative analyses will be conducted using NVivo software (NVivo qualitative data analysis Software; QSR International Pty Ltd Version 12, 2018).

Objective 1

A population-level descriptive analysis will be conducted comparing SBR and NMR by FBH+ or FPH, and by maternal characteristics (age, parity, residence, and education status). Crude risk ratios with its 95% confidence interval will be computed for comparison of SBR and NMR between FBH+ and FPH overall and by study site using the meta-analysis methods with Random Effects. Regression models will be fitted to assess determinants of adverse pregnancy outcomes using Generalised Estimation Equations to adjust for clustering of stillbirths or neonatal deaths within the same woman, therefore taking into account design effect. We will use paradata to assess differences in average time taken to complete the FBH+ and FPH.

Objective 2

We will undertake descriptive analyses of selected indicators including fertility intentions, selected pregnancy outcomes (TOP, miscarriage, birthweight, gestational age), as well as birth and death registration. This will include estimates of frequency of reported TOP and miscarriage, coverage of reported birthweight, gestational age, birth and death certification, and of fertility intentions. The answerability of new/refined questions will be assessed by describing patterns of non-response and heaping, where appropriate. Variation in these indicators by pregnancy outcome, maternal or interviewer characteristics will be assessed. For gestational age, internal consistency in the survey will be assessed by comparing women's reporting of gestation in months compared to weeks, and reported maternal perception of the birth to be preterm. Survey paradata will also be analysed to assess time taken to complete and frequency of repeated corrections to relevant survey questions.

Objective 3

Women-level data from the survey will be individually matched with the routine HDSS surveillance data to establish matching rates for stillbirths and neonatal deaths. We will assess determinants of reporting or not reporting of these outcomes in the survey by women's and interviewers' characteristics and HDSS settings (geographic, socioeconomic, cultural, data collection methodologies, etc.). We will also assess levels of agreement between the survey and the routine HDSS data over the same time period at a population level for several indicators, such as SBR, NMR, miscarriage, TOP, birthweight, and gestational age. Predictors of disagreement (eg, length of recall, maternal education, etc.) will be examined. We recognize that neither HDSS nor survey data can be considered 'gold standard' and that the difference in measurement might be in both or either direction and variable by site.

Objective 4

For qualitative data, we will use a grounded theory approach for analysis, with an iterative process involving reading the text, detecting potential emerging themes and outliers, comparing themes and searching for relationships, as well as building theoretical models. A priori coding will be done, with a codebook listing potential codes developed before the analysis begins, to guide the process, and new codes identified from data included as analysis proceeds. Results will be presented with verbatim

quotes from respondents. Reliability will be checked by multiple members of the team, two from each site, independently coding data.

DISCUSSION

The EN-INDEPTH is the first randomised comparison of two survey methods to capture pregnancy outcomes, the current DHS-7 FBH+ and FPH. This is a large-scale study (at least 68 000 births) based in five high-burden countries, including one site in South Asia and four sites across West and East Africa. The study is powered to be able to detect a 15% difference in the estimated SBR, but it is also expected to inform our understanding of survey capture of other pregnancy outcomes, notably neonatal deaths, birthweight, gestational age and abortions (spontaneous and induced). Even if the results show a convincing increase in capture of stillbirths or other pregnancy outcomes, a key operational question is whether the FPH takes longer. The software used for our study (Survey Solutions) allows recording of the paradata, including precise timing by section of the questionnaire. This study will, therefore, enable us to conduct more detailed analyses of time spent by question and section, as well as by the fertility context. Underreporting of pregnancy and adverse outcomes may be affected by socio-cultural barriers and survey data collection procedures, so the mixed methods assessment of barriers and enablers to reporting and recording will be valuable.

In addition to omission of events, household surveys are known to have important limitations in the measurement of stillbirths and neonatal deaths, including displacement of reported day of death and misclassification between stillbirths and neonatal deaths [36,37]. Distinguishing between stillbirths and neonatal deaths requires detection of signs of life at birth, notably assessment of heart rate. Recall by a mother in a survey requires her to know whether there were signs of life at birth and for her to report this. Differences in assessment at birth, perceptions of viability, availability of neonatal resuscitation and cultural and religious factors - all potentially have a role in whether a mother will report her baby's death as a stillbirth or an early neonatal death. This study will examine any differences between reporting of these outcomes in the survey compared to HDSS data, and explore women's perceptions of stillbirth and neonatal deaths, but will not have the ability to assess "true" stillbirths based on lack of accurate measures of heart rate at birth. Another important misclassification is between early fetal death and late fetal death/stillbirth with a threshold of 28 weeks, based on errors in gestational age measurement and reporting. Again, this study will not have "true" gestational age based on first trimester ultrasound.

Birthweight and gestational age measurements are important from individual, clinical and public health perspectives. From a clinical perspective, they are important to identify liveborn neonates at increased risk of mortality and morbidity, for example, those preterm (born at <37 completed weeks of gestation) or low birthweight (<2500g), to enable provision of extra care [12]. From a classification perspective, this information is critical to differentiate between miscarriages and stillbirths. Studies have shown that data on perceived size at birth recorded in surveys are not consistent with data recorded from health cards and that the quality of recalled birthweight data are variable [38]. In addition, little is known on community perceptions of the importance of birthweight and how this may influence reporting. This study seeks to provide further insights on how to obtain better birthweight data in surveys. Although a gestational age in months is collected in the five-year reproductive calendar in DHS surveys, concerns have been raised regarding the validity of these data. Whilst months are used to differentiate between miscarriages and stillbirths, they are currently not reported in most survey reports and are not used in the estimation of population-level measures, such as preterm birth. In this study, we will assess standard and modified questions that seek to capture gestational age, as well as the internal validity of the reporting of gestational age in months, in weeks, and reported maternal perception of the birth as preterm. In the Matlab site, where more accurate gestational age information is captured in the HDSS through early routine urinary pregnancy testing following a missed period, we will assess the validity of these questions by comparing the information captured in the EN-INDEPTH survey to the HDSS data. This will provide important new information on the feasibility of the use of these questions in household surveys. However, the frequent HDSS household visits might increase women's awareness of pregnancy duration and improve the reporting of gestational age.

Comparable data on abortion are limited [39]. In countries where abortion is illegal, data are under-reported for fear of prosecution or stigma [18]. In contexts where abortion is legal, data may also be problematic. This study will add to the literature by testing the feasibility to asking a small set of abortion questions.

Fertility intentions are subject to substantial variability over time [40]. As part of a multi-country research protocol developed by the "STEP-UP" consortium that the LSHTM is co-leading with the Population Council [33], a set of new questions were developed to improve the measurement of ambivalence. We will test these questions, and because the study is nested in the HDSS sites, we will be able to prospectively assess the link between un-intendedness and pregnancy outcomes. The information collected will support a rigorous assessment of reasons for unmet need for family planning, as well as to assess whether unwanted childbearing is linked to negative pregnancy outcomes.

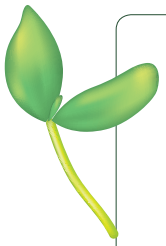
Registering a child's birth is a critical first step to protect the rights of every child, and non-registration might prevent the child from accessing health and education services. UNICEF estimates suggest that more than 230 million children under the age of 5 have not had their birth registered [41]. Household surveys represent the largest source of data on birth registration in low- and middle-income countries, [41]. Yet, registration and notification procedures vary, so more research into context specific survey questions is required [42]. In consultation with experts in the Child Protection team in UNICEF HQ, we selected birth registration questions to be evaluated in this study.

Whilst this study has strengths in terms of being randomised for the primary objective, as well as being multi-site and large-scale, pregnancies resulting in a stillbirth are less likely to be captured than those resulting in a live birth, especially if there is no frequent surveillance to capture new pregnancies and live births, even within relatively robust pregnancy surveillance systems. Hence, one limitation of this method is that women with only a stillbirth in the last five years and no live birth would not be surveyed, potentially underestimating the true SBR. In addition, the sites selected for this study have slightly different surveillance systems. Since we are collecting the data in the context of the demographic surveillance system, respondents might already be familiar with these survey questions and may be more likely than women in other settings to report pregnancy losses. If this is the case, we might have a higher reporting of events than in other settings. However, if respondents in the community are indeed more likely to report such events, this should affect women in the HDSS site equally and should not affect the randomised experiment.

The choice of Survey Solutions as our data collection tool might affect comparability with other studies using PAPI or CAPI using the CSPro Windows tablet interface. The main difference between these methods of data collection is that both PAPI and CAPI using the CSPro Windows tablet interface have a roster for the data collection of reproductive histories. In order to minimise such effect, we developed a summary of the reproductive history to mimic the reproductive history roster used in the original PAPI or CAPI using the CSPro Windows tablet application. In addition, another strength of this study is that our customised survey questionnaire developed with Survey Solutions Designer allows the inclusion of interviewer's instructions and these are visible to the interview for each question as the interviewer progresses during the interviewer visit. This addition might improve data quality.

CONCLUSIONS

Most of the 5.5 million deaths around the time of birth [1] occur in countries with the least data. Whilst improvement in CRVS and routine facility data systems is crucial, in the meantime the poorest countries rely on household surveys, and equity considerations should drive investment to improve the quality of data capture, especially for the large burden of pregnancy outcomes. We anticipate that the results of this study will inform improved tools and how these tools are applied and enable better measurement of the often-hidden outcomes associated with stigma and suffering of women in many countries. Better data alone will not change these outcomes, but counting and visibility is a crucial first step to change.



Acknowledgements: We would like to thank the Expert Advisory Group members for their inputs (alphabetical order): Fred Arnold; Peter Byass; Trevor Croft; Attila Hancioglu; Stéphane Helleringer; Kobus Herbst; Shane M Khan; Sunita Kishor; Florina Serbanescu; Turgay Unalan. We also acknowledge Michael Wild and the World Bank Survey Solutions Development Team for their technical support with the survey setup. We also would like to acknowledge the technical input provided by Claudia Cappa from UNICEF for the design of the birth registration questions. Appreciation to Claudia DaSilva, Fion Hay, Jacob Saah and Samuelina Arthur for their administrative support. We thank the women in the survey and the data collection staff from the HDSS sites at Bandim, Dabat, IgangaMayuge, Kintampo, and Matlab.

Ethics and consent to participate: This study was granted ethical approval by the institutional review boards in all operating countries as well as from Institutional Ethical Review Committee of the London School of Hygiene & Tropical Medicine (Appendix S2 in **Online Supplementary Document**).

Availability of data and material: Data sharing and transfer agreements have been made and signed by all collaborating partners.

Disclaimer: The information and views set out in this publication are those of the author(s) and do not necessarily reflect the views or the official opinion of the associated employer(s). The latter may not be held responsible for the use that may be made of the information contained therein.

Funding: This work is funded by the Children's Investment Fund Foundation (CIFF) by means of a grant to LSHTM (PI JEL), and a sub-award to the INDEPTH MNCH working group managed in Accra and with technical leadership by MakSPH (PI PW). A competitive RFA was issued to all 53 INDEPTH HDSS sites and the top five applicants were selected (as described above). SK is employed by ICF working on DHS programme, which is funded by USAID.

Authorship contributions: The study was conceptualised by JEL in 2014, and the initial protocol was developed with JEL, PW and HB and discussed in workshops in Kampala in 2015 and then refined in June 2016 with the five selected INDEPTH sites. AB joined in April 2016 and led the development of the data collection system, coordinated the research implementation activities and led the multi-country research implementation workshop that was held in Bangladesh in April 2017. AB with HB and JA and DK led the development of the questionnaire. VSG joined in mid-2016, coded the questionnaire into the software application and led the setup of multi-country data collection, data management, and monitoring systems. The data collection system was initially piloted in Bandim site with particular insights from ABF and SMT. All the site teams attended the design and implementation workshops and input to the concepts, data collection tools and training materials. DK led the development of the training materials and with HB is leading the work on barriers and enablers to data collection. The manuscript was drafted by AB and JEL, with inputs from VSG, HB, JA, and DK. All the authors reviewed and helped to revise the manuscript.

Competing interests: The authors completed the Unified Competing Interest form at www.icjme.org/coi_disclosure.pdf (available upon request from the corresponding author), and declare no conflicts of interest.

Additional material
Online Supplementary Document

REFERENCES

1 United Nations Inter-agency Group for Child Mortality Estimation (UN IGME). Levels & Trends in Child Mortality: Report 2017, Estimates Developed by the UN Inter-agency Group for Child Mortality Estimation. New York: United Nations Children's Fund; 2017.

2 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:e98-108. Medline:26795602 doi:10.1016/S2214-109X(15)00275-2

3 World Health Organization. Trends in maternal mortality: 1990 to 2015: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. Geneva: World Health Organization; 2015.

4 Lawn JE, Blencowe H, Waiswa P, Amouzou A, Mathers C, Hogan D, et al. Stillbirths: rates, risk factors, and acceleration towards 2030. Lancet. 2016;387:587-603. Medline:26794078 doi:10.1016/S0140-6736(15)00837-5

5 World Health Organization, United Nations Children's Fund. Every Newborn: an action plan to end preventable deaths. Geneva: World Health Organization; 2014.

6 United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development. New York: UN Publishing; 2015.

7 United Nations. The Global Strategy for Women's, Children's and Adolescents' Health (2016-2030). New York: United Nations; 2015.

8 Day TL, Ruysen H, Gordeev VS, Gore-Langton GR, Boggs D, Cousens S, et al. "Every Newborn Birth Indicator Research Tracking in Hospitals" (EN-BIRTH) Observational Study Protocol in Tanzania, Bangkadesh and Nepal: Validating indicators for coverage and quality of maternal and newborn health care. J Glob Health. 2019;9:010902. doi:10.7189/jogh.09.010901

9 Moxon SG, Ruysen H, Kerber KJ, Amouzou A, Fournier S, Grove J, et al. Count every newborn; a measurement improvement roadmap for coverage data. BMC Pregnancy Childbirth. 2015;15 Suppl 2:S8. Medline:26391444 doi:10.1186/1471-2393-15-S2-S8

10 World Health Organization. WHO technical consultation on newborn health indicators: every newborn action plan metrics. Ferney Voltaire, France: 3-5 December 2014. Geneva: World Health Organization; 2015.

REFERENCES

11 Lawn JE, Blencowe H, Pattinson R, Cousens S, Kumar R, Ibiebele I, et al. Stillbirths: Where? When? Why? How to make the data count? Lancet. 2011;377:1448-63. Medline:21496911 doi:10.1016/S0140-6736(10)62187-3

12 Lawn JE, Blencowe H, Oza S, You D, Lee AC, Waiswa P, et al. Every Newborn: progress, priorities, and potential beyond survival. Lancet. 2014;384:189-205. Medline:24853593 doi:10.1016/S0140-6736(14)60496-7

13 United Nations Children's Fund. Statistics and Monitoring: Multiple Indicator Cluster Surveys. Available: https://www.unicef.org/statistics/index_24302.html. Accessed: 1/2/2018.

14 International ICF. Demographic and Health Survey. Demographic and Health Surveys. Available: http://www.dhsprogram.com/. Accessed: 1 February 2018.

15 Blanc AK, Wardlaw T. Monitoring low birth weight: an evaluation of international estimates and an updated estimation procedure. Bull World Health Organ. 2005;83:178-85. Medline:15798841

16 Boerma JT, Weinstein KI, Rutstein SO, Sommerfelt AE. Data on birth weight in developing countries: can surveys help? Bull World Health Organ. 1996;74:209-16. Medline:8706237

17 Robles A, Goldman N. Can accurate data on birthweight be obtained from health interview surveys? Int J Epidemiol. 1999;28:925-31. Medline:10597993 doi:10.1093/ije/28.5.925

18 Sedgh G, Filippi V, Owolabi O, Singh SD, Ashew I, Bankole A, et al. Insights from an expert group meeting on the definition and measurement of unsafe abortion. Int J Gynaecol Obstet. 2016;134:104-6. Medline:27062249 doi:10.1016/j.ijgo.2015.11.017

19 Ghana Statistical Service (GSS), Ghana Health Service (GHS), and Macro International. Ghana Maternal Health Survey 2007. Calverton, Maryland, USA: GSS, GHS, and Macro International; 2009.

20 Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF Macro. Ghana Demographic and Health Survey 2008. Accra, Ghana: GSS, GHS, and ICF Macro; 2009.

21 Stanton C. Perinatal Mortality in the Philippine: An investigation into the Use of Demographic Survey Data for the Study of Perinatal Mortality and its Determinants. PhD Thesis. Baltimore, Maryland: School of Hygiene and Public Health; 1996.

22 Espeut D, Becker S. The validity of birth and pregnancy histories in rural Bangladesh. J Health Popul Nutr. 2015;33:17. Medline:26825676 doi:10.1186/s41043-015-0027-8

23 Haws RA, Mashasi I, Mrisho M, Schellenberg J, 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;71:1764-72. Medline:20541305 doi:10.1016/j.socscimed.2010.03.051

24 Johnson K, Grant M, Khan S, Moore Z, Armstrong A, Sa Z. Fieldwork-related factors and data quality in the Demographic and Health Surveys program. DHS Analytical Studies No. 19. Calverton, Maryland, USA: ICF Macro; 2009.

25 Biswas A, Rahman F, Eriksson C, Halim A, Dalal K. Social Autopsy of maternal neonatal deaths and stillbirths in rural Bangladesh: qualitative exploration of its effect and community acceptance. BMJ Open. 2016;6:e010490. Medline:27554100 doi:10.1136/bmjopen-2015-010490

26 Boyden JY, Kavanaugh K, Issel M, Eldeirawi K, Meert K. Experiences of African American parents follwoing perinatal or pediatric death: a litterature review. Death Stud. 2014;38:374-80. Medline:24666143 doi:10.1080/07481187.2013.766656

27 Ellis A, Chebsey C, Storey C, Bradley S, Jackson S, Flenady V, et al. Systematic review to understand and improve care after stillbirth: a review of parents' and healthcare professional experinces. BMC Pregnancy Childbirth. 2016;16:16. Medline:26810220 doi:10.1186/s12884-016-0806-2

28 EN-INDEPTH Technical Secretariat. ENAP & INDEPTH Research Protocol Design Workshop. Kampala, Uganda. 15-17 June2016.

29 Kadobera D, Waiswa P, Peterson S, Blencowe H, Lawn J, Kerber K, et al. Comparing performance of methods used to identify pregnant women in the Iganga-Mayuge Health and Demographic Suirveillance Site, Uganda. Glob Health Action. 2017;10:1356641. Medline:28799450 doi:10.1080/16549716.2017.1356641

30 INDEPTH; Maternal Newborn Health Working Group. 2010. Available: http://www.indepth-network.org/groups/working-groups/maternal-and-newborn-health. Accessed: 1/2/2018.

31 INDEPTH. INDEPTH Network. 2016. Available: http://www.indepth-network.org/about-us. Accessed: 1/2/2018.

32 The World Bank. Survey Solutions CAPI/CAWI platform: Release 5.26. Accessed: 1 February 2018.

33 Machiyama K, Casterline J, Mumah J, Huda F, Obare F, Kabiru C, et al. Reasons for unmet need for family planning, with attention to the measurement of fertility preferences: protocol for a multi-site cohort study. Reprod Health. 2017;14:23. Medline:28183308 doi:10.1186/s12978-016-0268-z

34 ICF. Demographic and Health SurveyInterviewer's Manual. Rockville, Maryland: ICF; 2017.

35 The World Bank. Survey Solutions Manuals. 2018. Available: http://support.mysurvey.solutions. Accessed: 1/2/2018.

36 Liu L, Kalter H, Chu Y, Kazmi N, Koffi A, Amouzou A, et al. Understanding Misclassification between Neonatal Deaths and Stillbirths: Empirical Evidence from Malawi. PLoS One. 2016;11:e0168743. Medline:28030594 doi:10.1371/journal.pone.0168743

37 Hill K, Choi Y. Neonatal mortality in the developing world. Demogr Res. 2006;14:429-52. doi:10.4054/DemRes.2006.14.18

38 Channon AA, Padmadas SS, McDonald JW. Measuring birth weight in developing countries: does the method of reporting in retrospective surveys matter? Matern Child Health J. 2011;15:12-8. Medline:20063179 doi:10.1007/s10995-009-0553-3

39 Johnson BR, Mishra V, Lavelanet F, Khosla R, Ganatra B. A global database of abortion laws, policies, health standards and guidelines. Bull World Health Organ. 2017;95:542-4. Medline:28670021 doi:10.2471/BLT.17.197442

40 Machiyama K, Baschieri A, Dube A, Crampin M, Glynn J, French N, et al. An assessment of Childbearing Preferences in Northen Malawi. Stud Fam Plann. 2015;46:161-76. Medline:26059988 doi:10.1111/j.1728-4465.2015.00022.x

41 United Nations Children's Fund. Every Child's Birth Right: Inequities and trends in birth registration. New York: UNICEF; 2013.

42 United Nations Children's Fund. A Passport to Protection. A guide to birth registration programming. New York: UNICEF; 2013.

Chapter 5: Randomised comparison of the Full Birth History with additional questions on pregnancy losses and the Full Pregnancy History

5.1. Introduction

This chapter provides a randomised comparison of the reproductive module used in the latest version of DHS-7 "full birth history" versus a "full pregnancy history" module for estimation of stillbirth rate and neonatal mortality rate (Objective 2). It also examines the variation in response time to completion for questions in the FBH+ and FPH survey modules (Objective 3) This chapter was published April 2020 in the Lancet Global Health[41]. The manuscript was published under a Creative Commons Attribution 4.0 International License (CC BY 4.0) - http://creativecommons.org/licenses/by/4.0/ and the published manuscript is included in full below. The web appendix referenced in the paper is available at https://www.sciencedirect.com/science/article/pii/S2214109X20300449#ecomp10

5.2. List of Figures

Figure 1 – Five Health and Demographic Surveillance System sites with summary statistics from the Every Newborn-INDEPTH study

Figure 2 – Every Newborn-INDEPTH study profile

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

Figure 4 – Time taken to complete questions for the two survey modules (n=65 511) in the Every Newborn-INDEPTH study

5.3. List of Tables

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

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

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

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

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)

5.4. Citation

Akuze Joseph, Blencowe Hannah, Waiswa Peter, Baschieri Angela, Gordeev Vladimir S, Kwesiga Doris, Fisker Ane B, Thysen Sanne M, Rodrigues Amabelia, Biks Gashaw A, Abebe Solomon M, Gelaye Kassahun A, Mengistu Mezgebu Y, Geremew Bisrat M, Delele Tadesse G, Tesega Adane K, Yitayew Temesgen A, Kasasa Simon, Galiwango Edward, Natukwatsa Davis, Kajungu Dan, Enuameh Yeetey AK, Nettey Obed E, Dzabeng Francis, Amenga-Etego Seeba, Newton Sam K, Tawiah Charlotte, Asante Kwaku P, Owusu-Agyei Seth, Alam Nurul, Haider M M, Imam Ali, Mahmud Kaiser, Cousens Simon, Lawn Joy E and the Every Newborn-INDEPTH Study Collaborative Group. **Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in five countries: the Every Newborn-INDEPTH study**. The Lancet Global Health, 2020. 8(4): p. e555-e566.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1512635	**Title**	Mr
First Name(s)	Joseph		
Surname/Family Name	Akuze Waiswa		
Thesis Title	Measurement of stillbirths and neonatal deaths in standardised population-based surveys		
Primary Supervisor	Hannah Blencowe		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Lancet Global Health as: Joseph Akuze, Hannah Blencowe, Peter Waiswa, Angela Baschieri, Vladimir S Gordeev, Doris Kwesiga, et al. (2020) Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in 69,176 pregnancies in five Health and Demographic Surveillance sites: Every Newborn-INDEPTH study. Lancet Global Health 2020-04. DOI: 10.1016/s2214-109x(20)30044-9. PMID: 32199123		
When was the work published?	April 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	Not Applicable		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Not Applicable
Please list the paper's authors in the intended authorship order:	Not Applicable
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I developed the detailed research questions jointly with Drs. Hannah Blencowe, Angela Baschieri, Vladimir Gordeev, and Professors. Peter Waiswa, Simon Cousens and Joy Lawn. I conducted the literature reviews and with guidance from Prof. Simon Cousens and Dr. Hannah Blencowe, I developed the detailed analysis plan and led all the data management, development of statistical code and undertook all analyses. I drafted the initial version of the paper and coordinated the revisions and submission of this manuscript until it was published in the Lancet Global Health. All co-authors reviewed and gave approvals of the final manuscript

SECTION E

Student Signature	Joseph Akuze Waiswa
Date	27th September 2021

Supervisor Signature	Hannah Blencowe
Date	28th September 2021

5.5. Research paper

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

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.

Copyright © 2020 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

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,

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

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.

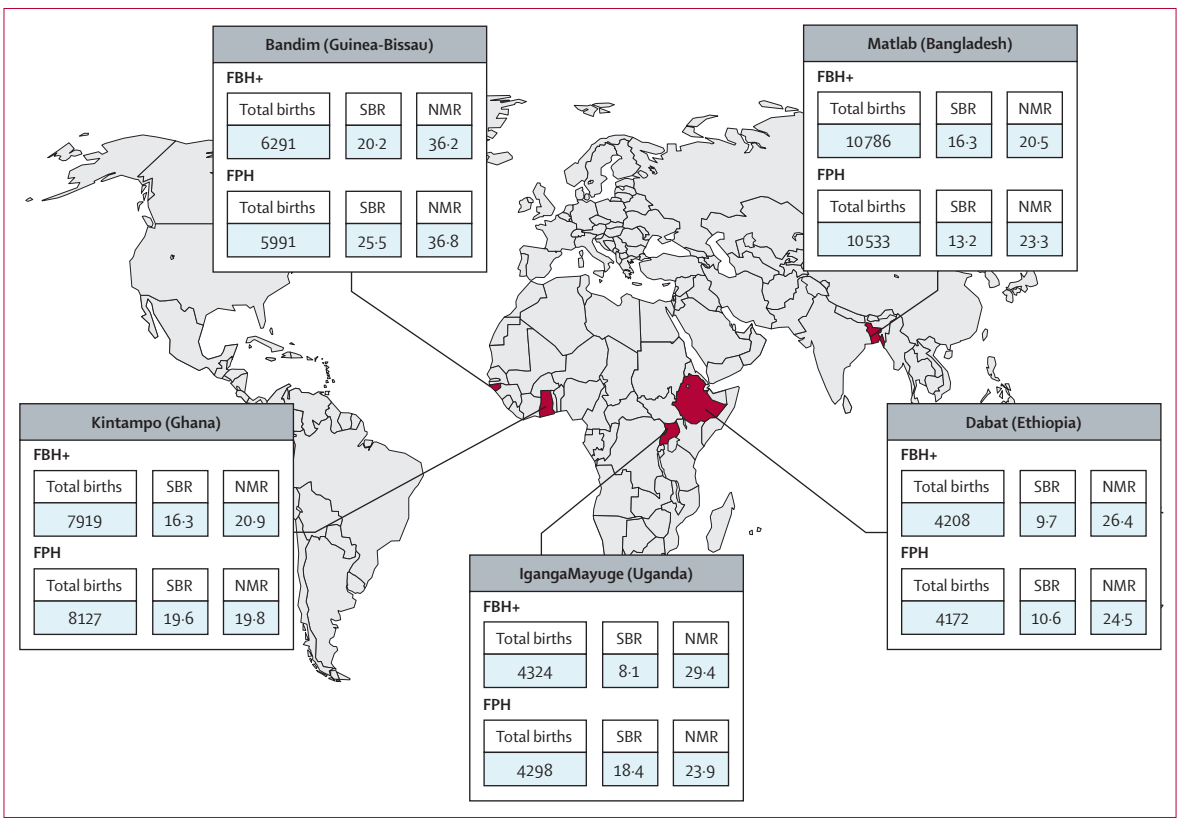


Figure 1: **Five Health and Demographic Surveillance System sites with summary statistics from the Every Newborn-INDEPTH study**
FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. NMR=neonatal mortality rate. SBR=stillbirth rate.

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

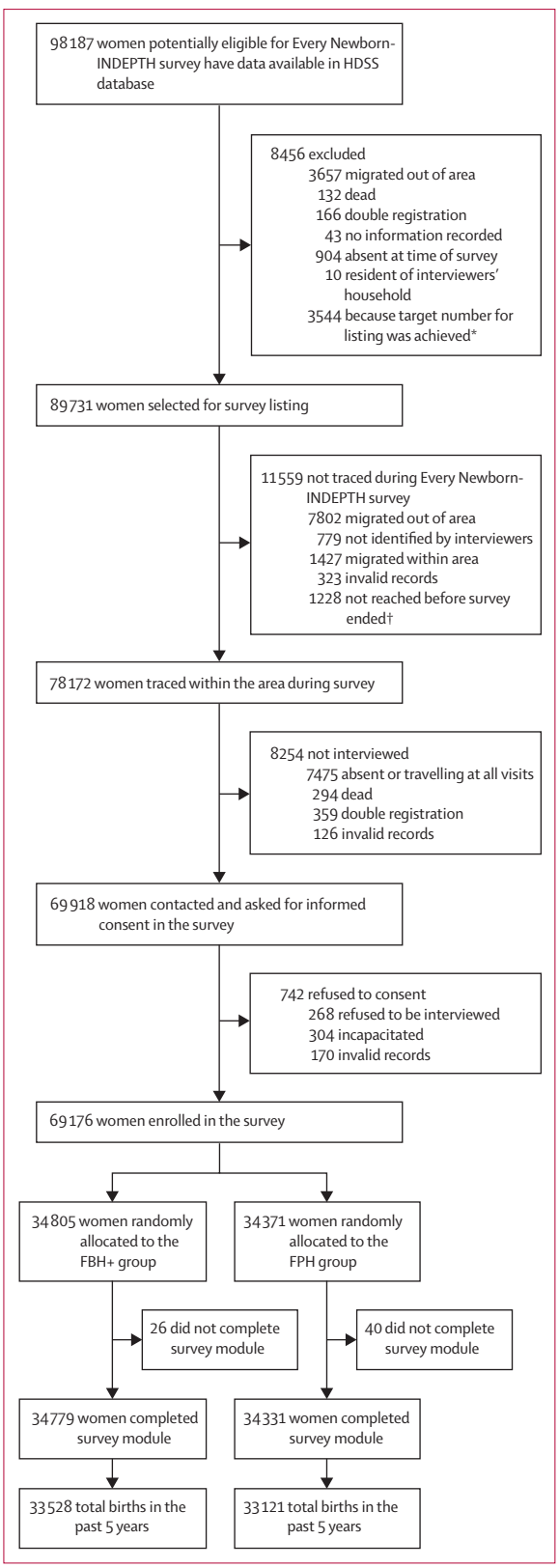


Figure 2: **Every Newborn-INDEPTH study profile**
FBH+=full birth history module with additional questions on pregnancy losses. FPH=full pregnancy history module. HDSS=Health and Demographic Surveillance System. *In IgangaMayuge, when creating the Every Newborn-INDEPTH study listing, some women were 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**

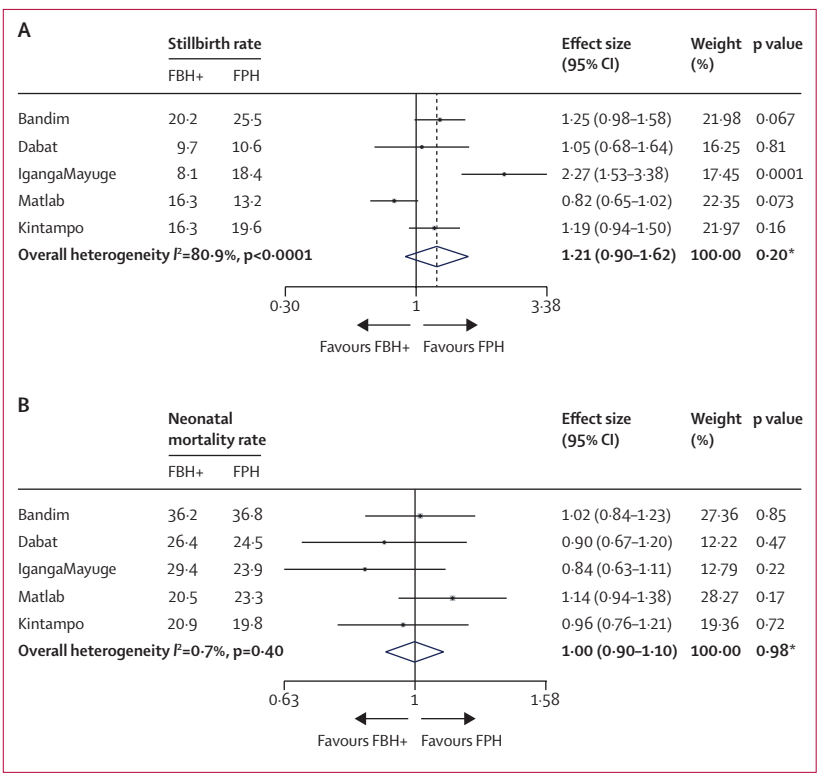


Figure 3: **Forest plot comparing stillbirth (A) and neonatal mortality (B) between FBH+ and FPH modules, by HDSS site**

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)**

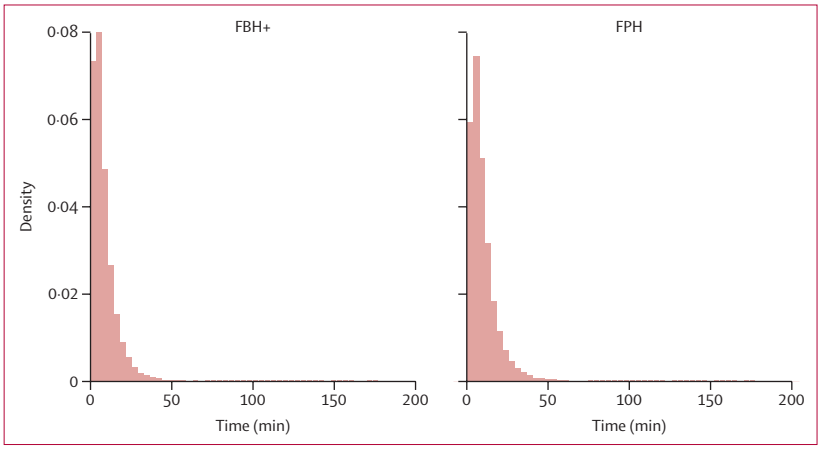


Figure 4: **Time taken to complete questions for the two survey modules (n=65 511) in the Every Newborn-INDEPTH study**

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.

Chapter 6: Further assessments of effects of time taken for completion of FBH+ and FPH survey modules and data quality using Paradata

6.1. Introduction

This chapter further examines the effects of time taken for completion of FBH+ and FPH survey modules (Objective 3). It also examines the data quality aspects of the two modules focusing on data missingness and correction patterns and makes suggestions for future research regarding improving stillbirth and neonatal deaths data quality.

In this chapter, my major contributions were 1) providing data management for the Paradata, 2) contributions to the conceptualisation of the manuscript and development of the research questions that were asked to answer the additional gaps in knowledge about data quality, corrections, correction types and how it affects the overall timing and implementation of the survey modules. I also produced some of the graphics in the manuscript.

In additional to the analyses done on this chapter, I am conducting further research to examine 1) the use of paradata and metadata for improving response rate, 2) relationship between correction rates and correction types and being supervised, 3) individual and team productivity, 4) estimating contact attempts without success and response propensity, and 5) the impact of workload for countries with additional questions on timing and corrections.

This chapter was published February 2021 in BMC Population Health Metrics.[40] The manuscript was published under a Creative Commons Attribution 4.0 International License (CC BY 4.0) - http://creativecommons.org/licenses/by/4.0/ and the published manuscript is included in full below. The web appendix referenced in the paper is available at https://pophealthmetrics.biomedcentral.com/articles/10.1186/s12963-020-00241-0

6.2. List of Figures

Figure 1 – Data collection cycle showing survey and paradata: EN-INDEPTH survey

Figure 2 – Module structure for questions in the two arms of the EN-INDEPTH survey

Figure 3 – Time to complete questions regarding maternity history (section 2) for two survey modules (N = 60,871)

Figure 4 – Question corrections by question type, nature and structure

Figure 5 – Data collection and correction as a sequence index plot by length of interview

6.3. List of Tables

Table 1 – Interview process details: number of timestamped entries, response time and corrections

Table 2 – Question characteristics associated with answer correction probabilities and frequencies

Table 3 – Classification of answer correction type and possible explanations

Table 4 – Answer correction types per question type, nature, and structure

6.4. Citation

Vladimir Sergeevich Gordeev, Joseph Akuze, Angela Baschieri, Sanne M. Thysen, Francis Dzabeng, M. Moinuddin Haider, Melanie Smuk, Michael Wild, Michael M. Lokshin, Temesgen Azemeraw Yitayew, Solomon Mokonnen Abebe, Davis Natukwatsa, Collins Gyezaho, Seeba Amenga-Etego, Joy E. Lawn, Hannah Blencowe, the Every Newborn-INDEPTH Study Collaborative Group. **Paradata analyses to inform population-based survey capture of pregnancy outcomes: EN-INDEPTH study**. BMC Population Health Metrics https://doi.org/10.1186/s12963-020-00241-0

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1512635	**Title**	Mr
First Name(s)	Joseph		
Surname/Family Name	Akuze Waiswa		
Thesis Title	Measurement of stillbirths and neonatal deaths in standardised population-based surveys		
Primary Supervisor	Hannah Blencowe		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	BMC population metrics as: Vladimir Sergeevich Gordeev, Joseph Akuze, Angela Baschieri, Sanne M. Thysen, Francis Dzabeng, M. Moinuddin Haider, Melanie Smuk, Michael Wild, Michael M. Lokshin, Temesgen Azemeraw Yitayew, Solomon Mokonnen Abebe, Davis Natukwatsa, Collins Gyezaho, Seeba Amenga-Etego, Joy E. Lawn, Hannah Blencowe, the Every Newborn-INDEPTH Study Collaborative Group. Paradata analyses to inform population-based survey capture of pregnancy outcomes: EN-INDEPTH study. London School of Hygiene and Tropical Medicine. BMC Population Health Metrics https://doi.org/10.1186/s12963-020-00241-0		
When was the work published?	08 February 2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	Not Applicable		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Not Applicable
Please list the paper's authors in the intended authorship order:	Not Applicable
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I led the overall data management of the EN-INDEPTH paradata which was analysed for this paper. Jointly with Drs. Vladimir Gordeev, and Hannah Blencowe, Professor Joy Lawn and I contributed the conceptualisation of this paper and to the development and refining of detailed research questions. I met regularly with Vladimir Gordeev and Hannah Blencowe to review the preliminary results from the data analysis. I supported Dr. Vladimir Gordeev in drafting and revising of the manuscript. I developed some of the graphics in the manuscript and supported the peer-review process providing responses to reviewers' comments on the manuscript after submission.

SECTION E

Student Signature	Joseph Akuze Waiswa
Date	27th September 2021

Supervisor Signature	Hannah Blencowe
Date	28th September 2021

6.5. Research paper

D: Paradata analyses to inform population-based survey capture of pregnancy outcomes: EN-INDEPTH study

Gordeev et al. Population Health Metrics 2021, 19(Suppl 1):10
https://doi.org/10.1186/s12963-020-00241-0

Population Health Metrics

From Every Newborn-INDEPTH study: Improving the measurement of pregnancy outcomes in population-based surveys

RESEARCH Open Access

Paradata analyses to inform population-based survey capture of pregnancy outcomes: EN-INDEPTH study

Vladimir Sergeevich Gordeev[1,2*], Joseph Akuze[2,3,4], Angela Baschieri[2], Sanne M. Thysen[5,6,7], Francis Dzabeng[8], M. Moinuddin Haider[9], Melanie Smuk[10], Michael Wild[11], Michael M. Lokshin[11], Temesgen Azemeraw Yitayew[12], Solomon Mokonnen Abebe[12], Davis Natukwatsa[13], Collins Gyezaho[13], Seeba Amenga-Etego[8], Joy E. Lawn[2†], Hannah Blencowe[2†] and the Every Newborn-INDEPTH Study Collaborative Group

Abstract

Background: Paradata are (timestamped) records tracking the process of (electronic) data collection. We analysed paradata from a large household survey of questions capturing pregnancy outcomes to assess performance (timing and correction processes). We examined how paradata can be used to inform and improve questionnaire design and survey implementation in nationally representative household surveys, the major source for maternal and newborn health data worldwide.

Methods: The EN-INDEPTH cross-sectional population-based survey of women of reproductive age in five Health and Demographic Surveillance System sites (in Bangladesh, Guinea-Bissau, Ethiopia, Ghana, and Uganda) randomly compared two modules to capture pregnancy outcomes: full pregnancy history (FPH) and the standard DHS-7 full birth history (FBH+). We used paradata related to answers recorded on tablets using the Survey Solutions platform. We evaluated the difference in paradata entries between the two reproductive modules and assessed which question characteristics (type, nature, structure) affect answer correction rates, using regression analyses. We also proposed and tested a new classification of answer correction types.

(Continued on next page)

* Correspondence: v.gordeev@qmul.ac.uk; vladimir.gordeev@lshtm.ac.uk
†Joy E. Lawn and Hannah Blencowe are joint senior authors.
[1]Institute of Population Health Sciences, Queen Mary University of London, London, UK
[2]Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene & Tropical Medicine, London, UK
Full list of author information is available at the end of the article

© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Results: We analysed 3.6 million timestamped entries from 65,768 interviews. 83.7% of all interviews had at least one corrected answer to a question. Of 3.3 million analysed questions, 7.5% had at least one correction. Among corrected questions, the median number of corrections was one, regardless of question characteristics. We classified answer corrections into eight types (no correction, impulsive, flat (simple), zigzag, flat zigzag, missing after correction, missing after flat (zigzag) correction, missing/incomplete). 84.6% of all corrections were judged not to be problematic with a flat (simple) mistake correction. Question characteristics were important predictors of probability to make answer corrections, even after adjusting for respondent's characteristics and location, with interviewer clustering accounted as a fixed effect. Answer correction patterns and types were similar between FPH and FBH+, as well as the overall response duration. Avoiding corrections has the potential to reduce interview duration and reproductive module completion by 0.4 min.

Conclusions: The use of questionnaire paradata has the potential to improve measurement and the resultant quality of electronic data. Identifying sections or specific questions with multiple corrections sheds light on typically hidden challenges in the survey's content, process, and administration, allowing for earlier real-time intervention (e.g.,, questionnaire content revision or additional staff training). Given the size and complexity of paradata, additional time, data management, and programming skills are required to realise its potential.

Keywords: Survey, Paradata, Neonatal, Newborn, Answer correction type, Survey design

Key findings

What is new?

- **What was known already:** Paradata are widely used in the field of survey methodology in high-income countries to monitor on-going fieldwork progress and identify issues with specific questions but have been little-used to date in low- and middle-income countries and for maternal, newborn, and child health data collection or research.
- **What was done:** We analysed paradata from the EN-INDEPTH survey administered to 65,768 women of reproductive age in five countries. We assessed which question characteristics used to capture pregnancy outcomes affected duration of section completion, data correction rates, or were associated with multiple corrections and whether these differed by two maternity history modules (full pregnancy history (FPH) and full birth history (FBH+)).

What was found?

- **Corrections to questions were common:** affecting 83.7% of interviews, with a median of two questions corrected per interview and one correction per question when corrected (maximum of 28 corrections). 7.5% of the 3.3 million questions analysed had at least one correction.
- **Simple one-time corrections most common:** accounting for 84.6% of all corrections.
- **In variation in corrections by maternity history module:** number and type of corrections were similar between FPH and FBH+.
- **In variation in corrections by question characteristics:** number and type of corrections were affected by question characteristics. The proportion of corrected questions was 3.3% higher for questions with notifications (9.8%) than for questions without notifications (6.5%).
- **Duration of question completion:** was not affected by question characteristics (type, content, structure) or history type. Avoiding corrections has the potential to reduce interview duration and reproductive module completion by 0.4 min.

What next in measurement and research?

- **Measurement improvement now:** Paradata can be used to identify questions with multiple corrections, informing question editing or targeted training during and after survey completion. Encoding ranges and instant error notifications in the reproductive modules could reduce data missingness and prompt for timely data correction. Paradata analyses are skill- and time-consuming, but, if automatised, can be used for real-time data collection monitoring and data quality control.

Key findings *(Continued)*

- **Research needed:** Studies could examine interviewer productivity and possible fatigue related to the length of the interview, the number of corrections, and correction types. The real-time dashboard monitoring and reporting systems using paradata could be evaluated in terms of associations with data quality and usefulness for survey management. Qualitative interviews with both respondents and interviewers would help to identify and verify factors affecting correction frequency to inform better questionnaire design and training adjustment.

Background

High-quality routine health data on maternal, newborn, and child health (MNCH) can be used to monitor, identify gaps, and take action to improve quality of care, optimise health system performance, and enable informed decision-making. Routine health management information systems vary in their completeness and quality across low- and middle-income countries (LMICs), and in many cases, are not able to provide the high-quality coverage data required for assessing and guiding health programmes [1, 2]. Household surveys, notably Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS), remain the primary sources of data for the outcome and coverage indicators for children and women for most low- and middle-income countries. However, despite existing quality control mechanisms in the survey process, data quality, including missingness, age displacement, and heaping, remains a challenge [3]. Optimising survey data efficiency and quality requires more information regarding the survey process and performance [4].

The shift from paper-based to computer-assisted personal interviewing (CAPI)-based data collection (e.g., using tablets and smartphones) has enabled inclusions of inbuilt validation and consistency checks, as well as a real-time

review of collected data [5]. In addition to the main survey dataset (which contains only the final respondent's answers), it is also possible to collect the survey's paradata (Fig. 1). Paradata contain information on the process of how data for each observation in the main survey dataset was collected and include detailed timestamped records of all survey actions including survey administration, interview process, as well as a detailed history of all the survey's data entry and correction [6, 7]. For example, paradata can show the order in which the questions were answered or corrected and reveal the content of deleted responses, which otherwise are not stored in the main survey dataset.

Even though paradata are widely used in the field of survey methodology in high-income countries [8], the use to date in LMICs and MNCH research has been limited. In household surveys, timestamped paradata can be used to monitor ongoing fieldwork progress and identify issues with specific questions or questionnaire sections based on the length of interviews or item response time [8]. Additional analyses can identify drivers behind item non-response and response time (e.g., survey locale; interviewer or respondent characteristics; survey's content—questions' type, nature, and structure) [9]. Paradata can also reveal determinants of data correction in relevant core survey questions during interviews as well as answer correction patterns (even though there is currently no agreed standardised terminology). Hence, paradata could lead to the overall improvements in data quality through targeted training [8] as well as improving questionnaire and survey design (structure and content) and survey implementation (process).

In this paper, we examine how paradata can be used to inform and improve questionnaire design and survey implementation in a large household survey collecting information on pregnancies and births using full pregnancy histories (FPH) and full birth histories with additional questions on pregnancy losses in the past 5 years (FBH+). This paper is one of a series of papers from the Every Newborn International Network for the Demographic Evaluation of Populations and Their Health (EN-INDEPTH) study in five Health and Demographic Surveillance System (HDSS) sites in sub-Saharan Africa and Asia.

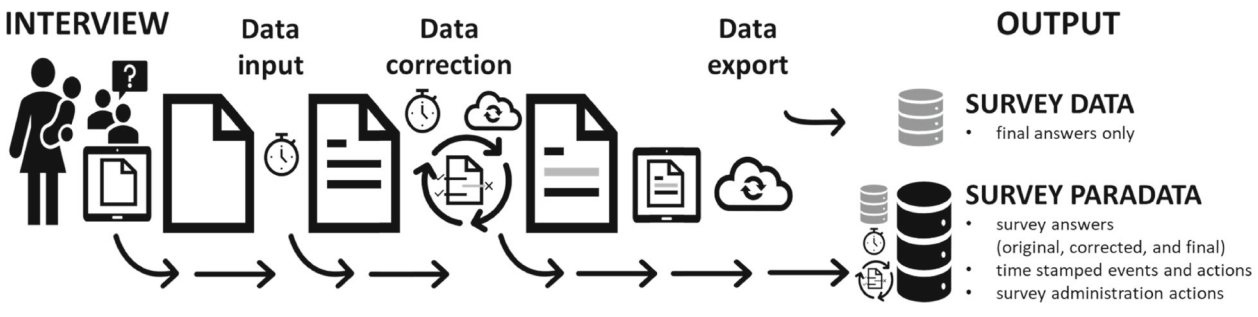
This paper has three objectives:

1. To assess the differences in paradata timestamped entries between two reproductive modules (FPH and FBH+);
2. To determine whether question characteristics (type, nature, structure) affect the duration of section completion and answer correction rates;
3. To propose and test classification of answer correction types and determine whether they differ by two reproductive modules.

Methods

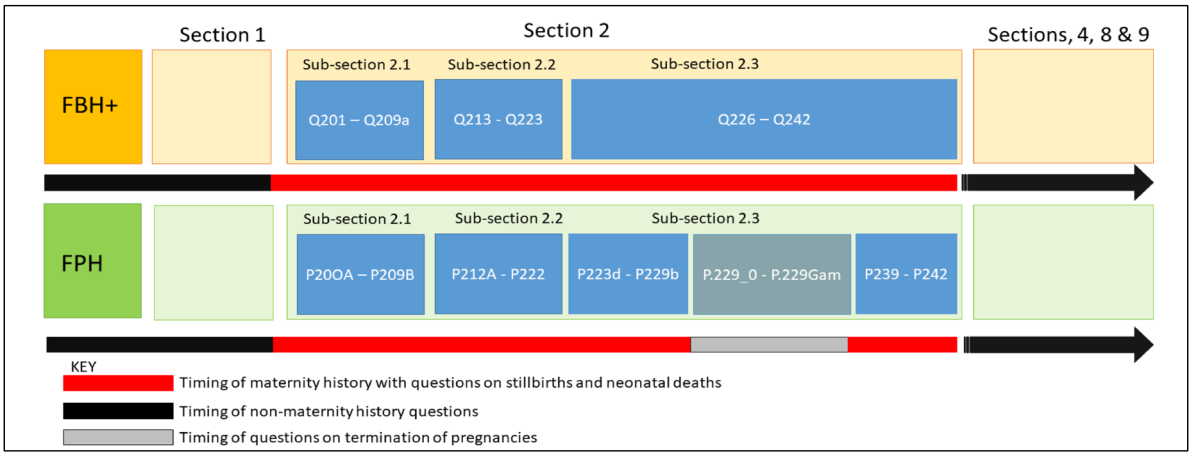
Overall EN-INDEPTH study design and data sources

The EN-INDEPTH study aimed to compare two approaches of collecting maternity history (FPH and FBH+) to examine whether the two methods yield different estimates of stillbirth rates and neonatal mortality rates and to determine whether there is a difference in completion time for these two approaches. The study protocol and main findings can be found elsewhere [10, 11]. Briefly, the EN-INDEPTH survey reached 69,176 women of reproductive age in five HDSS sites (Bandim in Guinea-Bissau, Dabat in Ethiopia, IgangaMayuge in Uganda, Matlab in Bangladesh, and Kintampo in Ghana). Participants of the EN-INDEPTH study were randomly assigned (1:1) to be interviewed using a questionnaire containing either an FPH or an FBH+ module (section 2 in Fig. 2). The EN-INDEPTH study used the World Bank's Survey Solutions CAPI/CAWI (computer-assisted web interviewing) data collection and management platform (hereafter Survey Solutions) [12] to collect face-to-face responses to the questionnaire (Additional file 1). The choice of the



Notes: Paradata dataset contains detailed timestamped records of all survey actions. It can store data regarding actions and timing undertaken for the survey administration (assignment and reassignment of an interview, approval or rejection of answers by supervisors, switching of the language (translation) of the interview), interview process (interview started, paused, resumed, restarted, completed), as well as a detailed history of survey's data entry and correction.

Fig. 1 Data collection cycle showing survey and paradata: EN-INDEPTH survey



Notes: FPH - full pregnancy history module. FBH+ - full birth history module with additional questions on pregnancy losses. Section 1 – Consent form and background of an interview; Section 2 – Reproduction module; Section 4 – Pregnancy and postnatal care; Section 8 – Fertility preferences; Section 9 – Household characteristics.

Fig. 2 Module structure for questions in the two arms of the EN-INDEPTH survey

software and an overview of the data collection process and procedures are detailed elsewhere [13].

The analyses in this paper focus on the reproduction section (section 2) of the questionnaire only (Fig. 2). Section 2 contains three subsections. In subsection 2.1, women were asked to state their total lifetime number of liveborn children (FPH and FBH+) and a total number of pregnancy losses (FPH only). In subsection 2.2, women were asked for details about their lifetime pregnancies (FPH) or lifetime livebirths (FBH+) and answer a subset of questions for each instance. In subsection 2.3, women were asked in the FBH+ group about pregnancy losses in the past 5 years, while the FPH module contained an additional set of questions on termination of pregnancy (TOP), which we also included in the analyses (Additional file 1).

Data processing

Paradata were exported in a tab-delimited format from the Survey Solutions platform [12], with each line corresponding to one recorded event (example in Additional file 2). Data from all sites were fully anonymised and required subsets of data extracted and merged using the R software [14]. We included only timestamped entries related to answers and corrections from section 2. We excluded entries that related to assigned but never conducted interviews and duplicate entries that resulted from updating questionnaire and software. All subsequent analyses were performed using Stata 16.1 [15].

Methods by objective

Objective 1: To assess differences in paradata timestamped entries between two reproductive modules (FPH and FBH+)

Time taken to complete a question was estimated based on the difference between the timestamps of the previously answered question and the current one (based on answered questions order) and separate per observation (in case of parity > 1). For questions with corrections, the timestamp of the final answer was taken as the question's timestamp. To exclude implausible values, but allow for multiple corrections during the module completion and/or switching between questionnaire sections during the interview [16], we restricted analyses of time taken to complete maternity history section of the survey to interviews lasting 0–180 min.

We categorised all questions by type, nature, and structure. The question types included single-select (e.g., 'Was that baby a twin?'); multi-select (e.g., 'Who assisted with the delivery of this baby?'); numerical computational (e.g., 'How many children do you have?'); date-related (e.g., 'What was the date of birth for this baby?'); and free-text input (e.g., 'What is the name of your baby?'). There were three categories based on questions' nature: two groups of potentially sensitive questions (death-related—relating to death and/or pregnancy loss, and TOP-related questions) and regular (non-sensitive) questions. Lastly, considering question structure, there were questions with built-in error notifications (e.g., displaying "value outside the range, please correct"), warnings appearing in capital red letters, or any other prompts for correction (e.g., when answers for age based on birth and age at last birthday did not match) and those without such notifications. Differences between FPH and FBH+ were evaluated using descriptive statistics and independent sample t test. Statistical significance level was defined at the 5% level.

Objective 2: To determine whether question characteristics affect the duration of section completion and answer correction rates

Differences in the duration of response time and proportion of corrections by question characteristics

(type, nature, structure) were evaluated using descriptive statistics. A two-part model was used to analyse which question characteristics are associated with the likelihood of question correction (generalised linear model (GLM) with a binomial distribution and logit link) and the number of corrections (GLM with gamma distribution and log link function). Explanatory variables included question characteristics. Models were adjusted for respondent's characteristics and location, with interviewer clustering accounted as a fixed effect. Statistical significance level was defined at the 5% level.

Objective 3: To propose and test classification of answer correction patterns and determine whether data correction patterns differ by two reproductive modules

In paradata, the process of data collection where all answers are entered and corrected is recorded as an ordered list of answers (sequence). To understand this process of data entry and correction better, we ordered all interviews based on the total number of questions asked during the interview and the number of answers (length of a sequence). We distinguished between original answers and corrections and visually inspected the resulting sequence index plot [17].

Whenever an answer to the same question has multiple corrections, these corrections can form a distinct pattern. For example, corrections can be single or multiple; the value of the original answer and the last correction may or may not match; correction entries may have identical or different values and may lead to missing data. As currently there is no classification of answer correction types, we developed and tested one using our survey data. We then used descriptive statistics to examine whether answer correction patterns vary by question characteristics and two reproductive modules.

Results are reported in accordance with STROBE Statement checklists for cross-sectional studies [18] (Additional file 3).

Results

Objective 1: To assess differences in paradata timestamped entries between two reproductive modules (FPH and FBH+)

Number of timestamped entries

We analysed 3.6 million timestamped entries corresponding to 3.3 million answered questions and their correction for 65,768 interviews, of which 52.1% related to FPH module (32,744 interviews), which by design contained more questions than the FBH+ reproductive module (33,024 interviews) (Table 1 and Additional file 4). Among all entries, 18.5% related to the pregnancy or birth history (sub-section 2.1), 66.2% to the roster (sub-section 2.2), and 15.3% to reproduction subsections of FPH and FBH+ reproductive modules (sub-section 2.3) (Fig. 2). The median number of timestamped answers per interview was 48 (52 and 45 for FPH and FBH+, respectively).

Type, nature, and structure of questions

FPH and FBH+ modules contain 98 and 66 possible uniquely formulated question/answer fields, respectively (Additional files 1 and 4). FPH reproductive module contains 52 single-select questions, 26 numerical computational, one date-related, and 17 free-text and two multi-select types of questions. FBH+ module contains 35 single-select questions, 28 numerical computational, one date-related, and two free-text and no multi-select types of questions. FPH has 18 questions related to death/pregnancy loss and 39 questions related to TOP (including country-specific questions). FBH has 27 questions related to pregnancy loss/death. The rest of the questions are regular (non-sensitive) by nature. A quarter of questions in FPH and about roughly a third in FBH+ have built-in error notifications.

Most of the timestamped entries related to single-select questions (66.6%), followed by numerical computational (32.1%), date-related (1.2%), and less than 0.03% being free-text and multi-select types of questions (Additional file 4). The proportion of timestamped entries per question type between modules was very similar. In terms of the questions' nature, most of the timestamped entries were for regular questions (87.2%) and not related to the two groups of potentially sensitive questions (death and/or pregnancy loss, and TOP-related questions). In terms of structure, about a third of timestamped entries were for questions that had built-in error notifications, warnings, or other prompts for a correction.

The average duration of section and question completion

The median number of questions answered per one interview was 44: 49 for FPH and 41 for FBH+, as FPH contained an additional set of TOP-related questions absent in FBH+. The median duration of section 2 completion was 7.3 min (Table 1, Fig. 3). The average time taken to complete the reproduction module was 1.1 min longer for the FPH (mean = 11.4 min) than the FBH+ (10.3 min). The median response time per question was around 0.1 min overall and for both modules.

Corrections per interview and questions

Overall, 83.7% of all interviews had at least one corrected answer to a question, slightly higher for FPH module than FBH+ (84.6% and 82.8%, respectively) (Table 1). The median number of corrected questions per interview was two, and a median number of corrections was three. The median time spent on corrections per one interview was 0.3 min (the mean

Table 1 Interview process details: number of timestamped entries, response time and corrections

Indicator	Overall			FPH			FBH+			*P* value[$]
	Mean (SD)	Median	Range		Median	Range	Mean (SD)	Median	Range	
N timestamped entries per interview[a]	55.1 (33.4)	48	1–335	57.7 (34.9)	52	1–335	52.5 (31.5)	45	6–289	< 0.001
N questions answered per 1 interview	50.7 (29.6)	44	1–223	53.2 (31.1)	49	1–223	48.4 (27.8)	41	6–194	< 0.001
Response time per 1 interview[b], min	10.8 (14.3)	7.3	0.06–179.9	11.4 (14.8)	7.8	0.06–179.9	10.3 (13.3)	6.9	0.23–179.8	< 0.001
Response time per 1 question[b], min	0.4 (3.5)	0.08	0–179.9	0.4 (3.6)	0.08	0–179.9	0.4 (3.6)	0.07	0–179.9	
N corrected questions per 1 interview	3.8 (4.8)	2	0–112	3.9 (4.9)	3	0–112	3.6 (4.8)	2	0–110	< 0.001
N corrected question per 10 questions	0.8 (0.9)	0.6	0–29.5	0.8 (0.9)	0.6	0–29.5	0.8 (0.9)	0.62	0–20.4	< 0.01
N corrections per 1 interview	4.4 (6.2)	3	0–227	4.6 (6.3)	3	0–227	4.2 (6.1)	2	0–149	< 0.01
Time spent on correction per 1 interview, min	1.9 (10.2)	0.3	0–179.6	2.0 (10.2)	0.3	0–174.5	1.9 (10.2)	0.2	0–179.6	< 0.01
Time spent on correction per 1 question, min	0.6 (5.6)	0.08	0–179.9	0.6 (5.5)	0.08	0–179.8	0.6 (5.6)	0.08	0–179.9	
Response time per 1 interview, if all corrections avoided, min	9.4 (10.1)	6.9	0.06–179.5	9.9 (10.6)	7.4	0.06–179.5	8.8 (9.4)	6.5	0.2–179.4	< 0.001
Response time per 1 question, if corrections avoided, min	1.7 (4.5)	0.8	0.5–177.1	1.7 (4.5)	0.8	0.5–177.1	1.7 (4.5)	0.9	0.5–176.1	
N interviews, *n* (%)	65,768 (100.0)			32,744 (49.8)			33,024 (50.2)			
N interviews with at least 1 correction, n (%)	55,066/65,768 (83.7)			27,721/32,744 (84.6)			27,345/33,024 (82.8)			

Total percentages may not add up or exceed one hundred due to rounding up
FPH full pregnancy history module, *FBH+* full birth history module with additional questions on pregnancy losses
[$]*P* values for independent sample *t* test that compared means for two groups
[a]All timestamped entries, including answer corrections
[b]All answers, accounting for correction time

time spent on correction was 0.1 min longer for FPH than FBH+), and the median time to correct one question was 0.08 min. Without corrections, the median response time per interview would be lower by 5.5% or 0.4 min.

Objective 2: To determine whether question characteristics (type, nature, structure) affect the duration of section completion and answer correction rates

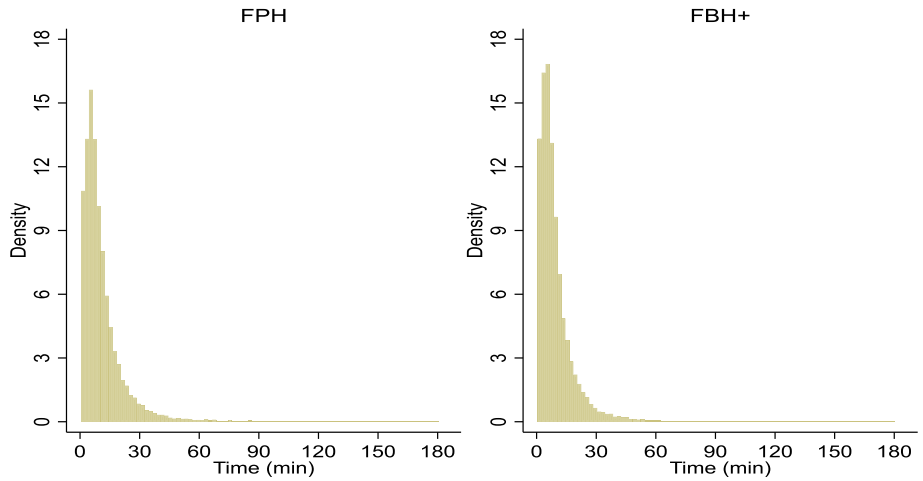
Duration of question completion

The median response time per question type was longest for free-text and multi-select questions (0.6 min), followed by date-related (0.2 min), numerical computational (0.1 min), and single-select questions (0.05 min) (Fig. 4). The median response time only slightly varied per question's nature, with the longest median response time for TOP-related questions (0.1 min). Questions with built-in error notifications had a median response time of 0.1 min compared with 0.07 min for questions with no built-in error notifications.

Proportion of corrections

Of all asked and answered questions, 7.5% had at least one or more corrections (Fig. 4, Additional file 5). Among all questions, the most frequently (by absolute number) corrected questions were single-select and numerical computational types of questions, regular questions, and questions without built-in notifications. However, the highest proportion of corrections within the question type group was multi-select questions (44.8%), followed by free-text (16.2%), numerical computational (9.7%), single-select (6.6%), and date-related (4%) questions. The proportion of corrections was similar based on the question's nature (around 6–8%). The proportion of corrected questions was 3.3% higher for questions with notifications (9.8%) than for questions without notifications (6.5%).

Among corrected questions, the median number of corrections was one, regardless of question characteristics. The maximum number of corrections was the highest for single-select (n = 28) and numerical computational (n = 23) types of questions. In terms of questions' nature, the maximum number of corrections was among regular questions



Notes: We restricted analyses of time taken to complete Section 2 (subsections 2.1-2.3) to interviews lasting 0–180 min to exclude implausible values. FPH - full pregnancy history module. FBH+ - full birth history module with additional questions on pregnancy losses.

Fig. 3 Time to complete questions regarding maternity history (section 2) for the two survey modules (*N* = 60,871)

(*n* = 28), followed by death-related questions (*n* = 23) and TOP-related questions (*n* = 19). Regarding the question structure, the maximum number of corrections was highest for questions with no built-in notifications (*n* = 28).

Based on the results of the regression analyses (Table 2, model 1), date question type (reference—single) and death-related questions (reference—regular) decreased the probability of making corrections. All other question characteristics increased the probability of making answer corrections when compared to their reference groups. Question characteristics (numeric, date, multi-select, death- and TOP-related) were positively associated

	All questions							Per corrected question		
Question characteristics	**N questions asked**		**N questions with corrections**		**Response time per 1 question, min**			**N corrections**		**Max N corrections**
	n	%	n	%	Mean (SD)	Median	Range	Mean (SD)	Median	n
Overall	3,340,749		251,168	7.5	0.37 (3.60)	0.08	0-179.9	1.14 (0.51)	1	28
Type										
Single-select	2,248,565	67.3	147,713	6.6	0.33 (3.35)	0.05	0-179.9	1.14 (0.49)	1	28
Multi-select	192	0.01	86	44.8*	1.99 (4.95)	0.58	0.07-48.0	2.13 (2.33)	1	19
Numerical computational	1,048,996	31.4	101,611	9.7	0.46 (4.03)	0.10	0-179.9	1.15 (0.52)	1	23
Date-related	42,582	1.3	1,691	4.0	0.43 (2.70)	0.20	0-176.1	1.19 (0.64)	1	8
Free-text	414	0.01	67	16.2*	2.14 (9.25)	0.58	0-139.3	1.21 (0.59)	1	5
Nature										
Regular	2,913,144	87.2	221,948	7.6	0.37 (3.58)	0.08	0-179.9	1.13 (0.48)	1	28
Death-related	346,112	10.4	22,673	6.6	0.37 (3.66)	0.07	0-179.9	1.22 (0.68)	1	23
TOP-related	81,493	2.4	6,547	8.0	0.39 (3.11)	0.13	0-164.2	1.19 (0.62)	1	19
Structure										
Built-in error notification	1,012,871	30.3	99,538	9.8	0.45 (3.95)	0.10	0-179.9	1.15 (0.52)	1	23
No built-in error notification	2,327,878	69.7	151,630	6.5	0.34 (3.40)	0.07	0-179.9	1.14 (0.50)	1	28

Notes: Total percentages may not add up or exceed one hundred due to rounding up. Subjective correction proportion colour-coding: green - 0-5% - excellent, lime - 6-10% very good, yellow - 11-15% good, orange - 16-20% moderate, and red - >20% - poor. * - Results for guidance only and should be treated with caution, given the design, and purpose of the question.

Fig. 4 Question corrections by question type, nature and structure

Table 2 Question characteristics associated with answer correction probabilities and frequencies

Characteristics	Model (1)						Model (2)					
	Part 1 correction (yes/no)			Part 2 number of corrections			Part 1 correction (yes/no)			Part 2 number of corrections		
Variables	Coef	Robust Std. Err.	*p* value	Coef	Robust Std. Err.	*p* value	Coef	Robust Std. Err.	*p* value	Coef	Robust Std. Err.	*p* value
Type, Ref (single)												
Numerical computational	0.056	0.012	< 0.001	0.050	0.007	< 0.001	-0.138	0.014	< 0.001	0.026	0.008	0.001
Date-related	− 0.531	0.025	< 0.001	0.055	0.013	< 0.001	-0.411	0.028	< 0.001	0.072	0.015	< 0.001
Free-text	0.850	0.146	< 0.001	0.031	0.061	0.611	0.864	0.164	< 0.001	0.041	0.075	0.581
Multi-select	2.248	0.145	< 0.001	0.590	0.118	< 0.001	2.337	0.158	< 0.001	0.598	0.123	< 0.001
Nature, Ref (regular)												
Death-related	− 0.166	0.009	< 0.001	0.079	0.004	< 0.001	-0.187	0.011	< 0.001	0.078	0.005	< 0.001
TOP-related	0.205	0.015	< 0.001	0.040	0.007	< 0.001	0.316	0.017	< 0.001	0.055	0.008	< 0.001
Structure, Ref (no notification)												
Yes, with notification	0.400	0.012	< 0.001	− 0.041	0.007	< 0.001	0.486	0.014	< 0.001	− 0.026	0.008	0.001
Module, Ref (FPH)												
FBH+	0.015	0.009	0.101	− 0.003	0.003	0.248	0.012	0.009	0.169	− 0.003	0.003	0.333
Constant	− 2.661	0.007	< 0.001	0.125	0.002	< 0.001	-2.818	0.056	< 0.001	0.1465	0.016	< 0.001
Observations	3,340,189			250,608			2,247,142			152582		
R-squared	0.004			0.004			0.014			0.014		
Root MSE	0.263			0.511			0.250			0.494		

Model (1) unadjusted, model (2) adjusted for respondent's characteristics (age, education, parity, wealth quintile) and location, with interviewer clustering accounted as a fixed effect. Both models accounted for clustering of individual responses within individual women (interview)
FPH full pregnancy history module, *FBH+* full birth history module with additional questions on pregnancy losses

with the number of corrections. Questions with notifications were negatively associated with the number of corrections. There was not enough evidence of an association between belonging to either of the two reproductive modules with either the probability of making a correction or the number of corrections.

After adjusting for respondent's characteristics and location, with interviewer clustering accounted as a fixed effect (Table 2, model 2), all question characteristics remained significantly associated with the probability of making answer corrections when compared to their reference groups; however, the numeric type changed the direction of the association. Numeric, date, multi-select, and death- and TOP-related questions continued to be positively associated with the number of corrections, while the questions with notifications remained negatively associated with the number of corrections.

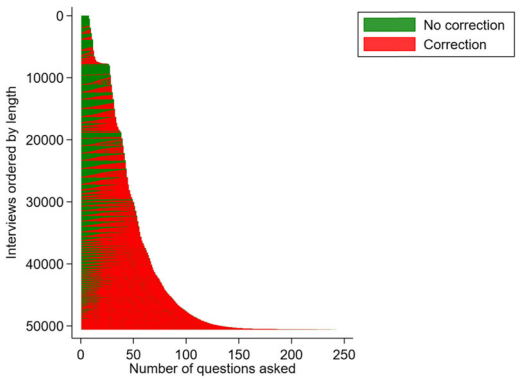
Objective 3: To propose and test classification of answer correction types and determine whether they differ by two reproductive modules

Correction patterns

In line with our findings for objectives 1 and 2, the visual inspection of the sequence index plot (Fig. 5) showed that most of the interviews had corrections to answers. Only a smaller number of shorter interviews seemed to have no or a limited number of corrections. As the number of asked questions during the interview increased, so did the number of answer corrections. Based on the number of corrections per question, we identified 23 correction patterns, ranging from one to 28 corrections (Additional file 6). Most questions had single correction (89.0%), followed by multiple corrections (two and three corrections, 8.8% and 1.6%, respectively). The remaining 0.6% of questions had four and more corrections per question. We also observed that among these correction patterns, the original answer (first entry) sometimes matched the final answer correction (last entry), while for others, it did not. Some patterns consisted of either repetitive sequences of identical entries or a combination of different entries.

Correction types

By combining characteristics of answer correction process (single vs multiple corrections; first and last answer match vs do not match) and correction pattern (different vs identical entries), we developed a classification of answer correction types. We distinguished between eight possible answer correction types after accounting for possibly missing answers after corrections (Table 3). Building on the



Notes: All interviews are presented as a horizontal sequence line, consisting of data entries with and without data correction. Correction group includes both original corrected questions and corrections. Interviews are sorted by the number of questions asked in Section 2 and the number of answers (sequence length). The red colour (representing corrected questions and corrections) might visually overlay and mask questions without corrections (presented by green colour) due to the resolution and a high number of interviews.

Fig. 5 Data collection and correction as a sequence index plot by the length of interview

original terminology used by Elliot (1934) for cycles and pattern market behaviours and using elements of the original terminology for corrective waves (i.e., flat, zigzag) [19], we named correction types as no correction, impulsive, flat (simple), zigzag, flat zigzag correction, missing after correction, missing after a flat (or zigzag) correction, missing, or incomplete. Additionally, we distinguished between non-problematic (no correction, repetitive entry of identical values, or accidental mistake) and problematic (multiple corrections, first and last entry do not match, or missing after correction) correction types.

We tested our proposed classification using our survey data. Out of eight proposed answer correction types (Table 3), we observed only five, including no correction type (Table 4). Among answers with corrections, a flat (or simple) correction was the most frequent answer correction type (84.6%), followed by zigzag and flat zigzag corrections (5.3–5.5%), and the impulsive correction type (4.6%). As we did not treat the 'Do not know' as a missing answer and had no observations with truly missing answers, we did not observe the other four proposed answer correction types.

The flat correction type was the most frequently observed correction type (84.1%), irrespective of question type, nature, or structure (Table 4). A zigzag correction was the second most frequent type of correction for single-select questions (6.9%), while impulsive and flat zigzag correction types were the second most frequently observed correction types for date-related (12.6%) and numeric computational (33.7%) question types, respectively. Accounting for the question's nature, a zigzag correction was the second most frequently observed correction type for death-related questions and TOP-related questions. For regular questions, the observed proportions were similar. Among questions with built-in notifications, the flat zigzag correction was the second most frequent type, while for questions with no built-in notifications, it was the zigzag correction.

Discussion

This is the first study to examine the feasibility and usefulness of paradata to enhance household survey capture of pregnancy outcomes to inform the content, timing, process, and administration of questions. We delve further into our earlier findings on the lack of statistically significant differences in response times for FPH or FBH+ modules' completion [11]. Around 84% of interviews had at least one correction to questions; however, most of them were simple one-time corrections. We identified four out of eight proposed answer correction types (impulsive, flat (simple) correction, zigzag correction, and flat zigzag correction) and found that question characteristics (type, content, structure) could affect the probability of making answer correction, be associated with the number of corrections, and vary in answer correction types. They remained to be significant predictors even after adjusting for respondent's characteristics and location, with interviewer clustering accounted as a fixed effect. At the same time, the correction patterns based on the number of corrections per question and answer correction types were very similar between the two reproductive modules. The latter two were also not found to be significantly associated with the probability of making answer corrections and the number of answer corrections.

Even though the median number of corrected questions per interview and corrections per one question was relatively low (2 and 1), the maximum numbers of corrected questions and corrections per question were high (110 and 28, respectively). Not only does this add to the duration of section completion (as shown in our results), and ultimately to interview duration, it could also potentially shed light on questions that are poorly understood or misunderstood by either respondent or interviewer. For example, in our survey, the highest proportion of corrections was among the numerical computational (9.7%) and TOP-related (8.0%) questions. At the same time, even higher proportions of corrections were observed for multi-select questions (44.8%) and free-text (16.2%) questions. They also had the strongest association with the probability of making answer corrections. However, these results should be treated with caution as these questions constitute less than 0.1% among all questions asked.

Moreover, given the design and related data entry process (i.e., sequential data entry for a combination of multiple answer options and noting down and correcting

Table 3 Classification of answer correction type and possible explanations

Answer correction pattern (from first to final answer entry)	Multiple (> 1) corrections	First and last entry match	Correction type	Problematic	Possible explanation
A[a]	NO	YES	No correction	NO	n/a
A > A > A	YES	YES	Impulsive	NO	Multiple tapping (e.g., due to tablet's lag response)
A > B[b]	NO	NO	Flat (simple) correction	NO	Accidental mistake, correction, following clarification or mistake
A > n[c] > A	YES	YES	Zigzag correction	YES	Corrections after multiple additional clarifications and confirmation (e.g., change in responses) and changed back to the original answer
A > n > B	YES	NO	Flat zigzag correction	YES	Correction, following clarifications
A > …	NO	NO	Missing after correction	YES	Accidental mistake corrected after clarification, request to skip or not to record the answer
A > B > … or A > n > B > … or A > n > A > …	YES	NO	Missing after flat (zigzag) correction	YES	Clarification and consequent refusal to answer or request to skip or not to record the answer
…	NO	YES	Missing/incomplete	YES	Unknown

[a]stands for the original (first entered) answer
[b]represents a different answer in content than the original answer A
[c]represents a number of answer corrections between first and last answer in sequence

the free text), distinguishing between answers, answer combinations, and corrections might not always be straightforward. Nonetheless, we suggest that these types of questions get additional attention during training sessions, with more time being allocated to explaining and practising asking these questions, with additional guidance and supervision provided during the fieldwork. One could also consider limiting even further these types of questions in household surveys.

We also developed and tested a new classification of answer correction types. We found this classification useful and suggest it for future studies. For example, we found that almost 90% of all corrections were simple mistake corrections (which is less worrisome) or impulsive and repetitive answers (most likely due to non-responsive screen); hence, they should not be considered problematic. This suggested that the reasons for the remaining multiple zigzag corrections (around 10%) lie elsewhere. We speculate

Table 4 Answer correction types per question type, nature, and structure

Question characteristics	*n*	Flat (simple)		Impulsive		Zigzag correction		Flat zigzag correction	
		n	%	*n*	%	*n*	%	*n*	%
Overall	250,608	212,057	84.6	11,438	4.6	13,288	5.3	13,825	5.5
Type									
Single-select	147,157	123,818	84.1	7613	5.2	10,124	6.9	5602	3.8
Multi-select	86	47	54.7	4	4.7	6	7.0	29	33.7
Numerical computational	101,607	86,850	85.5	3604	3.5	3080	3.0	8073	7.9
Date-related	1691	1290	76.3	213	12.6	76	4.5	112	6.6
Free-text	67	52	77.6	4	6.0	2	3.0	9	13.4
Nature									
Regular	221,447	188,670	85.2	9917	4.5	10,774	4.9	12,086	5.5
Death-related	22,628	17,871	78.9	1390	6.1	1905	8.4	1462	6.5
TOP-related	6533	5516	84.4	131	2.0	609	9.3	277	4.2
Structure									
Built-in error notification	99,532	85,244	85.6	3492	3.5	2979	3.0	7817	7.9
No built-in error notification	151,076	126,813	83.9	7946	5.3	10,309	6.8	6008	4.0

that the remaining corrections were made following additional clarifications or confirmations of previous or later answers (Table 1). However, to verify our assumptions and identify other factors that affect correction probability and frequency (e.g., the exact wording or any other contextual factors) and even further unpack reasons behind answer corrections, field observations and qualitative interviews with both respondents and interviewers will be necessary. This once again underlines the importance of adequate timing dedicated to the data collection training, extensive field questionnaire testing, and effective supervision and guidance.

We believe that our proposed classification of answer correction types accounts for several dimension of the answer correction process. However, we would like to invite other researchers to evaluate our classification, improve and optimise it further, and test its usefulness and applicability in other types of surveys and research settings. Using our survey data, we identified only five answer correction types (including no corrections), lacking missing or incomplete answers after correction. This is primarily due to a lack of 'missing' (or empty in content) timestamped entries in our paradata dataset (which by default is not possible) and our decision not to treat 'Do not know' answers as 'missing' entries. However, we also much acknowledge effective training and comprehensive training manuals, diligent work of our data collectors, and their dedication to prompt interviewees and complete all relevant fields, which we believe aspired minimisation of any missing data. About a third of all questions in our questionnaire had inbuilt error notification, prompts, and warnings. Our results suggest that such notifications are effective since the proportion of corrected questions was 3.3% higher for questions with notifications compared to those without them, and having notifications was significantly associated with the probability of making corrections but negatively associated with the number of such corrections. Additionally, Survey Solutions application had a built-in colour coding indicating survey section completion (red for incomplete and unanswered questions and green for complete), which prompted data collectors to answer all questions. For example, during some training sessions at several data collection sites, interviewers were insisting on learning how to achieve completeness 'having all sections colour-coded as green' in all survey sections and were 'somewhat unhappy' to finish the exercise with one or more sections remaining incomplete (or red).

Strength and limitations

Given our focus on corrections during interviewing, in our analysis, we excluded implausible and impractical values (over 180 min for section completion duration) but allowed for multiple corrections during the module completion and switching between questions and questionnaire sections. We assumed that such restrictions could provide meaningful and practical insights into face-to-face data collection process, even if it would exclude and not account for long breaks in the interviews (stopped and resumed several days/weeks/months later) or other errors in timestamps (e.g., resulting from a change of a tablet's calendar set up from local form to the Gregorian calendar during data collection). However, we recognise that this decision could be considered as one of the limitations of the study, as it potentially did not capture corrections based on office data quality and error checks (following which questionnaires were returned and/or reassigned back to the interviewer for correction in-field). Moreover, in our analyses, we specifically focused only on a subset of paradata that related to answers and corrections. We did not utilise the data with timestamped events that related to process-related activities (e.g., interviewer or supervisor comments; enabling and disabling questions; declaring answers as valid or invalid based on the passing or failing of programmed validation rules; switching between the questionnaire's translations; recalculating system variable values based on manual correction), which could be considered another limitation of the study. As paradata were not readily available for export at the beginning of our data collection (due to software limitations), we did not evaluate individual and team productivity (e.g., average hours per contact attempt, contact attempts without success, number of interviews per workday), or estimate the response likelihood and perform measurement error evaluation [6]. However, we relied on experience from our local data collection teams and invested additional time into training and field testing.

Research gaps for improving measurements of MNCH indicators in household surveys

Given a lack of other studies that have used paradata in MNCH field, we cannot compare our findings directly to other studies, and we would like to stimulate the wider use of survey paradata to advance survey design and implementation for collecting information on pregnancies and births and for other purposes.

Paradata provide a wealth of information and could augment surveys, particularly overseeing the data collection process. Not surprisingly, it has already found use in other health and medical areas. For example, similar to our study, paradata were previously used in telemedicine research to estimate time spent to complete a questionnaire [20] and to examine completion and impact of push notifications on data completion in behaviour risk assessment [21]. Other applications in

health include examining the role of paradata in non-response adjustment process [22], underreporting errors and finding suggestions for methodological improvement for future surveys [23] and examining response time at the level of individual questions [24]. Other studies examined practical use of paradata, for example, as an interactive web-based data visualisation tool, providing survey staff with the information to monitor data collection daily [25]. Using paradata (along with metadata and embedded data) can also improve response rates, identify bias, and give a possible explanation for apparent outlier responses, providing an efficient method of conducting web-based Delphi surveys [26]. Overall, using paradata in health research suggests that paradata could be valuable in quantifying recruitment efforts and aid the development and evaluation of new recruitment strategies [27].

Future analyses could investigate the relationship between correction rates and correction type and being supervised by a supervisor or other colleagues, which potentially could prompt additional corrections under peer pressure. Other potential uses of paradata in MNCH research could include effort indicators, tracking individual and team productivity, estimating contact attempts without success, and response propensity. Outcome indicators and case status indicators can also include non-interviews by type and refusal patterns by respondent characteristics. Paradata in MNCH research can also be used to generate a dashboard/monitoring system or a validation system for collected data against external sources of information, hence, automatically flagging incorrect entries in the interviews.

Conclusion

Accurate estimation of coverage indicators from household surveys is vital but contingent on data quality; hence, a better understanding of how to improve the questionnaire design and survey implementation is crucial. Paradata have the potential to enhance survey design and implementation for collecting information on pregnancies and births, leading to improved metrics of measurement in maternal and newborn health research. They can help to identify questions and sections with multiple corrections and shed light on typically hidden challenges in the survey's content, process, and administration. Overall, our experience suggests that given the size of paradata and their complex structure, analysis is not always straightforward, and consideration should be given to the additional data management and programming skills required. Nonetheless, paradata provide a wealth of data, can improve the process of data collection using live survey monitoring, and can add value in improving survey data quality as well as efficiency.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12963-020-00241-0.

Additional file 1. Detailed overview of questions in Section 2 in FPH and FBH+.

Additional file 2. Example of survey paradata structure.

Additional file 3. STROBE guidelines checklist.

Additional file 4. Types of questions in Section 2 in FPH and FBH+.

Additional file 5. Detailed overview of questions correction by question type, content and structure and reproductive module.

Additional file 6. Correction patterns, 23 groups by the number of corrections per question.

Additional file 7. Ethical approval of local Institutional Review Boards.

Abbreviations

CAPI: Computer-assisted personal interviewing; CAWI: Computer-assisted web interviewing; DHS: Demographic and Health Surveys; EN-INDEPTH: Every Newborn International Network for the Demographic Evaluation of Populations and Their Health; FBH+: Full birth history (+ denotes additional questions on pregnancy losses); FPH: Full pregnancy history; HDSS: Health and Demographic Surveillance System; LMIC: Low- and middle-income countries; MICS: Multiple Indicator Cluster Survey; MNCH: Maternal, newborn, and child health; TOP: Termination of pregnancy

Acknowledgements

This supplement is dedicated to the memory of Professor Peter Byass, who was the Senior External Editor of the supplement. Peter died suddenly in August 2020 and will be greatly missed by the EN-INDEPTH study team and the entire global health community.

We thank the 118 interviewers and many Health and Demographic Surveillance System staff participating in this study for their hard work and dedication to this project. Many thanks to Samuelina Arthur, Claudia DaSilva, Olivia Nakisita, Maria Cesay, and the relevant site staff for their administrative support.

We acknowledge the core funders for all sites/institutions.

We express appreciation to the EN-INDEPTH expert advisory group: Fred Arnold, Peter Byass, Trevor Croft, Kobus Herbst, Sunita Kishor, Florina Serbanescu, Stephane Helleringer, Turgay Unalan, Shane Khan, Attila Hancioglu. Additionally, we would like to thank members of the World Bank Survey Solutions technical support team for their guidance and support during survey implementation.

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.

The Every Newborn-INDEPTH Study Collaborative Group:

Senior External Supplement Editors: Peter Byass; Stephen M Tollman; Hagos Godefay

Technical Supplement Editors: Joy E. Lawn; Peter Waiswa; Hannah Blencowe

Managing Supplement Editors: Judith Yargawa; Joseph Akuze (data and statistics)

Other EN-INDEPTH Collaborative Group Members:

By team: PI followed by other members in alphabetical order

Bandim: Ane B Fisker (PI); Justiniano SD Martins; Amabelia Rodrigues; Sanne M Thysen

Dabat: Gashaw Andargie Biks (PI); Solomon Mokonnen Abebe; Tadesse Awoke Ayele; Telake Azale Bisetegn; Tadess Guadu Delele; Kassahun Alemu Gelaye; Bisrat Misganaw Geremew; Lemma Derseh Gezie; Tesfahun Melese; Mezgebu Yitayal Mengistu; Adane Kebede Tesega; Temesgen Azemeraw Yitayew

IgangaMayuge: Simon Kasasa (PI); Edward Galiwango; Collins Gyezaho; Judith Kaija; Dan Kajungu; Tryphena Nareeba; Davis Natukwatsa; Valerie Tusubira

Kintampo: Yeetey AK Enuameh (PI); Kwaku P Asante; Francis Dzabeng; Seeba Amenga-Etego; Alexander A Manu; Grace Manu; Obed Ernest Nettey; Sam K Newton; Seth Owusu-Agyei; Charlotte Tawiah; Charles Zandoh

Matlab: Nurul Alam (PI); Nafisa Delwar; M Moinuddin Haider; Md. Ali Imam; Kaiser Mahmud

LSHTM/ Makerere School of Public Health: Angela Baschieri; Simon Cousens; Vladimir Sergeevich Gordeev; Victoria Ponce Hardy; Doris Kwesiga; Kazuyo Machiyama

Published: 8 February 2021

About this supplement

This article has been published as part of *Population Health Metrics Volume 19 Supplement 1, 2021: Every Newborn-INDEPTH study: Improving the measurement of pregnancy outcomes in population-based surveys*. The full contents of the supplement are available online at https://pophealthmetrics.biomedcentral.com/articles/supplements/volume-19-supplement-1.

Authors' contributions

JEL conceptualised the EN-INDEPTH study. All EN-INDEPTH collaborative group and site teams contributed to the design of the study protocol and undertook data collection. VSG, together with HB and JEL, developed the detailed research questions and overall analysis plan for this paper. These were refined with inputs from the wider EN-INDEPTH study collaborative group including AB at a multi-country workshop in Entebbe Uganda in February 2019. SMT, FD, MMH, TAY, SMA, DN, CG, and SIA implemented the study at the sites. MW and ML were consulted specifically on the Survey Solutions software. JA prepared the survey data files. VSG and MS undertook quantitative data analysis. The manuscript was drafted by VSG, with inputs from HB and JEL. All authors reviewed and agreed on the final version of the manuscript.

Funding

The EN-INDEPTH study (including publication costs) was funded by the Children's Investment Fund Foundation (CIFF) by means of a grant to LSHTM (PI *Joy E. Lawn*) and a sub-award to the INDEPTH MNCH working group with technical leadership by Makerere School of Public Health (PI *Peter Waiswa*).

Availability of data and materials

Data sharing and transfer agreements were jointly developed and signed by all collaborating partners. The datasets generated during the current study are deposited online at https://doi.org/10.17037/DATA.00001556 with data access subject to approval by collaborating parties.

Ethics approval and consent to participate

The EN-INDEPTH study was granted ethical approval by the Institutional Review Boards in all operating countries, as well as, from the Institutional Ethical Review Committee of the London School of Hygiene & Tropical Medicine (Additional file 7). Respondents of every successful interview gave written consent/ascent after being informed of the objective and data use, the procedure of the interview, risks and benefits of participating in the study, right to withdraw from interview anytime point of time, and not responding to questions where they feel discomfort. The study ensures the respondent's privacy at data collection and confidentiality at data use.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Institute of Population Health Sciences, Queen Mary University of London, London, UK. [2]Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene & Tropical Medicine, London, UK. [3]Department of Health Policy, Planning and Management, Makerere University School of Public Health, Kampala, Uganda. [4]Centre of Excellence for Maternal Newborn and Child Health Research, Makerere University, Kampala, Uganda. [5]Bandim Health Project, Bissau, Guinea-Bissau. [6]Research Centre for Vitamins and Vaccines, Statens Serum Institut, Copenhagen, Denmark. [7]Department of Clinical Research Open Patient data Explorative Network (OPEN), University of Southern Denmark, Odense, Denmark. [8]Kintampo Health Research Centre, Kintampo, Ghana. [9]Health Systems and Population Studies Division, icddr,b, Dhaka, Bangladesh. [10]Department of Medical Statistics, London School of Hygiene & Tropical Medicine, London, UK. [11]The World Bank, Washington DC, USA. [12]Dabat Research Centre Health and Demographic Surveillance System, Dabat, Ethiopia. [13]IgangaMayuge Health and Demographic Surveillance System, Makerere University Centre for Health and Population Research, Makerere, Uganda.

References

1. Commission on Information Accountability for Women's and Children's Health. Keeping promises, measuring results: commission on information and accountability for women's and children's health. Geneva: Commission on information and accountability for Women's and Children's Health, World Health Organization; 2011.
2. Countdown to 2015 & Health Metrics Network. Monitoring maternal, newborn and child health: understanding key progress indicators. Geneva: World Health Organization; 2011.
3. Hansen SE, Benson G, Bowers A, Pennell B-E, Lin Y-cJ, Duffey B, Hu M, Hibben KC. Survey quality. In: Guidelines for best practice in cross-cultural surveys. Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan; 2016.
4. Hancioglu A, Arnold F. Measuring coverage in MNCH: tracking progress in health for women and children using DHS and MICS household surveys. PLoS Med. 2013;10:e1001391.
5. Leisher C. A comparison of tablet-based and paper-based survey data collection in conservation projects. Soc Sci. 2014;3:264–71.
6. Kreuter F, editor. Improving surveys with paradata: analytic uses of process information. Hoboken: Wiley; 2013.
7. Nicolaas G: Survey paradata: a review. 2011. http://eprints.ncrm.ac.uk/1719/1/Nicolaas_review_paper_jan11.pdf. [Accessed June 2020].
8. Choumert-Nkolo J, Cust H, Taylor C. Using paradata to collect better survey data: evidence from a household survey in Tanzania. Rev Dev Econ. 2019;23:598–618.
9. Couper MP, Kreuter F. Using paradata to explore item level response times in surveys. J R Stat Soc Ser A Stat Soc. 2013;176:271–86.
10. Baschieri A, Gordeev VS, Akuze J, Kwesiga D, Blencowe H, Cousens S, Waiswa P, Fisker AB, Thysen SM, Rodrigues A, 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.
11. Akuze J, Blencowe H, Waiswa P, Baschieri A, Gordeev VS, Kwesiga D, Fisker AB, Thysen SM, Rodrigues A, Biks GA, 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:e555–66.
12. The World Bank. Survey solutions CAPI/CAWI platform: release 5.26. Washington DC: The World Bank; 2018.
13. Thysen SM, Tawiah C, Blencowe H, Manu G, Akuze J, Haider MM, Alam N, Azemeraw T, Baschieri A, Biks GA, et al. Electronic data collection in a multi-site population-based survey: EN-INDEPTH study. BMC Population Health Metrics. 2021:19(Supplement 1).
14. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019.
15. StataCorp. Stata Statistical Software: release 16. College Station: StataCorp LLC; 2019.
16. Yan T, Olson K. Analysing paradata to investigate measurement error. In: Kreuter F, editor. Improving surveys with paradata: analytic uses of process information. Hoboken: Wiley; 2013. p. 73–95.
17. Brzinsky-Fay C, Kohler U, Luniak M. Sequence analysis with Stata. Stata J. 2006;6:435–60.
18. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335(7624):806–8. https://doi.org/10.1136/bmj.39335.541782.AD. PMID: 17947786; PMCID: PMC2034723.
19. Elliott RN. The wave principle. Los Angeles: Alanpuri Trading; 1938.
20. Romano MF, Sardella MV, Alboni F. Web health monitoring survey: a new approach to enhance the effectiveness of telemedicine systems. JMIR Res Protoc. 2016;5:e101.
21. Oakley-Girvan I, Lavista JM, Miller Y, Davis S, Acle C, Hancock J, Nelson LM. Evaluation of a mobile device survey system for behavioral risk factors (SHAPE): app development and usability study. JMIR Form Res. 2019;3:e10246.
22. Substance Abuse and Mental Health Services Administration (US). CBHSQ methodology report. In: Incorporating level of effort paradata in the NSDUH nonresponse adjustment process. Rockville: Substance Abuse and Mental Health Services Administration (US); 2013.
23. Hu M, Gremel GW, Kirlin JA, West BT. Nonresponse and underreporting errors increase over the data collection week based on paradata from the National Household Food Acquisition and Purchase Survey. J Nutr. 2017;147:964–75.

24. Craig BM, Runge SK, Rand-Hendriksen K, Ramos-Goni JM, Oppe M. Learning and satisficing: an analysis of sequence effects in health valuation. Value Health. 2015;18:217–23.
25. Murphy JJ, Duprey MA, Chew RF, Biemer PP, Harris KM, Halpern CT. RTI press occasional papers. In: Interactive visualization to facilitate monitoring longitudinal survey data and paradata. Research Triangle Park: Research Triangle Institute (RTI) Press; 2019.
26. Helms C, Gardner A, McInnes E. The use of advanced web-based survey design in Delphi research. J Adv Nurs. 2017;73:3168–77.
27. Langeheine M, Pohlabeln H, Ahrens W, Rach S. Consequences of an extended recruitment on participation in the follow-up of a child study: results from the German IDEFICS cohort. Paediatr Perinat Epidemiol. 2017;31:76–86.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

Chapter 7: Population-level comparison of pregnancy outcomes between EN-INDEPTH survey and HDSS data

7.1. Introduction

This chapter provides an in-depth analysis of available data from the EN-INDEPTH survey and HDSS data across five well-phenotyped HDSS sites in Africa and Asia to improve the measurement of pregnancy outcomes in population-based surveys and HDSS data systems (Objective 4)

7.2. List of Figures

Figure 7.1 - Linkage of pregnancy episodes captured by HDSS and EN-INDEPTH survey for four HDSS sites

Figure 7.2 - Distribution of pregnancy outcomes between EN-INDEPTH survey and Health and Demographic Surveillance Sites by module

Figure 7.3 - Distribution of pregnancy outcomes between EN-INDEPTH survey and Health and Demographic Surveillance Sites by module

Figure 7.4 - Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to Health and Demographic Surveillance Sites

Figure 7.5 - Ratio of Stillbirth Rate in EN-INDEPTH survey compared to Health and Demographic Surveillance Sites

Table S 1: Definition of independent variables to be used in this analysis

Table S 2: Questions for assessing the HDSS system for pregnancy surveillance

Table S 3: Reasons for exclusion of data by year from the supplementary analysis

Table S 4: Overview of Health and Demographic Surveillance Systems in five EN-INDEPTH study sites

7.3. List of Tables

Table 7.1 - EN-INDEPTH study sites with overview of the varying Health and Demographic Surveillance Systems

Table 7.2 - Characteristics of continuously resident women included in the EN-INDEPTH survey, for four HDSS sites

Table 7.3 - Characteristics of interviewers who participated in the EN-INDEPTH survey, by sites, for four HDSS sites

Table 7.4 - Comparison of pregnancy outcome indicators between EN-INDEPTH survey and Health Demographic Surveillance Sites by module

Table S 5a: Comparing pregnancy outcome between EN-INDEPTH survey and Health Demographic Surveillance Sites by module and year

Table S 5b: Comparison of pregnancy outcome indicators between EN-INDEPTH survey and Health Demographic Surveillance Sites by module (Truncated)

Figure S 1a: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to Health and Demographic Surveillance Sites++ by year

Figure S1b: Ratio of Stillbirth Rate in EN-INDEPTH survey compared to Health and Demographic Surveillance Sites++ by year

Figure S 2a: Distribution of pregnancy outcomes between EN- INDEPTH survey and Health and Demographic Surveillance Sites by module

Figure S2b: Distribution of pregnancy outcomes between EN- INDEPTH survey and Health and Demographic Surveillance Sites by module

Figure S 3a: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to Health and Demographic Surveillance Sites++

Figure S 3b: Ratio of stillbirth rate in EN-INDEPTH survey compared to Health and Demographic Surveillance Sites++

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1512635	**Title**	Mr
First Name(s)	Joseph		
Surname/Family Name	Akuze Waiswa		
Thesis Title	Measurement of stillbirths and neonatal deaths in standardised population-based surveys		
Primary Supervisor	Hannah Blencowe		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Not Applicable		
When was the work published?	Not Applicable		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	Not Applicable		
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMC Population Health Metrics
Please list the paper's authors in the intended authorship order:	Joseph Akuze*, Hannah Blencowe, Ane B Fisker, Bisrat Misganaw Geremew, Davis Natukwatsa, Francis Dzabeng, M Moinuddin Haider, Kazuyo Machiyama, Angela Baschieri, Vladimir Gordeev, Joy E Lawn, Peter Waiswa^, Simon Cousens^ and the Every Newborn-INDEPTH study

	collaborative author group
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I led the conceptualisation and writing of the manuscript. I developed and revised the detailed research questions with feedback from co-authors. I conducted the literature reviews, and data management. With guidance from Prof. Simon Cousens and Dr. Hannah Blencowe, I developed the detailed analysis plan, and conducted the data analysis. I drafted the initial draft of the paper. I led the writing and preparation of manuscript for submission and coordinated the reviews process by all authors.

SECTION E

Student Signature	Joseph Akuze Waiswa
Date	27th September 2021

Supervisor Signature	Hannah Blencowe
Date	28th September 2021

7.4. Research paper

E: Comparing population-level pregnancy outcomes between population-based surveys and prospective demographic surveillance systems: an EN-INDEPTH study

SUPPLEMENT TITLE: Every Newborn-INDEPTH multi-country study to improve population-based survey data for pregnancy outcomes

PAPER TITLE: Comparing population-level estimates of pregnancy outcomes between population-based surveys and prospective health demographic surveillance systems: an EN-INDEPTH study

Joseph Akuze*, Hannah Blencowe, Ane B Fisker, Bisrat Misganaw Geremew, Davis Natukwatsa, Francis Dzabeng, M Moinuddin Haider, Kazuyo Machiyama, Angela Baschieri, Vladimir Gordeev, Joy E Lawn, Peter Waiswa^, Simon Cousens^ and the Every Newborn-INDEPTH study collaborative author group:

*Corresponding Author

^ Senior Authors

Authors' contact email addresses

Joseph Akuze: joseph.waiswa@lshtm.ac.uk
Hannah Blencowe: hannah-jayne.blencowe@lshtm.ac.uk
Ane B Fisker: a.fisker@bandim.org; afisker@health.sdu.dk
Bisrat Misganaw Geremew: bisratcsa@gmail.com
Davis Natukwatsa: natukwatsa.davis@gmail.com
Francis Dzabeng: francis.dzabeng@gmail.com
M Moinuddin Haider: moin@icddrb.org
Kazuyo Machiyama: Kazuyo.Machiyama@lshtm.ac.uk
Angela Baschieri: angela.baschieri@lshtm.ac.uk
Vladimir Gordeev: v.gordeev@qmul.ac.uk
Joy E. Lawn: joy.lawn@lshtm.ac.uk
Peter Waiswa: pwaiswa@musph.ac.ug
Simon Cousens: simon.cousens@lshtm.ac.uk

Word count (excluding title page, abstract, references, figures and tables): 5634

Tables to go at the end of the manuscript (n= 3)

Figures to go at the end of the manuscript (n= 3)

Web appendix/supplementary materials (separate file)

Keywords: neonatal, stillbirth, survey, surveillance, probabilistic data linkage.

ABSTRACT

BACKGROUND: Surveys are important sources of stillbirth rate (SBR) and neonatal mortality rate (NMR) estimates. The EN-INDEPTH study compared a Full Pregnancy History (FPH) and a Full Birth History plus pregnancy loss questions (FBH+) in five Health and Demographic surveillance system (HDSS) sites (Matlab-Bangladesh, Kintampo-Ghana, Dabat-Ethiopia, IgangaMayuge-Uganda, Bandim-Guinea-Bissau). The FPH and FBH+ produced similar NMR estimates, but higher SBR estimates in FPH with between-site heterogeneity. This paper explores between-site variation using HDSS data and reviews HDSS systems.

METHODS: We reviewed the HDSS systems in all sites, including document and site profile reviews and consultations with site representatives. In four sites, excluding Bandim, we compared numbers and rates (NMRs, and SBRs) from HDSS data for women continuously resident since January 1 2012 with those from both arms of the EN-INDEPTH survey using descriptive and graphical methods. Site-specific risk ratios were combined using standard meta-analysis techniques. Supplementary analyses, excluding HDSS site-years with data collection challenges, were performed.

RESULTS: Pregnancy surveillance systems differed between sites in the frequency of surveillance visits, main respondent reporting pregnancy outcomes, pregnancy testing, and data capture and surveillance modes. Reported challenges to quality of pregnancy and outcome surveillance included misclassification between stillbirth and neonatal deaths and missing data.

In three of the four sites, both arms of the EN-INDEPTH survey reported more pregnancies than HDSS. Only Matlab's HDSS, with a high frequency of surveillance rounds and routine early pregnancy testing, recorded more pregnancies than the EN-INDEPTH survey. Overall, survey data produced higher estimated SBRs and NMRs than the HDSS data (SBR-FBH+: RR=1.13, 95%CI (0.79, 1.63); SBR-FPH:RR=1.20, 95%CI (0.76, 1.90); NMR-FBH+:RR=1.19, 95%CI (0.98, 1.43); NMR-FPH:RR=1.15, 95%CI (1.01, 1.30)). However, in Matlab, HDSS recorded more stillbirths and miscarriages than the survey in the FPH arm. In Kintampo, HDSS recorded more stillbirths than the survey, even though the HDSS recorded fewer pregnancies.

CONCLUSIONS: Pregnancy surveillance methods and data quality in HDSS sites is highly variable. HDSS system strengthening (in terms of survey rounds, standardisation of approaches, and quality assurance) is needed to increase the completeness of pregnancy surveillance and recording of pregnancy outcomes to enable HDSS to achieve their potential to increase understanding and improve population pregnancy and newborn health.

KEY FINDINGS TEXT BOX

WHAT IS NEW?

What was known already:

- Standardised household surveys are an important source of stillbirth and neonatal data in high burden countries. Prospective Health and Demographic Surveillance Sites' (HDSS) systems enable epidemiological and demographic studies within a defined cohort, but few studies have examined data capture and quality for pregnancies and outcomes.
- The EN-INDEPTH study compared two survey approaches in five sites and found that Full Pregnancy History (FPH) produced higher stillbirth rate estimates than a Full Birth History plus pregnancy loss questions (FBH+), however, there were between site variations. No difference in NMRs were found.

What was done:

To explore between site variation in the EN-INDEPTH study:

- A formative description was conducted on the five HDSSs' pregnancy surveillance processes and methods.
- SBRs and NMRs computed for the period after January 2012 were compared between EN-INDEPTH survey estimates and HDSS estimates in four sites (N = 47,464).
  - Supplementary analyses were conducted excluding data-years with HDSS data collection challenges identified (Dabat: 2012-2013; 2017-2018, IgangaMayuge: 2017-2018, Matlab: 2017-2018 and Kintampo: 2016-2018).

WHAT WAS FOUND?

Comparison of processes and methods of pregnancy surveillance systems (five HDSS):

- Sites were established at different times, with different purposes, mainly focused on childhood outcomes, and hence have differing pregnancy and birth outcome surveillance systems.
- Differences in data collection, frequency of surveillance rounds, choice of respondent for pregnancy and outcome questions, use of pregnancy testing, use of household-held registers, indicator definitions, missing data, and varying misclassification between stillbirth and neonatal deaths hinder cross-site comparisons. In recent years, data loss during migration to new electronic systems presents a further challenge to comparability for Dabat.

Population-level comparison (four HDSS):

- The EN-INDEPTH survey recorded many more pregnancies, irrespective of the survey module used, than the HDSS, except in Matlab where the HDSS recorded slightly more pregnancies than the survey.
- Survey data produced higher neonatal mortality rates (NMR) estimates than HDSS data in all sites and study arms except in the FBH+ arm in Dabat. No substantial between-site variation was observed.
- Survey data produced, on average, higher Stillbirth Rate (SBR) estimates than HDSS data, but there was substantial between-site variation. In both survey and HDSS data, SBR estimates from FPH were higher than FBH+ in all sites except Matlab

WHAT NEXT IN MEASUREMENT AND RESEARCH?

- **Measurement improvement now:**
  - **HDSS:** Pregnancy surveillance methods and data quality in HDSS sites is highly variable. Understanding the different contexts and methodological variation is important when interpreting results from HDSS data.
  - **Surveys:** The Demographic and Health Surveys program is using FPH in its 8th phase, based on results from the recent EN-INDEPTH study, which showed a better capture of stillbirths in the FPH than in the FBH+ arm except in Matlab. Further analysis indicates that the survey captured stillbirths better than the HDSS except in Matlab. Yearly comparisons of the number of reported pregnancies between HDSS and survey found inconsistencies in all sites, although more years were affected in sites with infrequent pregnancy surveillance.
- **Research needed:** Individual-level data linkage between surveys and HDSS would further enhance our understanding of differences in data capture and quality for improving stillbirth and neonatal mortality estimates. More standardised methods of obtaining accurate data and data quality assessment are important for improving both HDSS and survey data.

BACKGROUND

Annually, an estimated 2.0 million stillbirths and 2.4 million neonatal deaths occur globally. Low-income and middle-income countries (LMICs), which often have weaker civil registration and vital statistics (CRVS) systems, bear the highest burden of these adverse birth outcomes.[3,10,42,43] Substantial progress during the Millennium Development Goals era halved stillbirths and neonatal deaths.[44] However, further reductions are required to meet the Sustainable Development Goals and Every Newborn plan (ENAP) targets of 12 or fewer neonatal deaths per 1,000 livebirths and 12 or fewer stillbirths per 1,000 total births in every country before the year 2030.[44,45] With less than a decade left, achieving these targets will depend on improved care and services, increased availability of data, and better gestational age assessment to reduce misclassification, among other innovations.[44,45] LMICs rely largely on large-scale household surveys, notably Demographic Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS), to gather nationally representative data on stillbirths and neonatal deaths, however, these are known to underestimate adverse pregnancy outcomes.[46-48] These surveys capture such data through maternity histories, either a Full Birth History or Full Pregnancy History.[12,38,46, 49]

Health and Demographic Surveillance Sites (HDSS) provide a longitudinal picture by following a population cohort, gathering demographic, socio-economic, and other data, including pregnancies, births, deaths, and migration at regular intervals. The number of pregnancy events recorded in an HDSS varies by the type of pregnancy surveillance system.[50] Most HDSSs are located in LMICs, and some belong to networks with similar research, such as the International Network for the Demographic Evaluation of Populations and their Health Network (INDEPTH Network), ALPHA Network, and South African Population Research Infrastructure Network (SAPRIN).[51 48] The 53 INDEPTH sites follow close to 4 million people located in Africa, Asia and Oceania. This geographical diversity provides an opportunity for multi-site and multi-country epidemiological studies to compare and test various innovations.[48,51] Although HDSS systems have great potential, their differences (i.e., definitions of key indicators, missing data, different data capture systems, and some indicators not being part of every HDSS's priority list for surveillance) hamper comparability.[1]

A multi-country study, "The Every Newborn-International Network for the Demographic Evaluation of Populations and their Health (EN-INDEPTH) study", was undertaken in five HDSS

within the INDEPTH network.[51] The EN-INDEPTH study undertook a randomised comparison of two methods of retrospective recording of pregnancy outcomes: a Full Birth History with additional questions on pregnancy losses (FBH+), the standard approach used in DHS-7, and a Full Pregnancy History (FPH). The five sites were selected out of fourteen applications in response to a call from ENAP metrics that was advertised within the INDEPTH Network in 2015.[36,51] These five sites were selected based on each having a population greater than 30,000 people; recorded annual stillbirth rate (SBR) and neonatal mortality rate (NMR) greater than 15 per 1,000 livebirths, presence of pregnancy surveillance with stillbirths and neonatal deaths included; and availability of co-funding from the sites.[13] The EN-INDEPTH study found the stillbirth rate (SBR) was, on average, 21% higher when using the FPH module rather than the FBH+, but with strong evidence of between-site heterogeneity. Neonatal mortality rates (NMRs) recorded by the two modules were similar in every site.[17] Details of the study protocol and the results relating to the primary objective are published elsewhere.[36,50]

This paper, which is one of a series of papers from the EN-INDEPTH study, seeks to document the pregnancy surveillance systems and to understand further the differences in population-level estimates observed in the main EN-INDEPTH study by addressing the following objectives:

1) Review data collection processes and methods for pregnancy and birth outcomes surveillance for the five HDSS.
2) Compare SBRs and NMRs estimated from the two arms of the EN-INDEPTH survey (FBH+/ FPH) with the routine HDSS data for four HDSS with comparable data (excluding Bandim, which does not collect retrospective pregnancy outcome information).

METHODS

Overall study design and setting

The EN-INDEPTH multi-site study was a retrospective population-based cross-sectional survey conducted in five prospective HDSS located in Africa and Asia; Bandim in Guinea-Bissau, Dabat in Ethiopia, IgangaMayuge in Uganda, Matlab in Bangladesh, and Kintampo in Ghana. The survey was undertaken between July 28, 2017, and August 13, 2018, amongst women of reproductive age (15-49 years). The EN-INDEPTH study's eligibility criteria were slightly

different in each site i.e., all women of reproductive age in Dabat, IgangaMayuge and Kintampo; women with a known and recorded pregnancy and birth outcome in the last five years in Matlab and Bandim. In the Bandim rural site, a random sample of eligible women was taken.[36,50]

The EN-INDEPTH survey included 69,176 women who were randomised between the FBH+ arm (n=34,805) and the FPH arm (n=34,371).[17] The survey questions were based on DHS-7 questionnaires (i.e. FBH+, DHS's core module and FPH that was recently used in Nepal).[39] We recruited local interviewers who were familiar with the cultures and dialects used within the study sites. Interviewers completed an interviewer survey. Survey data from women and interviewers were collected using tablets with the Android operating system, and the data was recorded using the World Bank's Survey Solutions application.[52] All women were asked for information on all livebirths in their lifetime and pregnancy losses in the last 5 years (FBH+), or all pregnancies in their lifetime regardless of the outcome (FPH).[36,53] After completing data collection, syntax files were developed by statisticians at London School of Hygiene and Tropical Medicine (LSHTM) and Makerere University School of Public Health (MakSPH) and shared with the scientists in five sites who prepared anonymised data. The anonymised data were encrypted and then shared using secure Microsoft SharePoint folders.[52]

For the present analyses, the outcomes analysed were abortion/miscarriage (defined as induced/spontaneous fetal death at 6 or fewer months of gestation), stillbirth (fetal death at 7+ months of gestation), neonatal death (death before the first 28th day of life (0 – 27 days)), adverse pregnancy outcomes (defined as abortions/miscarriages or stillbirth), registered pregnancy (defined as a pregnancy with a known outcome i.e., livebirth, miscarriage, abortion, stillbirth and neonatal death) and survival beyond the neonatal period (survival beyond 28 days: post-neonatal survivors). Woman and interviewer characteristics are described in the supplementary document (Appendix 1: Table S1).

All quantitative analyses were undertaken using Stata (version 16.1; StataCorp, College Station, TX, USA) and RStudio (version 1.2.5033, RStudio, Boston, MA, USA).

Methods by objective

Objective 1: Review data collection processes and methods for pregnancy and birth outcomes surveillance for the five HDSS.

Before implementation of the EN-INDEPTH study, representatives from the sites were invited to the EN-INDEPTH metrics protocol design workshop held in Kampala in June 2016, during which we conducted an assessment of the maternal and newborn indicators and the HDSS systems' readiness for pregnancy surveillance using a standard template (Appendix 1: Table S2).[36] During this meeting small group discussions were undertaken to further discuss challenges faced in the capture of pregnancy outcomes within the HDSS, and to develop a proposal for enhancements to address one or more of these. Information submitted during the workshop was summarised by different themes (site information, data capture process, availability of outcome indicators, enhancements, and challenges affecting reporting of pregnancies). Sites shared tools and questionnaires used during pregnancy surveillance. Next, we reviewed and summarised information from the HDSS tools and questionnaires and their profiles on the INDEPTH Network website.[51] Following this, all site teams reviewed the draft results and provided further input and clarifications as required throughout the study implementation of this study.

The comparability of data collection processes was assessed by the HDSS, LSHTM, and MakSPH teams.

Objective 2: Compare SBRs and NMRs estimated from the two arms of the EN-INDEPTH survey (FBH+/ FPH) and the routine HDSS data for four sites with comparable data (excluding Bandim).

A list of study IDs and survey dates for women within four comparable sites (Dabat, IgangaMayuge, Matlab and Kintampo) who had been continuously resident within the HDSS since January 1 2012 and were interviewed in the EN-INDEPTH survey was created. Women were classified as continuously resident based on HDSS data. Survey and HDSS datasets for each site comprising information on all births and pregnancies amongst continuous residents that ended between January 1 2012 and the date of the women's EN-INDEPTH survey interview were compiled. HDSS datasets were compiled by the sites' data scientists basing on the women's study ID and survey date. HDSS data were extracted using a standard abstraction table in Excel (Supplementary document, Appendix 2). These HDSS data were then exported into Stata version 16.1 for summarising and later merged with the survey data for further data management and analysis.

The number of pregnancies, abortion/ miscarriages, stillbirths, livebirths, and neonatal deaths recorded in the EN-INDEPTH survey and HDSS were compared using descriptive statistical

methods. For each site for both the survey and HDSS data, we estimated the risk of stillbirth (SBR) and neonatal death (NMR) and calculated site-specific risk ratios (survey-SBR: HDSS-SBR and survey-NMR: HDSS-NMR) and their 95% confidence intervals (CIs). We combined the site-specific risk ratios using standard random-effects meta-analysis techniques.

Supplementary analysis

To understand the main results better, we undertook additional analyses, excluding data for site-years identified to have incomplete data years, and errors based on findings from objective one and extreme yearly differences in the number of outcomes compared between HDSS and survey (Appendix 1: Table S3).

In comparing reporting between Survey and HDSS for both primary and supplementary analysis, we treated the Survey and HDSS as independent data sources.

RESULTS

Objective 1: Data collection methods and processes for pregnancies and birth outcomes for the five HDSS

Site information

The five HDSSs were established at different points in time. Matlab, the oldest site, began in 1966, and IgangaMayuge, the youngest site, began in 2004. All sites are predominantly rural except Bandim and IgangaMayuge, with urban and peri-urban areas, respectively. The three-to-five-year population average and average number of households for the five sites ranged from 69,000 to 230,000 people and 16,000 to 53,000 households respectively (Table 7.1). Estimates for Total Fertility Rate (TFR)), birth rate and the out and in-migration per 1,000 population varied for sites with available data (Table 7.1).

Data capture processes

A unique identifier is assigned to livebirths in all sites, and this ID is linked to the mother. Only the Matlab and Bandim sites can partially link their HDSS data with health facility data. All sites except Matlab primarily collect data using a paper-based system, although, in recent years, efforts to use electronic data systems have been made in other sites. Matlab uses 100% female HDSS enumerators, and all other sites deploy both male and female enumerators (Table 7.1).

Pregnancy and outcome identification and surveillance visits vary from monthly (Bandim, urban) to once every six months or annual (Dabat, IgangaMayuge, and Kintampo). Dabat,

IgangaMayuge, and Kintampo had a fluctuating number of surveillance visits recorded per year for the period 2012 to 2018. Reasons for this included focusing on understanding data loss associated with migration from Microsoft Foxpro to Open-source Health and Demographic Surveillance System (OpenHDS system) and Open Data Kit (Dabat in 2017 and Kintampo in 2018) and funding and administrative and logistical challenges (IgangaMayuge and Kintampo) (Table 7.1).

In Dabat, IgangaMayuge, and Kintampo is any adult member (the main respondent) in the household can give an update on pregnancies and outcomes during surveillance visits if the woman is not present, whereas, in Matlab, the woman is the main respondent on issues of pregnancy and outcomes. If a woman died, the information about her death is obtained from the verbal autopsies which are conducted with the main caretaker (Table 7.1).

All sites except Bandim reported that children who were born and died between surveillance rounds are retrospectively recorded in the next surveillance round. Bandim's system thus differs from other sites. Pregnancies and children are followed prospectively only, and no retrospective information on children who were born and died between rounds is collected unless the pregnancy was registered.

Community informants/scouts/guides were used to identify pregnancies in all sites except Matlab. In Matlab, to further improve the quality of reconciliation between pregnancy outcomes that occur in-between surveillance rounds, the site keeps a household record register within the household; this register is reviewed and updated at each household visit by the HDSS enumerator. Furthermore, if a woman of reproductive age reports missing her last menstrual period, a urinary pregnancy test is performed to confirm the pregnancy if the status is unknown (Table 7.1).

Data on outcome indicators

All sites gathered data on stillbirths, neonatal deaths, birthweight, and gestational age. Data collection for abortion and miscarriages vary by site. Kintampo records only information for abortion and IgangaMayuge gathers data on miscarriages only. Bandim, with its recoding of outcomes of prospectively recorded pregnancies is likely to capture only miscarriages. In contrast, Dabat and Matlab collect and record information on both miscarriages and induced abortion. Only Matlab ascertains whether births are registered with the national CRVS for all births that happened within the HDSS until the second birthday. In Bandim this has been investigated for a sample, where only 16% of mothers of children aged two years or younger

in urban Bissau had received a birth certificate.[54] Apart from Dabat, all sites gather data on whether a death certificate for a recorded death was issued. Cause of death information is collected using verbal autopsies with the cause of death assigned by a physician in Bandim, Dabat, IgangaMayuge, Matlab, and Kintampo. Furthermore, in Matlab, International Classification of Diseases (ICD-10) codes are assigned to the cause of death (Table S4).

Enhancements and challenges affecting reporting and data quality of pregnancies.

All sites except Matlab suggested potential enhancements for improving capture and quality of data on pregnancies and their outcomes, such as improved training for field workers, standardising definitions, pregnancy testing, health campaigns, and understanding the socio-cultural norms that hinder reporting of adverse pregnancy outcomes (Table S4).

All sites reported misclassification between stillbirths and early neonatal deaths, missing data (except Matlab), and information on some key indicators as challenges that they face when comparing their HDSS data with other data capture systems (surveys and other demographic surveillance systems) (Table S4).

Objective 2: Population-level comparison of SBR and NMR rates between survey and HDSS

Bandim was excluded from population-level comparisons as HDSS event counts were not comparable with survey data due to differences in pregnancy surveillance processes. The following analyses include 81.2% (n=48,465) of women participating in the EN-INDEPTH survey from four sites who were continuous HDSS residents between January 1, 2012, and the survey interview date (Table 7.2 and Figure 7.1). Continuously resident women were well balanced between FBH+ and FPH groups for each site (Table 7.2). The distributions of pregnancy outcomes, interviewer, and woman characteristics for continuously resident women, varied across sites. In the survey Matlab and Dabat deployed all-female interviewers, a 50:50 distribution in IgangaMayuge, and predominantly male interviewers in Kintampo (Table 7.2-7.3).

Overall pregnancies and pregnancy outcomes

Amongst continuously resident women, 34,743 (72%) had at least one pregnancy recorded since January 1, 2012, in the HDSS and 37,258 (77%) in the survey (Table 7.4). 45,107 and 53,191 registered pregnancies since January 1, 2012, were recorded in the HDSS and survey, respectively. The EN-INDEPTH survey recorded more pregnancy outcomes than the HDSS data irrespective of survey module, except in Matlab where the HDSS recorded more pregnancies and outcomes

(Table 7.4). In the three sites where the survey recorded substantially more pregnancies than the HDSS, the survey unsurprisingly recorded more stillbirths and more neonatal deaths. However, in Matlab, where very similar numbers of pregnancies were recorded in both the HDSS and the survey in the FBH arm, the FPH arm potentially missed adverse pregnancy outcomes (FBH+: 16% in HDSS vs 14% in survey and FPH: 16% in HDSS vs 7% in survey) Table 7.4.

Neonatal deaths

From January 1, 2012, to the date of the survey, 40,370 livebirths in HDSS and 49,279 livebirths in survey were recorded, with 834 (2.1%) neonatal deaths were reported in the HDSS and 1188 (2.4%) in the survey (Table 7.4 and Figure 7.2 - 7.3). The estimated NMR ranged from 22.1 to 31.1 per 1,000 livebirths for survey data and from 16.3 to 27.8 per 1,000 livebirths for HDSS data among women in the FBH+ arm of the survey (Figure 7.4). Among women in the FPH arm, NMR ranged from 22.8 to 25.6 per 1000 livebirths for survey data and ranged from 18.6 to 22.7 per 1000 livebirths for HDSS data (Figure 7.4). Overall, the survey estimated a higher NMR than the HDSS surveillance systems, with differences of 19% (95% CI (-2% – 43%), p=0.071) among women in the FBH+ group and 15% (95% CI (1% - 30%), p=0.030) among women who received the FPH module. There was no strong evidence of heterogeneity between sites for the HDSS: survey NMR ratio in either the FBH+ (p=0.121, I^2 =48.4%) or FPH group (p=0.975, I^2= 0.0%) (Figure 3a). Annual estimates for NMR were generally higher in the survey compared to the HDSS (Figure S1a).

Stillbirths

From January 1, 2012, to the date of the survey, 40,977 total births in HDSS and 50,008 total births in the survey were recorded with 607 (1.4%) stillbirths were reported in the HDSS and 729 (1.5%) in the survey (Table 7.4 and Figure 7.2 – 7.3). The estimated SBR ranged from 7.2 to 17.0 per 1,000 total births for survey data and from 3.9 to 18.8 per 1,000 total births for HDSS data among women in the FBH+ arm (Figure 7.5). In the FPH arm, SBR ranged from 10.2 to 20.5 per 1000 total births for survey data and from 5.5 to 21.6 per 1000 total births for HDSS data (Figure 7.5). In both survey and HDSS data, SBR estimates from FPH were higher than FBH+ in all sites except Matlab (Table 3). Overall, the survey recorded a slightly higher SBR than the HDSS surveillance systems: 13% (95% CI (-21% – 63%), p=0.519) among women

in the FBH+ arm 20% (95% CI (-24% - 90%), p=0.444) among women in the FPH arm. There was evidence of heterogeneity between sites in both the FBH+ arm (p=0.024, I^2=68.2%) and FPH arm (p<0.001, I^2=85.2%), with a higher reported ratio of rates in Dabat in the FBH+ group and IgangaMayuge in both groups with the survey, but little difference in other sites (Figure 3b). Annual estimates for SBR were generally similar between the survey the HDSS data (Figure S1b).

Miscarriages and abortions

Some pregnancies in the HDSS and the survey resulted into miscarriages (HDSS: 2,631 (5.8%) and survey: 3,006 (5.8%)) and abortions (HDSS: 1,165 (2.6%) and survey: 177 (0.3%)). In two of the three sites providing information on miscarriages, more were captured using the survey than HDSS, irrespective of the module (Table 7.4 and Figures 7.2 and 7.3). However, in Matlab, a difference was observed between the FBH+ and FPH arms, in the FBH+ arm, the survey recorded more miscarriages (1436 (10.5%) in the survey vs 1318 (9.4%) in HDSS), whereas the HDSS recorded more miscarriages in the FPH arm, 368 (2.9%) miscarriages in the survey, compared to 1,300 (9.4%) in the HDSS. No information on miscarriages was available from Kintampo HDSS data (Table 7.4).

Additional supplementary analysis excluding pregnancy events that occurred in site-years (four years in Dabat; two years in IgangaMayuge and Matlab; and three years in Kintampo) with identified data collection issues found similar results (Tables S5a and S5b; Figures S2a-to-S3b).

Table 7. 1: Overview of HDSS in five EN-INDEPTH study sites

	Bandim	Dabat	IgangaMayuge	Matlab	Kintampo
Site Information					
Country	Guinea-Bissau	**Ethiopia**	Uganda	Bangladesh	Ghana
HDSS start year	1978	**1996**	2004	1966	1994
Area – km^2			155	184	7,162
Location	Guinea-Bissau, covering rural and urban	Gondor, Amhara regional state, 821 km northwest of Addis Ababa and 75km north	Iganga and Mayuge districts, approximately, 120km east of capital, Kampala along Kenya-Uganda highway	Matlab Upazila, in Chandpur district, 55km southeast of capital, Dhaka	Brong Ahafo region
Population (Households)	180,000 (36,000)	69,000 (16,000)	83,000 (16,000)	230,000 (53,000)	153,000 (32,000)
Migration rate (out)	Not Applicable	30.8 per 1,000 population (2012)	Not Applicable	2012-2018: ~58 per 1,000 person-year	27.3 per 1,000 population (2016)
Migration rate (in)	Not Applicable	32.6 per 1,000 population (2012)	Not Applicable	2012-2018: ~52 per 1,000 person-year	27.9 per 1,000 population (2016)
Birth rate	21.9 per 1,000 population (2017)	28.2 per 1,000 population (2012)	27.7 per 1,000 population (2017)	7 years CBR: ~22/1,000 population	20.7 per 1,000 population (2016)
Total Fertility Rate	Not available	3.8	4.3	2.6	4.1
Pregnancy identification					
Main respondent on pregnancies and outcomes	Not Applicable	Any adult member in household. If woman is not present.	Any adult member in household. If woman is not present.	Woman is the main respondent about pregnancy related issues and outcomes. If the woman died, then information about her death is obtained from Verbal Autopsy. Matlab records <10 pregnancy related death per year.	Any adult member in household. If woman is not present. Or Community Key Informant
Temporary absenteeism of women during data collection	Not Applicable	Yes – revisits are conducted up to four months. Percent of absenteeism is not computed	Yes - Revisits are conducted up to 4-6 months Percent of absenteeism is not computed	Yes – Approximately 20% of women are absent during data collection. Revisits are conducted up to 4 months or more	Yes – Approximately 25% of women are absent during data collection. Revisits are conducted to the households on the next round.
Methods for pregnancy identification	Recently started using community key informants in a sample of rural villages. Each woman as asked about her pregnancy status and a pregnancy ID is given if found pregnant. Both male and female informants are used	Local guides report pregnancies and births within 48 hours after they have been identified. 83% of the informants are female	Has 64 Community based "scouts" and Village Health Teams. Both male and female scouts are used	The women age 15-49 years old missing last menstruation are offered a urine test for pregnancy if pregnancy status is unknown. 100% of the enumerators are female	Community key informants are used to notify pregnancies. Each woman is asked about her pregnancy status. 5% data on household data generated are sampled and sent back for supervisors to verify and confirm all events captured by the data collector.
Data capture process					
Frequency of re-census the area	Every 2-4 years (Urban), each update round (rural)	Every 7 years	Each update round (Twice a year)	8 years or more	Last census 2003

	Bandim	Dabat	IgangaMayuge	Matlab	Kintampo
Average frequency of surveillance rounds	Urban: Monthly Rural: 2/year	2/year	2/year	6/year	1/year
Number of surveillance rounds since 2012	Not Applicable	2012: 2 rounds 2013: 2 rounds 2014: 2 rounds 2015: 2 rounds 2016: 2 rounds 2017: 1 round 2018: 1 round	2012: 2 rounds 2013: 1 round 2014: 1 round 2015: 1 round 2016: 2 rounds 2017: 1 round 2018: 1 round	2012: 6 rounds 2013: 6 rounds 2014: 6 rounds 2015: 6 rounds 2016: 6 rounds 2017: 6 rounds 2018: 6 rounds	2012: 3 rounds 2013: 3 rounds 2014: 2 rounds 2015: 2 rounds 2016: 1 round 2017: 1 round 2018: 1 round
Reasons for changes in number of surveillance rounds	Not Applicable	One round of data collection in 2017 due to data migration to Open-HDS. One round in 2018 due to errors encountered in linking data from Open Data Kit (ODK) to open-HDSS database. Data pre-2015 and post-2015 cannot be linked as well	Administrative and logistical bureaucracies and delays Funding challenges. When funds were not secured early enough, this would delay start of the next field round Kintampo migrated their Foxpro-based Household registration System (HRS2) to OpenHDS and Open Data Kit	Not Applicable	Funding challenges. When funds were not secured early enough, this would delay start of the next field round
Reconciliation of pregnancy outcomes that occur in-between surveillance rounds	For all registered pregnancies, a birth/end of pregnancy form is filled out	Pregnancy observation history form will be filled retrospectively, and livebirth and pregnancy loss is registered. For neonatal deaths a verbal autopsy interview is conducted, however, this is not linked back to main HDSS table except if livebirth was previously registered. Some pregnancies, outcomes and children are not captured by the HDSS system. For stillbirths and TOP, no cross-checking or validation is conducted	Livebirth and pregnancy losses are is retrospectively registered in the subsequent round. For neonatal deaths a verbal autopsy interview is conducted, however, this is not linked back to main HDSS table except if livebirth was previously registered. Some pregnancies, outcomes and children are not captured by the HDSS system. Sometimes a child is registered but their pregnancy is not. For stillbirths and TOP, no cross-checking or validation is conducted	Birth outcome is recorded and the HDSS keeps a household register within each household. The register is review and updated at each visit	The pregnancy is retrospectively registered, after which a pregnancy termination form will be completed together with a birth form if the woman is a resident of the HDSS
Pregnancy surveillance and outcome data linkage within HDSS database	Assign unique individual IDs to all livebirths. These are linked to residency tables and mother IDs	Assign unique individual IDs to all livebirths. Mother IDs are linked to non-live birth pregnancy outcome. Non-live birth	Assign unique individual IDs to all livebirths. These are linked to both location and mother IDs. Mother ID are linked to non-live	All birth outcomes are linked to the mother's ID	Assign unique individual IDs to all livebirths. These are linked to both compound and mother IDs. Mother ID are linked to non-live

	Bandim	Dabat	IgangaMayuge	Matlab	Kintampo
		outcomes are not given IDs, also death of post-neonatal survivor not previously registered is not captured in death table even if a verbal autopsy is conducted	birth pregnancy outcome. Non-live birth outcomes are not given IDs, also death of post-neonatal survivor not previously registered is not captured in death table even if a verbal autopsy is conducted		birth pregnancy outcome. Non-live birth outcomes are not given IDs, also death of post-neonatal survivor not previously registered is not captured in death table even if a verbal autopsy is conducted
Facility births	Urban: 65% Rural: 39%	17%	64%	68.9% (Intervention area: 86.4%; Comparison Area: 49.2%)[1]	61%
Links to facility data	In national hospital, not in rural	Pilot study ongoing	Not currently	Matlab hospital only (17% of births). The hospital is in intervention area.	No
Data collection	Paper based	Piloting electronic-based since mid-2017	Paper-based	Electronic (Galaxy Tab)	Paper-based but piloted electronic-based in mid-2018
Data migration from one system to another between 2012 and 2018	Not Applicable	Yes – In 2017, HDSS data migrated from paper-based (HRS2-microsoft FoxPro) to tablet-based data collection system (MySQL, OpenHDS suite)	No	No	Yes - In June 2018, HDSS data migrated from paper-based (HRS2-microsoft FoxPro) to tablet-based data collection system (MySQL, OpenHDS suite)
Data collection supervision	Yes	Yes – Three-personnel supervise on a bi-weekly basis	Yes – Field manager, five team leader and quality assurance officer supervise during the data collection	Yes – Field manager and supervisors perform the supervision. Supervisors hold monthly meeting with data collectors on data quality check Supervisors conduct field supervision visits on random days HDSS has ±5% re-visit depending on the performances of the data collectors, Data collectors' yearly performance evaluation based on data quality Supervisors report to the field manager every month.	Yes. Routine supervision is done by the field supervisors and coordinators
Previous experience with Full Birth History (FBH) or Full Pregnancy History (FPH)	Both	Both	Both	Only FBH	Both

[1] icddr,b (2016) Health and Demographic Surveillance System–Matlab, v. 49. Registration of health and demographic events 2014, Scientific Report No.133. Dhaka: icddr,b.

[2]Fisker, A. B., et al. (2019). "Differences in barriers to birth and death registration in Guinea-Bissau: implications for monitoring national and global health objectives." Trop Med Int Health 24(2): 166-174

Table 7. 2: Characteristics of continuously resident women included in the EN-INDEPTH survey, for four HDSS sites

	Dabat	IgangaMayuge	Matlab	Kintampo	Total
	n=11438	n=9403	n=21432	n=6192	n=48465
	n (%)	n (%)	n (%)	n (%)	
Questionnaire module assigned					
FBH+	5752 (50.3)	4725 (50.3)	10796 (50.4)	3089 (49.9)	24362 (50.3)
FPH	5686 (49.7)	4678 (49.8)	10636 (49.6)	3103 (50.1)	24103 (49.7)
Mother and baby characteristics					
Age					
15 – 19	2823 (24.7)	3176 (33.8)	1452 (6.8)	255 (4.1)	7706 (15.9)
20 – 24	1601 (14.0)	1446 (15.4)	5689 (26.5)	1195 (19.3)	9931 (20.5)
25 – 29	1758 (15.4)	870 (9.3)	6280 (29.3)	1684 (27.2)	10592 (21.9)
30 – 34	1510 (13.2)	961 (10.2)	4760 (22.2)	1555 (25.1)	8786 (18.1)
35+	3746 (32.8)	2943 (31.3)	3247 (15.2)	1498 (24.2)	11434 (23.6)
Missing	0 (0.0)	7 (0.1)	4 (0.0)	5 (0.1)	16 (0.0)
Education level					
No education**	5595 (48.9)	731 (7.8)	978 (4.6)	2148 (34.7)	9452 (19.5)
Primary only	3088 (27.0)	4346 (46.2)	4062 (19.0)	2994 (48.4)	14490 (29.9)
Primary & Secondary	1640 (14.3)	3712 (39.5)	13486 (62.9)	949 (15.3)	19787 (40.8)
Higher	1115 (9.8)	609 (6.5)	2904 (13.6)	97 (1.6)	4725 (9.7)
Missing	0 (0.0)	5 (0.0)	2 (0.0)	4 (0.1)	11 (0.0)
Socioeconomic status					
Poorest	2659 (23.3)	1999 (21.3)	4368 (20.4)	1151 (18.6)	10177 (21)
2	2028 (17.7)	2016 (21.4)	4251 (19.8)	1185 (19.1)	9480 (19.6)
3	2332 (20.4)	2042 (21.7)	4243 (19.8)	1193 (19.3)	9810 (20.2)
4	2252 (19.7)	1758 (18.7)	4293 (20.0)	1276 (20.6)	9579 (19.8)
Richest	2167 (19.0)	1588 (16.9)	4277 (20.0)	1387 (22.4)	9419 (19.4)
Parity as reported in the survey					
0	3369 (29.5)	3990 (42.4)	379 (1.8)	17 (0.3)	7755 (16.0)
1	1316 (11.5)	689 (7.3)	6439 (30.0)	1188 (19.2)	9632 (19.9)
2	1151 (10.1)	398 (4.2)	7479 (34.9)	1414 (22.8)	10442 (21.5)
3	952 (8.3)	472 (5.0)	4649 (21.7)	1204 (19.4)	7277 (15.0)
4	1084 (9.5)	584 (6.2)	1780 (8.3)	1204 (19.4	4652 (9.6)
5+	3566 (31.2)	3270 (34.8)	706 (3.3)	911 (14.7)	8453 (17.4)

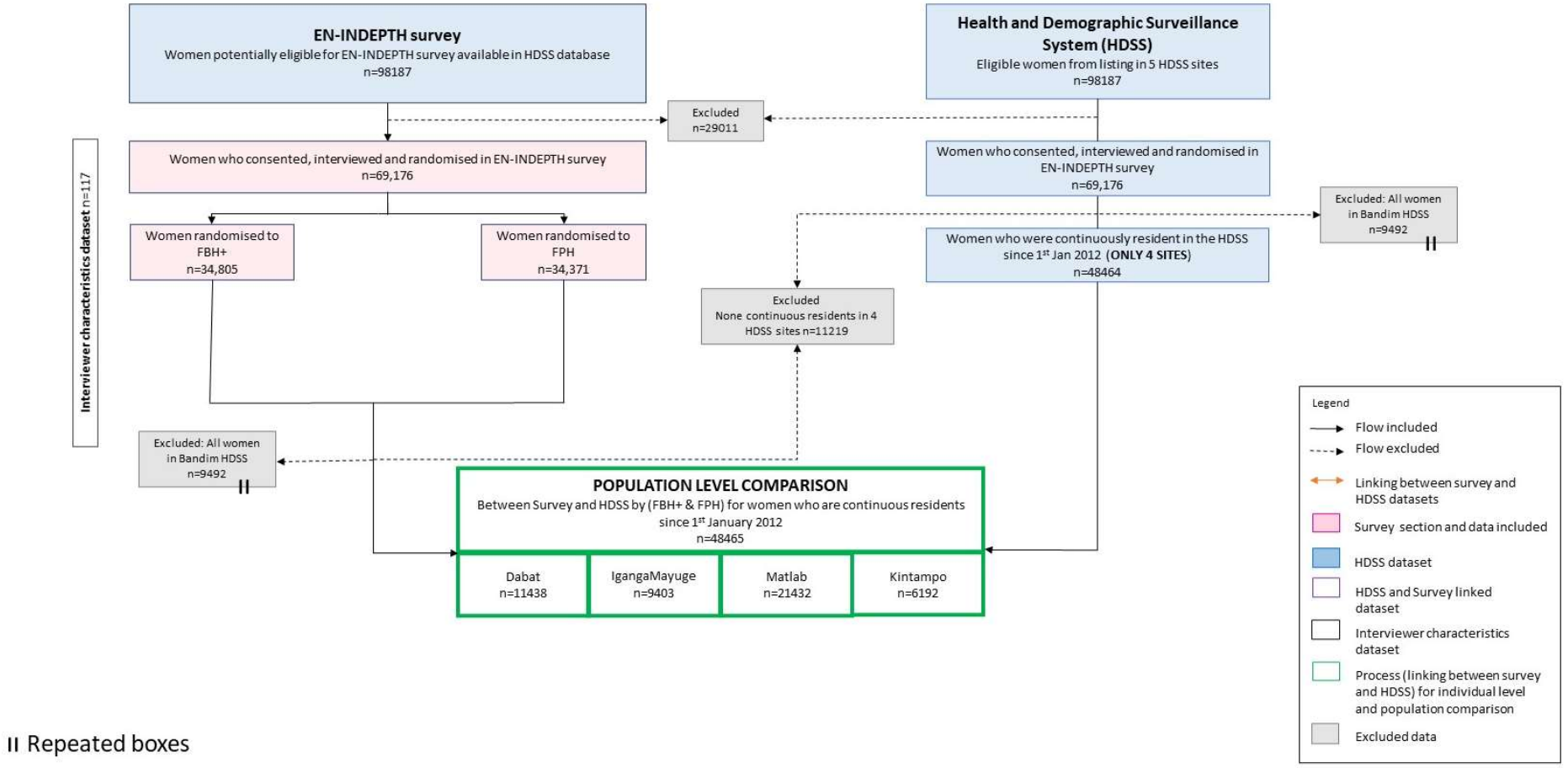
**This information was obtained from the EN-INDEPTH survey dataset. Bandim HDSS is excluded from this analysis; **never attended school or madrasa;*

Table 7. 3: Characteristics of interviewers who participated in the EN-INDEPTH survey, by sites, for four HDSS sites

Interviewer characteristics	Dabat	IgangaMayuge	Matlab	Kintampo	Total
	n=41	n=20	n=20	n=22	n=103
Gender	n (%)	n (%)	n (%)	n (%)	n (%)
Male	0 (0.0)	10 (50.0)	0 (0.0)	19 (86.4)	29 (28.2)
Female	41 (100.0)	10 (50.0)	20 (100.0)	3 (13.6)	74 (71.8)
Age					
<30	29 (70.7)	6 (30.0)	9 (45.0)	6 (27.3)	50 (48.5)
30+	12 (29.3)	14 (70.0)	11 (55.0)	16 (72.7)	53 (51.5)
Previous experience with the DHS survey					
Yes	36 (87.8)	19 (95.0)	20 (100.0)	13 (59.1)	88 (85.4)
No	5 (12.2)	1 (5.0)	0 (0.0)	9 (40.9)	15 (14.6)
Previous Experience with other surveys					
Yes	17 (41.5)	16 (80.0)	16 (72.7)	14 (70.0)	63 (61.2)
No	24 (58.5)	4 (20.0)	6 (27.3)	6 (30.0)	40 (38.8)

****This information was obtained from the EN-INDEPTH interviewer survey dataset. This was a self-administered questionnaire survey.***

Figure 7. 1: Linkage of pregnancy episodes captured by HDSS and EN-IDEPTH survey for four HDSS sites



II Repeated boxes

**** Bandim HDSS was excluded***

Table 7. 4: Comparison of pregnancy outcome indicators between EN-INDEPTH survey and HDSS by module

	Dabat++				IgangaMayuge++				Matlab++				Kintampo++			
	FBH+		FPH		FBH+		FPH		FBH+		FPH		FBH+		FPH	
No. of women interviewed in EN-INDEPTH	6327		6266		6788		6649		10809		10653		6049		6143	
No. women who were continuously resident in the HDSS*	5752		5686		4725		4677		10796		10636		3089		3103	
% of women interviewed who were continuously resident	90.9		90.7		69.6		70.3		99.9		99.8		51.1		50.5	
	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey
No. women with 1>= pregnancies ending after 1st January 2012*	2322	3083	2282	3103	1494	2125	1527	2176	10796	10559	10363	10067	2976	3068	2983	3077
Total pregnancies recorded*	2619	4769	2582	4770	2342	3260	2381	3340	14092	13713	13880	12495	3592	5391	3619	5453
Total live births recorded*#	2411 (92.1)	4648 (97.5)	2375 (92.0)	4644 (97.4)	2330 (99.5)	3051 (93.6)	2363 (99.2)	3039 (91.0)	12046 (85.5)	12068 (88.0)	11866 (85.5)	11862 (94.9)	3492 (97.2)	5001 (92.8)	3487 (96.4)	4966 (91.1)
Number of post-neonatal survivors	2344	4524	2321	4525	2277	2956	2315	2968	11801	11801	11621	11586	3435	4878	3422	4853
% of pregnancies that resulted into adverse pregnancy outcomes¬	3.1	5.1	3.2	5.1	2.8	9.3	2.8	11.1	16.3	13.9	16.3	7.3	5.2	9.5	5.7	11.0
Number of stillbirths recorded*	11 (0.4)	43 (0.9)	27 (1.0)	48 (1.0)	9 (0.4)	22 (0.7)	13 (0.5)	62 (1.9)	211 (1.5)	209 (1.5)	192 (1.4)	168 (1.3)	67 (1.9)	73 (1.4)	77 (2.1)	104 (1.9)
Number of early neonatal deaths (0-6 days) recorded*	41	80	37	92	48	78	40	60	204	227	198	223	44	95	49	88
Number of late neonatal deaths (7-27 days) recorded*	26	44	17	27	5	17	8	11	41	40	47	53	13	28	16	25
Number of abortions recorded*	NA	NA	NA	1 (0.0)	NA	NA	NA	23 (0.7)	517 (3.7)	NA	522 (3.8)	97 (0.8)	61 (1.7)	NA	65 (1.8)	56 (1.0)
Number of miscarriages recorded*	4 (0.2)	78 (1.6)	1 (0.0)	77 (1.6)	3 (0.1)	187 (5.7)	5 (0.2)	216 (6.5)	1318 (9.4)	1436 (10.5)	1300 (9.4)	368 (2.9)	NA	317 (8.8)	NA	327 (6.0)
Stillbirth Rate per 1000 total births (95% Confidence Interval)	4.5 (1.9 ,7.2)	9.2 (5 ,13.3)	11.2 (2.7 ,19.8)	10.2 (2.9 ,17.5)	3.9 (-3.8 ,11.5)	7.2 (-3.4 ,17.7)	5.5 (-6.6 ,17.5)	20.0 (-2.5 ,42.5)	17.2 (5.5 ,28.9)	17.0 (4.2 ,29.8)	15.9 (2.3 ,29.6)	14.0 (0.1 ,27.8)	18.8 (-13 ,50.6)	14.4 (-10.6 ,39.4)	21.6 (-17.3 ,60.5)	20.5 (-13.2 ,54.3)
Neonatal mortality rate per 1000 live births (95% Confidence Interval)	27.8 (21.2 ,34.4)	26.7 (19.7 ,33.7)	22.7 (10.6 ,34.8)	25.6 (14.1 ,37.1)	22.8 (4.3 ,41.2)	31.1 (9.3 ,53)	20.3 (-2.8 ,43.4)	23.4 (-1.2 ,47.9)	20.3 (7.5 ,33.1)	22.1 (7.4 ,36.8)	20.7 (5 ,36.3)	23.3 (5.3 ,41.2)	16.3 (-13.6 ,46.3)	24.6 (-8.2 ,57.4)	18.6 (-17.9 ,55.2)	22.8 (-13.1 ,58.6)

** Among women who were continuously resident in the HDSS between 1 January 2012 and the survey interview date; # Number of live births includes (Number of post-neonatal survivors plus number of neonatal deaths);¬Adverse pregnancy outcomes refers to abortions, miscarriages, stillbirths, and neonatal deaths; ++ n(%) is what is reported, for percentages, total pregnancies is the denominator for the percentage*

Figure 7. 2: Distribution of pregnancy outcomes between EN-INDEPTH survey and HDSS by module

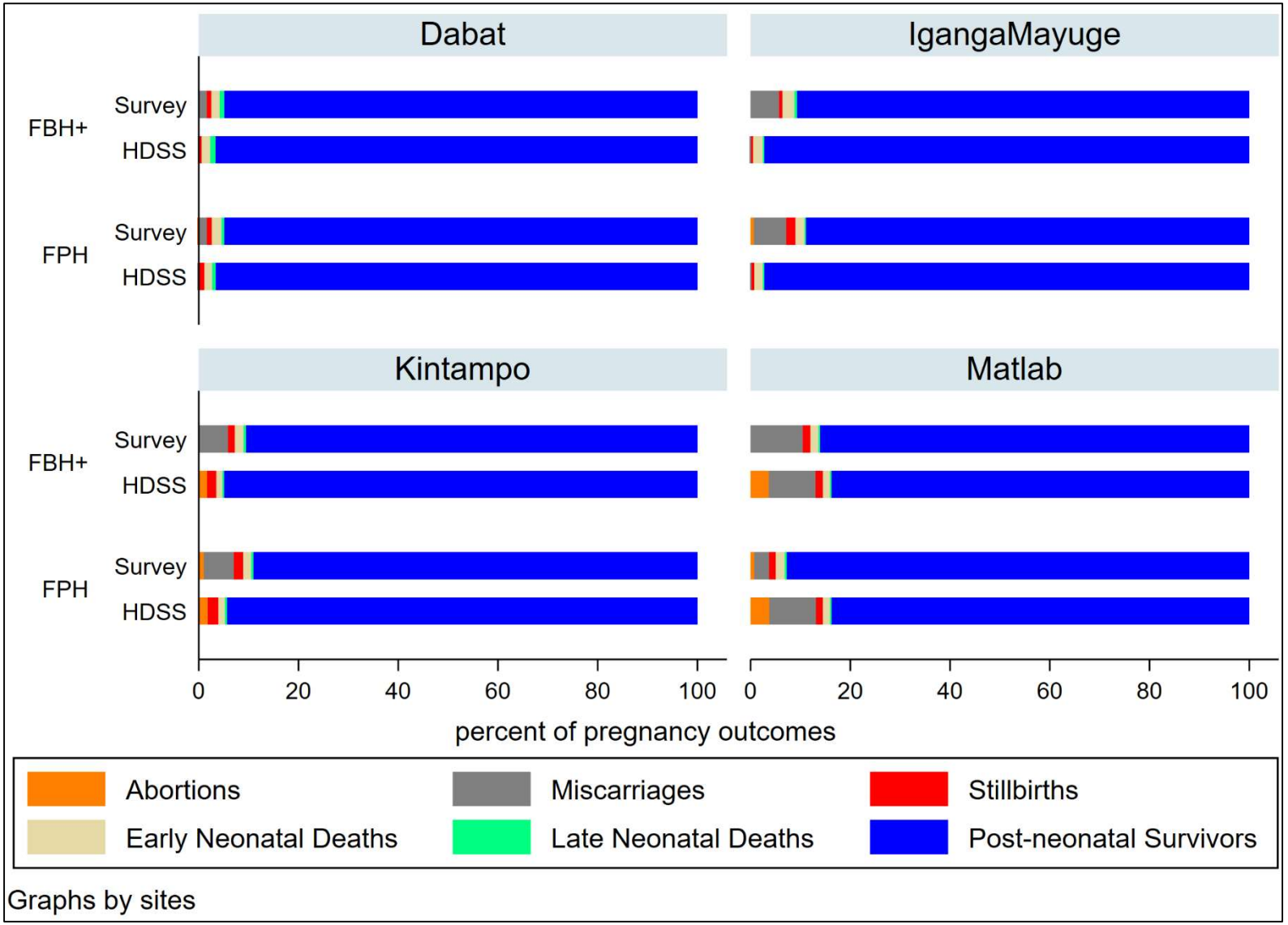


Figure 7. 3: Distribution of pregnancy outcomes between EN-INDEPTH survey and HDSS by module

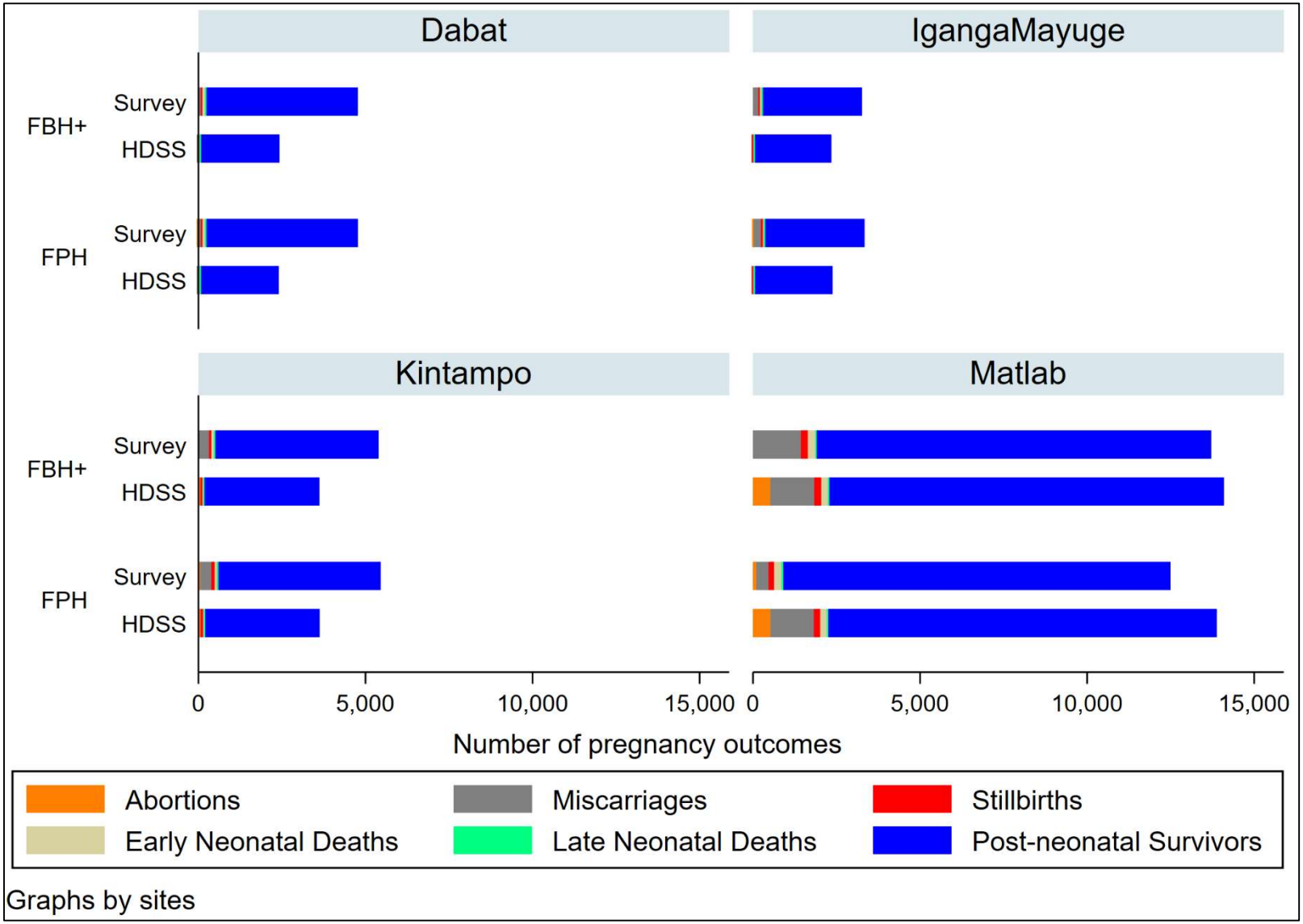
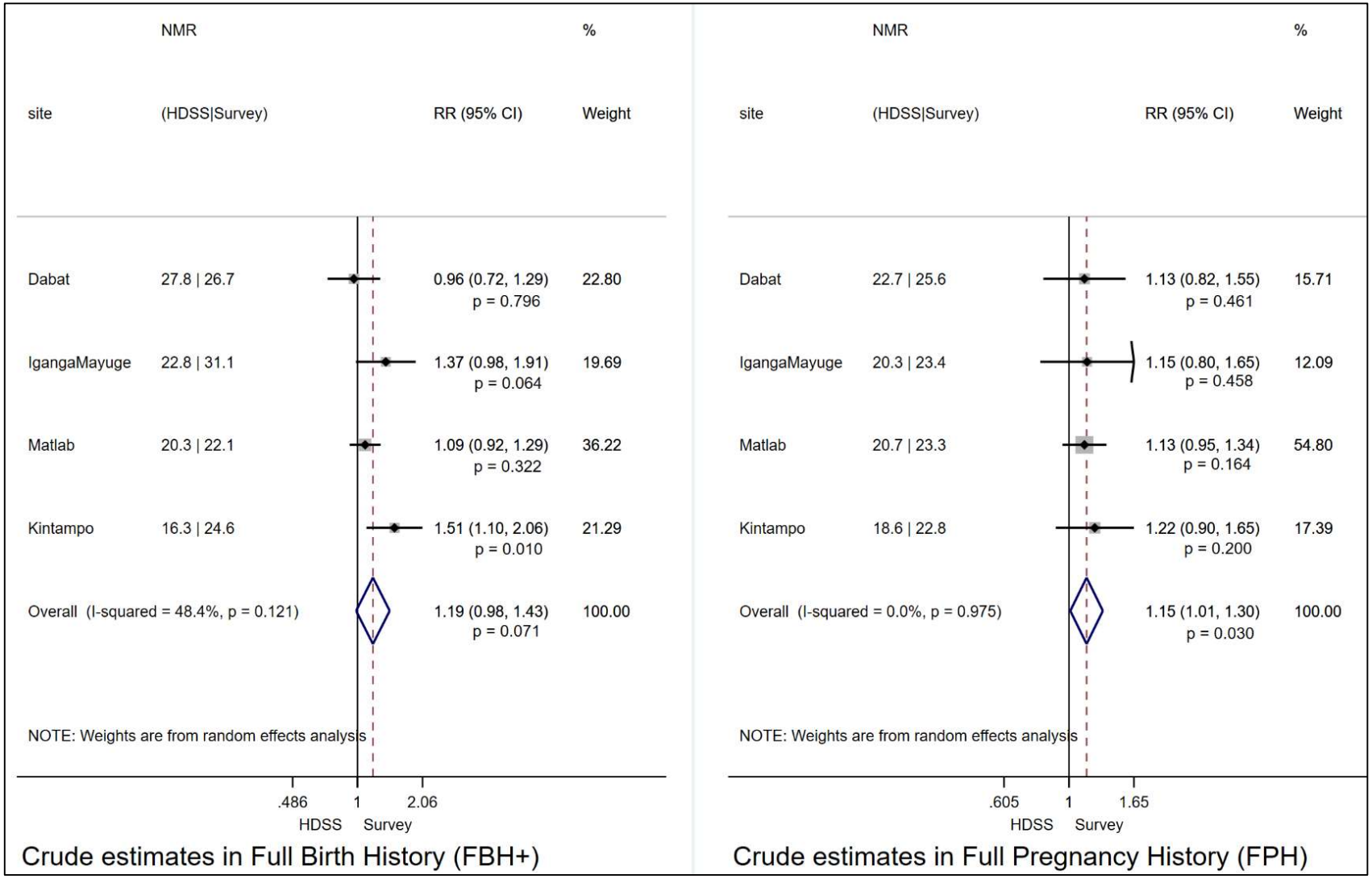
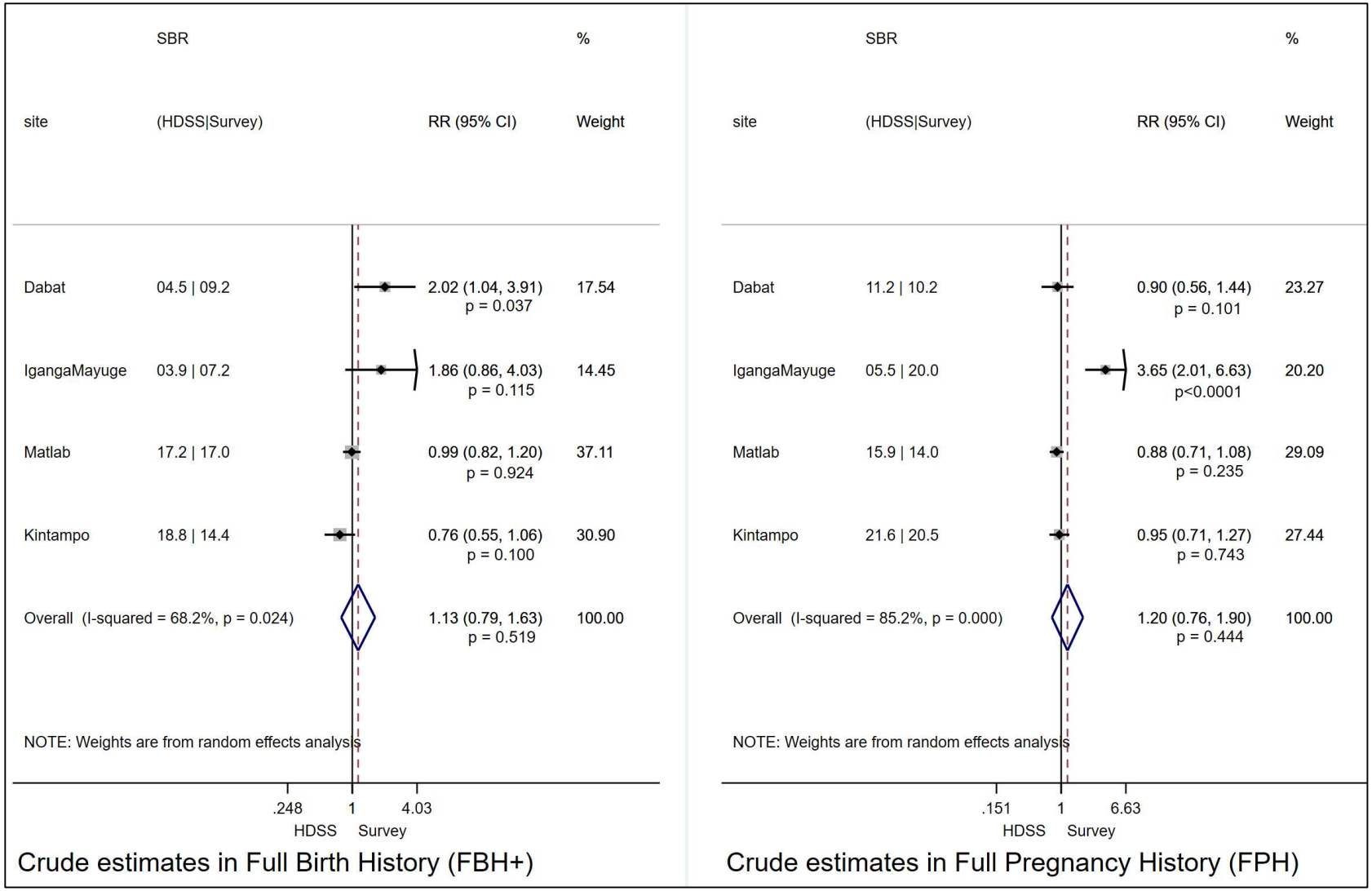


Figure 7. 4: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to HDSS



++ Used a simplified assumption that the HDSS and survey women are independent. Analysis was among women who are continuous residents since 1st January 2012

Figure 7. 5: Ratio of Stillbirth Rate in EN-INDEPTH survey compared to HDSS



++ Used assumption that the HDSS and survey women are independent. Analysis was among women who are continuous residents since 1st January 2012

DISCUSSION

This study determined that the five sites were established at different times with varying purposes including studying childhood outcomes, migration, malaria and hence have different processes and methods for pregnancy and outcome surveillance. These differences may account for some of the observed between-site variation in capture of pregnancies and adverse pregnancy outcomes.

The frequency of surveillance visits is likely to be an important factor in improving capture of pregnancies and adverse pregnancy outcomes. Matlab with six rounds per year consistently captured a similar number of pregnancies in the routine HDSS and survey arms, in contrast to other sites with only 1 or 2 rounds most years. This fluctuation in surveillance visits in Dabat, IgangaMayuge and Kintampo was largely attributed to administrative, logistical, and funding challenges. These fluctuations may have negatively affected the capture of pregnancy outcomes, especially the adverse outcomes and the data quality in these sites leading to many events occurring between surveillance rounds being missed in-between surveillance visits.

Matlab HDSS was established for reproductive health outcomes and therefore has always placed a strong emphasis on pregnancy surveillance, with a household record register kept in the household and updated at each surveillance round, routine pregnancy testing, and frequent surveillance rounds. These characteristics may have positively influenced data quality, resulting in a high degree of completeness in the recording of pregnancies and their outcomes. In contrast, Dabat, IgangaMayuge, and Kintampo have fewer surveillance rounds and place greater reliance on the retrospective recording of outcomes that happen in-between surveillance rounds. Retrospective registration is likely to capture a higher proportion of surviving livebirths but has the potential to miss adverse pregnancy outcomes.[1] The likelihood of an HDSS informant reporting previously unregistered children who died or adverse pregnancy outcomes in-between surveillance rounds may depend on the age of death of the child or gestational age of pregnancy at termination, thus a decreased probability for recording stillbirths and early deaths, e.g. children who die before their naming ceremony.[1]

NMR estimates were generally higher in the survey than the HDSS although these estimates were closer to each other irrespective of the survey module. This is consistent with the notion

that retrospective recording of death after a livebirth is likely to be better compared to adverse pregnancy losses. The HDSS in Dabat, IgangaMayuge and Kintampo appears to have missed substantial proportions of pregnancies compared to the survey and suggesting these HDSS data should not be considered to represent a high-quality standard dataset. The parent study by Akuze et al., reported that interviewers in Matlab were familiar with the FBH and had no prior experiences with the FPH.[50] Our study found that in Matlab, the survey missed a small proportion of pregnancy outcomes in the FBH arm which interviewers were familiar with, this did not result in an underestimate of the NMR or SBR. Conversely, in the FPH arm, there was an under recording of pregnancies which did not end in a livebirth, but not of pregnancies which did end in a livebirth. Thus, FPH slightly overestimated NMR but underestimated SBR relative to HDSS. We anticipate that interviewers' training and understanding of the procedures and process of implementing a FPH was a potential problem compared to the FBH+ because Matlab spent only 3.5 days training their interviewers compared to two-to-three weeks at the other sites.[50,55]

Matlab had extensive experience with electronic data collection, IgangaMayuge, and Bandim did not have prior experience and Dabat and Kintampo piloted electronic data collection during the implementation of the EN-INDEPTH study. A related study examining electronic data collection experiences in the five sites found that the number of days for training interviewers in the use of the Survey Solutions data collection system in the EN-INDEPTH study ranged from four to ten days. Matlab, which had more experience with electronic data collection, had fewer training days than other sites.[55] Electronic data collection allows for data linkage which may improve pregnancy surveillance and data quality because of minimal data entry, transition errors and available alternate data to validate pregnancy outcomes, for example Matlab's well-established electronic data collection system is also linked to other data sources (i.e. Matlab hospital data) .

Prospective surveillance systems are expected to yield more accurate estimates of pregnancy outcomes, SBR and NMR than surveys. However, our analyses found survey NMR estimates were generally higher than those based on HDSS data, except in Dabat. This suggest that HDSS may not always be the "high-quality standard dataset" that is sometimes assumed. In addition to the robustness of pregnancy surveillance and frequency of round, important contributors to data quality and completeness reported by sites in this study were misclassification

between stillbirths and early neonatal deaths, and missing information on key indicators were major challenges. Sites should make efforts to improve reporting of perinatal deaths as well as NMR. Other recent related studies linking pregnancies and pregnancy outcomes between the HDSS and survey in two sites that participated in the EN-INDEPTH survey found evidence of omission and misclassification between pregnancy outcomes in both the survey and the HDSS system (preliminary analysis results not shown).[53,56]

The gender distribution of interviewers was different across sites with all-female teams in both Dabat and Matlab, a 50:50 distribution in IgangaMayuge, and predominantly male interviewers in Kintampo. This gender balance may have affected the capture of data in the survey across all sites. Although not conclusive from our study, deployment of an all-female enumerator team during EN-INDEPTH survey and HDSS update rounds in Matlab may have improved data capture as mothers may be more likely to be open and disclose their pregnancies to fellow women. A related study by Kwesiga et al., found that women were more comfortable disclosing certain outcomes if the enumerator was female because they felt they would understand them better.[57]

Measurement implications

Differences in HDSS processes and methods for pregnancy surveillance complicates comparisons and interpretation of estimates computed from HDSS data. Standardisation of processes, definitions, tools and questionnaires, methods of probing for events and enumerator gender may help to improve the comparability of data on stillbirths across sites.[50] Differences in data collection methods between the survey and HDSS, such as the process of probing, listing, and frequency of inquiry, could have affected the results and analysis in this study. Although standardising methods may improve comparisons between sites, this may not necessarily resolve all data quality issues relating to stillbirths and neonatal deaths. First steps to improve data quality could include introducing in-built pregnancy outcome related data checks in the HDSS system, such as checking reported stillbirth with gestational age, and conducting verbal autopsies on all reported stillbirths and neonatal deaths to reduce misclassification. Measures such as increasing the number of surveillances rounds per year, introducing innovations such as paper or mobile phone-based household pregnancy registers linked to free distribution of pregnancy testing have the potential to lead to improvements in capturing complete information on pregnancies, but require further investment., to ensure

correct measures of rates of adverse pregnancy outcomes, data should be analysed in ways, which account for them being incomplete.

In addition, differences in date reporting of pregnancy outcomes (i.e., date of birth and gestational age at birth) in both survey and HDSS may influence the outcome's classification, for example, an earlier birth date reported in the survey than actual or reported in HDSS could classify a neonatal death as a death outside the neonatal period. Similarly, gestational age at birth and birthweight are required for the correct classification of stillbirths and miscarriages.[58-60] These divergences need to be considered when interpreting the reported differences. However, gestational age and birthweight data are often of low quality in both surveys and surveillances systems.

Since different HDSS systems were set up with differing objectives and purposes, it is important to improve awareness of the differences in methods used, definitions of stillbirths and neonatal deaths and contexts in the different sites. A recent study by Waiswa and Akuze et al that reviewed 31 HDSS sites found that one site had a different stillbirths definition and therefore, its stillbirth data could not be directly compared with other sites with the standard stillbirth definition (fetal loss with a gestational age of >7 months).[1] Differences is methodology, definitions renders the available data incomparable.

Overall, our stillbirth and neonatal death estimates patterns from FBH+ and FPH are consistent with results from the main study's randomised comparison of FBH+ and FPH, where the FPH arm was found to higher SBR estimate compared to the FBH+ with between-site heterogeneity and small difference between NMR by study arm. However, our study's focus is the comparison was between the survey and the HDSS.[50]

Strengths and limitations of this research

Our study had several strengths. First, we compared stillbirths and neonatal deaths estimates across different maternity histories in four HDSSs with a large sample. To the best of our knowledge, our study is the first large multi-site, multi-country, and multi-module comparison of both HDSS and survey reporting and capture of stillbirths. Secondly, our survey was conducted within HDSS settings, and our analysis was limited to only women who were continuous residents within the sites for birth events recorded after January 1 2012 and before the date of the interview. We believe that this truncated period of recall reduces the

recall bias of reported events in the survey. Third, since we were able to correctly identify women who were interviewed in the survey based on study ID linked to HDSS identifiers, we compared the same group of women reporting pregnancy outcomes for the same period at a population-level between survey and HDSS. Finally, we conducted a formative review with input from sites representatives in addition to reviewing the HDSS profiles which give comprehensive insights into how the HDSS sites' pregnancy surveillance is conducted. Our comparison of the different systems gives further insights into understanding the differences and complexities between the pregnancy surveillance variations in different contexts.

Our study also had some limitations. In our population-level comparison of reporting between Survey and HDSS, we treated the Survey and HDSS as independent data sources. Although HDSS measurements could not have been affected by the survey measurements because the survey happened after the HDSS measurements, the fact that the survey respondents within the HDSS setting are subject to regular surveillance visits could have influenced women's responses to the survey positively because regular inquiry in the HDSS surveillance is likely to improve recall in the survey. Our comparison of FBH data from survey to HDSS (in Matlab if treated as a high-quality standard dataset) shows that the FBH performed very well, in a site where women are subjected to regular inquiry. This may not be the case in populations which are not under regular surveillance, and we need to be cautious in generalising our findings.[50] Lastly only 81% of the women in the original EN-INDEPTH survey were continuously resident in the HDSS and able to be included in this analysis, this may hinder the generalisability of findings to the entire populations especially amongst mobile populations

Our study identified several research gaps with respect to improving measurement of stillbirths and neonatal deaths in household surveys and HDSS. Examples of research gaps include (1) Assessment of methods known to improve identification and capture of pregnancy outcomes in both survey and HDSS to minimise omission (i.e., pregnancy testing to allow for early identification and using household registers). (2) Methods for assessing data quality and accuracy for stillbirths and neonatal deaths. (3) Data interoperability and linkage between HDSS and survey and with other sources of data on stillbirths and neonatal deaths (i.e., facility records, maternal cards/passports,). Closing these gaps will be important to improve data on stillbirths and neonatal deaths across these data platforms.

CONCLUSIONS

HDSS systems need to establish regular reporting and capture of pregnancies, especially stillbirths and neonatal deaths, to improve pregnancy surveillance. Although HDSS data is often considered the “high-quality standard” compared to survey data, our findings suggest this is not always the case for stillbirths and neonatal deaths for all HDSS systems, based on the survey approaches and processes they implement. Further work in understanding the different contexts and methodologies for pregnancy surveillance within and between sites is important for interpreting results from HDSS. Concluding regarding bias in HDSS estimates may not be wise without further investigation. HDSS sites need enhancements to strengthen their pregnancy surveillance systems to facilitate the earlier identification of pregnancies through innovations such as community pregnancy testing and choosing women as the main respondent about issues regarding pregnancies and outcomes.

LIST OF ABBREVIATIONS

DHS	Demographic and Health Survey
ENAP	Every Newborn Action Plan
EN-INDEPTH	Every Newborn-INDEPTH
FBH+	Full Birth History (+ denotes additional questions on pregnancy losses)
FGD	Focus Group Discussion
FPH	Full Pregnancy History
HDSS	Health and Demographic Surveillance System
ID	Identification Number
LMIC	Low- and Middle-Income Countries
MICS	Multiple Indicator Cluster Survey
SDG	Sustainable Development Goals
TOP	Termination of Pregnancy
UN	United Nations
VA	Verbal Autopsy
WHO	World Health Organization
OpenHDS	Open-source Health and Demographic Surveillance System

DECLARATIONS

Ethics and consent to participate.

The EN-INDEPTH study was granted ethical approval by the institutional review boards in all operating countries as well as from Institutional Ethical Review Committee of the London School of Hygiene & Tropical Medicine (see Annex 5).

Consent for publication

Not applicable.

Availability of data and material

Data sharing and transfer agreements were jointly developed and signed by all collaborating partners.

Competing interests

All authors completed unified competing interest forms and declare no conflicts of interest'

Funding

This work was funded by the Children's Investment Fund Foundation (CIFF) by means of a grant to LSHTM (PI JEL), and a sub-award to the INDEPTH MNCH working group with technical leadership by Makerere School of Public Health (PI PW).

Authors' contributions

The EN-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, PW, AB, JEL and SC developed the detailed research questions. JA HB and SC developed the overall analysis plan for this paper. 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. Analysis was undertaken by JA. SC provided statistical oversight. The manuscript was drafted by JA, HB, JEL and SC. All authors reviewed and helped to revise the manuscript. All authors reviewed and agreed the final version.*

Acknowledgements

We thank the 118 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 the site administrators staff for their administrative support.

We express appreciation to the EN-INDEPTH expert advisory group: Fred Arnold; Peter Byass; Trevor Croft; Kobus Herbst; Sunita Kishor; Florina Serbanescu; Stephane Helleringer; Turgay Unalan; Shane Khan; Attila Hancioglu.

We express appreciation to Judith Yargawa for coordinating the author and for her editorial services and Kimberly Peven for the support with some of the graphics.

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.

The Every Newborn-INDEPTH study collaborative author group:

Solomon Mokonnen Abebe; Joseph Akuze; Nurul Alam; Kwaku P Asante; Tadesse Awoke Ayele; Angela Baschieri; Gashaw Andargie Biks; Telake Azale Bisetegn; Hannah Blencowe; Simon Cousens; Tadess Guadu Delele; Nafisa Delwar; Francis Dzabeng; Seeba Amenga Etego; Yeetey AK Eyuameh; Ane B Fisker; Edward Galigawango; Kassahun Alemu Gelaye; Bisrat Misganaw Geremew; Lemma Derseh Gezie; Vladimir Gordeev; Collins Gyezaho; M Moinuddin Haider; Md Ali Iman; Judith Kaija; Dan Kajungu; Simon Kasasa; Doris Kwesiga; Joy E Lawn;

Kazuyo Machiyama; Kaiser Mahmud; Grace Manu; Alexander A Manu; Justiniano SD Martins; Tesfahun Melese; Mezgebu Yitayal Mengistu; Tryphena Nareeba; Davis Natukwatsa; Obed Ernest Nettey; Sam K Newton; Seth Owusu-Agyei; Victoria Ponce Hardy; Amabelia Rodrigues; Charlotte Tawiah; Adane Kebede Tesega; Sanne M Thuysen; Valerie Tusubira; Peter Waiswa; Temesgen Azmeraw Yitayew; Charles Zandoh.

Chapter 8: Discussion

8.1. Introduction

This thesis has investigated the various components of measurement of stillbirths and neonatal deaths in standardised population-based surveys in low and middle-income countries, which are often characterised by weak CRVS systems or poor data linkage or interoperability (the ability for databases to exchange and make use of information) between different data systems with data on stillbirths and neonatal deaths.

This Chapter discusses the main findings and methodological considerations about measurement of stillbirths and neonatal deaths in surveys, strengths and limitations, generalisability of findings, measurement changes and implications now and for future research. The PhD findings are presented in relation to the global health measurement agenda and other research. The synthesis and discussion include personal reflections.

The Chapter is structured by the objectives of the PhD presented in Chapter 2. The PhD is comprised of: 1) a historical review of Demographic and Health Surveys (Chapter 3); 2) Methods of the EN-INDEPTH study in (Chapter 4); 3) Randomised comparison of the FBH+ with additional questions on pregnancy losses and FPH (Chapter 5). 4) Further assessments of determinants of time taken for completion FBH+ and FPH survey modules and data quality using Paradata in Chapter 6; 4) and Population-level comparison of pregnancy outcomes between EN-INDEPTH survey and HDSS data (Chapter 7).

8.2. Overview of main findings

This research focused on improving our understanding of data capture for stillbirths and neonatal deaths in DHS surveys through reviewing previous practices and comparing two alternative approaches to inform future population survey methods.

Globally, there is a recent call to improve the measurement of stillbirths and neonatal deaths as countries strive to achieve the set SDG and ENAP targets before the year 2030.[61] Several factors, including survey processes, interviewer and respondent attributes, questionnaire and software designs, survey implementation, data management and analysis and quality of the data used to derive/compute the indicators, influence the overall SBR and NMR estimates computed from population-based surveys (Figure 8.1). These factors are interrelated, and our understanding of these different factors will go a long way in improving the capture and data

quality and measurement of stillbirths and neonatal deaths in surveys. In this PhD study, all aspects of survey methodology except survey processes were examined.

8.2.1 DHS Historical review

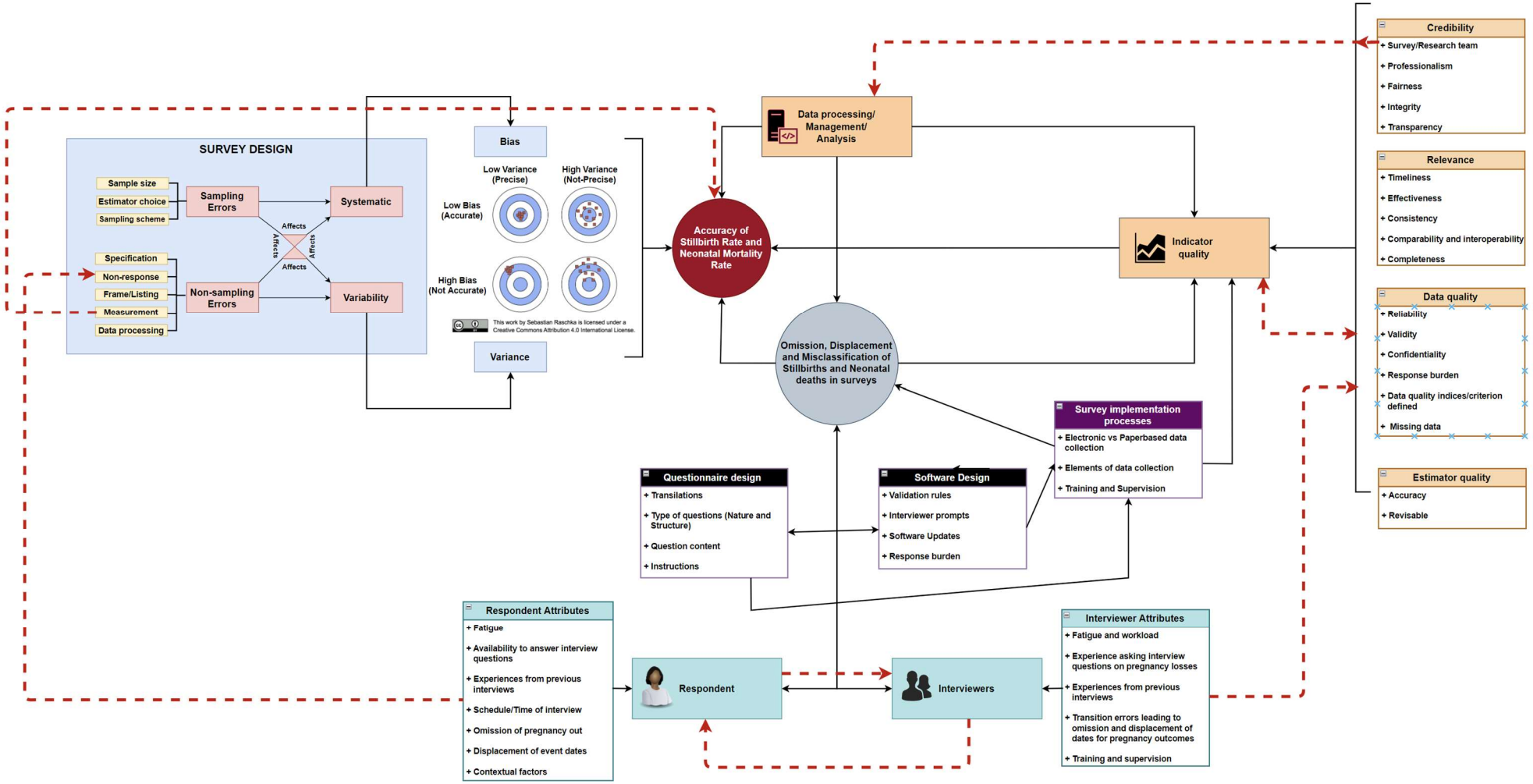
The first paper in Chapter 3 reviewed the DHS programme's forty-year existence in relation to data capture, changes in data quality, coverage, and measurement gaps for stillbirths and neonatal deaths. Since no previous studies had systematically examined the evolution of the DHS programme for data capture, quality, and measurement of these outcomes, in this PhD, POPLINE and PubMed databases, DHS website and methodological reports were searched and examined for information about the programme's implementation for the capture of data on stillbirths and neonatal deaths since 1982. Model questionnaires and data were analysed, focusing on data quality for the period 1982 to 2019.[62]

Evidence generated

DHS programme, as the largest global producer of data on stillbirths and neonatal deaths in over 90 countries, has conducted over 320 surveys since 1984 using two maternity histories; a FBH+ and FPH (implemented in fewer surveys by 17 countries). Although these surveys are credited as an important source of data on stillbirths and neonatal deaths, evidence from the data we reviewed shows that the data quality for stillbirths and neonatal deaths varied over time for both maternity histories used in the DHS programme, because of the several changes and improvements made to the DHS core model questionnaires over the years.

Several approaches and methods are used to capture pregnancies and related outcome data within the DHS programme (Forward Approach, Backward Approach, back truncated, and maternity calendars).[63] The Backward Approach was found to improve data capture for later events compared to the earlier pregnancy outcomes, although available evidence was not sufficient to warrant a major change in the DHS approach.[46] The back truncated approach was piloted in the Dominican Republic and Peru to examine whether this approach improved data quality. Available evidence showed that no additional improvements in terms of data quality were associated with the shortened maternity history compared to the full history. The reproductive calendar approach was found to improve tracking of contraceptive use, because it is easy to implement systematically and information on when a woman is pregnant and when pregnancy has ended is easy to track, therefore information on gestational age and

Figure 8.1: Conceptual framework showing potential measurement errors that affect the SBR and NMR estimates computed from retrospective survey data



stillbirths can be easily extracted from these calendars although data management expertise is required to extract this information for computing SBR.

Over the entire DHS' lifetime, questionnaires have been reviewed regularly to improve data capture of pregnancy outcomes, including stillbirths and neonatal deaths. These modifications made included 1) the addition of a question about day of death or birth option to a previously asked question with only month and year of birth/death, 2) addition of other new questions and 3) improvements in question wording and order in its seventh phase which was aimed at improving the overall data capture. Additionally, we found large variations in data quality for stillbirths and neonatal deaths from the previous surveys across all DHS phases. The highest magnitude of variation was detected for surveys with a FPH.[62] Since there was limited evidence on how FBH+ and FPH survey modules differ in terms of capture of stillbirths, we designed a study that randomised women to either a FBH+ or a FPH (Chapter 5).[36] The study was designed to examine what approaches (methodologies) and how settings and contextual factors can improve our understanding of SBR and NMR data quality through comparing two maternity histories previously used within the DHS programme. From this assessment, we learnt more about the application of these modules within the DHS programme.

8.2.2 Randomised comparison

In light of the forementioned limited research about the capture of stillbirths in standardised surveys using either a FBH or FPH; we designed a randomised multi-country study to bridge the knowledge gap in practice.[7,34,36,64,65] The second paper (Chapter 4) described the study protocol, software and site selection, and methodological approaches. The third paper (Chapter 5) examined how the use of FPH module in standardised surveys affects the capture of stillbirths and neonatal deaths and the additional time taken in implementing the FPH compared to a FBH+. A large study was conducted in five countries with a 1:1 parallel design, with over 68,000 women randomised to either a FBH+ or FPH using electronic data collection using the World Bank's Survey Solutions data collection platform.

Evidence generated

Electronic data collection platforms have shown improvements to data collection processes leading to improved quality, thus reducing overall data collection and data entry costs;

however, little or no evidence on how electronic platforms impact on survey implementation is available.[66,67] Survey Solutions, a freeware Computer Assisted Web Interviewing platform developed by the World Bank, was found to have the ability to collect paradata, which includes timestamps and the detailed audit trail of the survey implementation and management.[36,67] With Paradata, we were able to track all timestamped activities and determine how long it took to complete the survey modules for each woman.

Evidence generated from this PhD study shows that overall, the FPH reported a 21% higher SBR compared to a FBH+ with a relatively large p-value, with strong evidence of between site-heterogeneity, and in one site (Matlab) the FBH+ reported more stillbirths and higher SBR than FPH. There was no overall difference in reporting NMR between FPH and FBH+ modules, although variations were observed by site and parity. The mean completion time was slightly longer for FPH than for FBH+ (10·5 min vs 9·1 min). NMR estimates were similar to the national estimates in both modules, whereas SBR were lower than national estimates.[50] Even though the FPH reported higher SBR than FBH+, overall, the site-level estimates for SBR were lower than estimated national rates, with evidence of heterogeneity and potential misclassification between stillbirths and neonatal deaths being identified.

Furthermore, unintentional differences in implementation of sub-sections 2.1 (summary record of livebirths or pregnancies) and 2.2 (detailed record of livebirths or pregnancies) occurred between the FPH and FBH+. This difference in implementation was related to numerical validation of the number of pregnancies or births in the summary record compared to the detailed record. We observed that 14.5% of pregnancies were missing in sub-section 2.2 in Matlab in the FPH compared to <3% in other sites. No noticeable differences were observed for the FBH+ module.

8.2.3 Paradata assessments

The fourth paper (Chapter 6) further examined how survey implementation processes improve the quality of the data capture of stillbirths and neonatal deaths and how the question types (nature and structure) influence the timing and number of corrections using the EN-INDEPTH study's Paradata. There is limited use of Paradata for monitoring and improving the process of survey methodology in LMICs, particularly for maternal, newborn and child-health research compared to high-resource settings.[68-70] Missing data or

unanswered questions and survey actions can be traced from the Paradata by survey managers. Therefore, it is possible to ascertain whether an answer was deleted or was deliberately left blank. Paradata collected using electronic platforms can be analysed to determine the item level responses for asked questions in interviews as well as the item level response times in the survey, which is obtained from the timestamped Paradata.[67] Paradata from the EN-INDEPTH survey were analysed to determine whether question characteristics affected duration, and corrections by the maternity history module.[71]

Evidence generated

Corrections and errors during survey implementation (during or post-interview) may influence the overall data quality. Paradata analysis can inform, training for future surveys, and implementation of current/future surveys. Questions with multiple-choice and free-text responses as opposed to yes/no questions had the greatest number of corrections.

8.2.4 Population-level comparison

The fifth paper (Chapter 7) examined how data on stillbirths and neonatal deaths and their population-level estimates compare to routine HDSS capture of pregnancy outcomes. Findings from Chapter 5 show that there was evidence of between site variations in reported SBR but no differences in reported NMR between the FBH+ and FPH within each site. This Chapter aimed to further understand the survey by comparing with existing routine HDSS data system. Since the PhD was conducted in HDSS, with defined cohorts, it was possible to examine the between site heterogeneity by comparing population-level estimates from both HDSS and survey. However, although HDSS systems provide platforms for conducting epidemiological and demographic studies, there are few studies that had previously compared the data capture and quality of pregnancy outcomes in both HDSS and survey settings.[72-74] Despite selecting HDSS with stronger pregnancy or birth surveillance systems, the study identified substantial limitations in the HDSS data, and this paper investigated these further.

Evidence generated

A formative review and description of the five HDSSs' pregnancy surveillance processes and methods found that sites were established at different times, with different purposes, mainly

focused on childhood outcomes, and have differing pregnancy and birth outcome surveillance systems.

In three out of four sites, the survey capture of pregnancies and outcomes was more than HDSS. Survey data produced higher NMR estimates than HDSS data in all sites and study arms except in the FBH+ arm in Dabat. No substantial between-site variation in reported NMR was observed. However, it is a question of interpretation whether the survey's capture of more pregnancies and outcomes is closer to the truth. In the absence of a "gold standard" approach to data collection, this question was left unanswered. With optimal linkage, interoperability, data quality and improved surveillance a "high-quality standard" dataset can be achieved. Understanding the different contexts and methodological variations is crucial when interpreting results from HDSS. Individual-level data linkage between different datasets would further enhance our understanding of differences in data capture and quality for improving overall stillbirths and neonatal mortality estimates. With these attributes in mind, the desired "high-quality standard" dataset is certainly achievable but requires major improvements about these other aspects before it can be achieved.

Survey data produced, on average, higher SBR estimates than HDSS data, but there was substantial between-site variation in reported SBR. In both survey and HDSS data, SBR estimates from FPH were higher than FBH+ in all sites except Matlab. Matlab conducts more frequent pregnancy surveillance rounds and early pregnancy tests (once a woman misses a menstrual period). I think these two aspects have placed Matlab as a better site for optimal pregnancy surveillance. However, there were differences in implementation of both FPH and FBH+ during the EN-INDEPTH survey makes these findings inconclusive. Therefore, further investigations around data quality are needed.

There was high variability between HDSS pregnancy surveillance and data quality, sites with more frequent surveillance rounds recorded more pregnancies and their outcomes prospectively. These sites also had better linkage and interoperability between HDSS data and health facility data. Sites with more fluctuating pregnancy surveillance and update rounds were majorly characterised with longer retrospective recording of pregnancy outcome after the event. Reflections on how this impacts the data accuracy and quality suggests that additional investments in having frequent update rounds would improve the overall prospective capture of these events.

8.3. Generalisability of findings

DHS and MICS are both implemented in many low-resource settings where CRVS or registries are weak or non-existent[42], with limited or no linkage between community and health facility data.[75] Therefore, both DHS and MICS are major data sources for stillbirths and neonatal deaths estimates. They are implemented in more than 90 countries using standard methodologies and protocols implying that data and estimates generated from both MICS and DHS can be comparable across multiple countries and regions.[46,47] This PhD study utilised similar approaches and methodologies to the DHS in four African and one Asian countries; therefore, results can be potentially generalisable to populations from both regions, although findings from Matlab (the Asian site) are somewhat different from those from the African sites. I recommend the FPH module despite the Matlab findings because this PhD was powered to detect a difference for the pooled datasets, rather than the individual sites' datasets. Secondly, evidence from supplementary analyses on differences between the summary and detailed sub-sections of the FPH showed that in Matlab more than 14% events recorded in the FPH's summary sub-section were missed in the detailed sub-section. Even with this anomaly, the FPH still emerged with a higher SBR compared to the FBH+ in the pooled level analyses.

Results from this PhD study provided an evidence base for the DHS to shift from collecting a woman's birth history to her pregnancy history.[39] The shift was aimed to provide more reliable data about the woman's entire reproductive history rather than only livebirths as had previously been conducted in previous DHS surveys since 1984 as shown in Chapter 3.[62] With the FPH module to be utilised in the current DHS's eighth phase and subsequent phases, new information on stillbirths will be captured without losing any data on livebirths especially if combined with a promise of high-quality data on pregnancy outcomes. In this PhD, indicator definitions, methods and data quality measures are uniform and comparable to the standard DHS approaches, therefore, our tools, protocol and approaches can be adapted for future surveys on stillbirths and neonatal deaths.

The World Bank's Survey Solutions software is open source, and FBH+ and FPH survey modules used in this PhD are standard DHS modules, therefore, other studies and surveys can replicate this PhD's implementation in other settings. The DHS has used CSPro another open-source software on windows tablets and laptops. Until recently, CSPro did not have the ability

to capture Paradata. Furthermore, the android version of CSPro displays one question at a time and cannot be programmed to display a household roster. Given this background and from personal experiences, I think Survey Solutions' online and offline Computer-Assisted Personal Interviewing (CAPI) system and hierarchical model of survey implementation is advantageous over other packages.

In this PhD, robust methodologies and analytical strategies were utilised; for example, during analysis, we considered clustering of multiple stillbirths with one woman, this can be applied to the measurement of other pregnancy outcomes (beyond SB and NND) or similar outcomes in different surveys and contexts. Additionally, we found that questions with free-text and multiple-answer questions were characterised by multiple corrections and increased length of time. This finding can inform design and training for future surveys where problematic questions can be easily identified, and selective or extra efforts are put in place to ensure their understanding and implementation.

This PhD found differences in implementation of the FBH+ and FPH between sub-sections 2.1 (summary) and 2.2 (detailed record) in the EN-INDEPTH survey, especially in Matlab HDSS which had fewer days of training compared to other sites because interviewer were already familiar with the FBH+ approach.[20,50,55]

Future surveys, including the DHS, should take extra care when training and developing and programming FPH modules to allow for automatic validation checks and reconciliation between the summary and detailed records, reducing the certainty of missing or skipping pregnancies in either the summary or detailed records.

8.4. Strengths of the PhD

This PhD has several strengths, including its design, implementation, and methodological and contextual aspects.

Stillbirth and neonatal death data capture in the DHS

The first systematic review of data capture for stillbirths and neonatal deaths within the DHS for its 40-years of existence was conducted in this PhD. This review considered several aspects, notably approaches, tools, and data quality assessments. The DHS is an important source of these data in more than 90 LMICs; we found that data quality and accuracy was

variable, despite the various efforts and changes that have been implemented over the years to improve the capture of SB and neonatal deaths within the programme. The systematic review literature, data and questionnaires utilised laid the foundation for the subsequent works within this PhD.

Implementation factors

This PhD study was centred on a randomised comparison between two survey maternity histories. Data were collected in five sites with varying contexts and experiences as part of the EN-INDEPTH survey, which was a large study with 98% coverage of the planned sample size (66,649 out of 68,000 births that happened in the last five years since the day of interview) in five sites in five countries. Furthermore, we implemented questionnaire modules similar to DHS (1) DHS-phase 7 core module with a FBH+ and 2) recent FPH implemented in Nepal) and these were implemented uniformly across all five sites through the Survey Solutions Application.[50] These questionnaires have been tested in previous surveys within the DHS programme; therefore, we anticipate that this PhD study's data, information, and results can be generalised to other similar settings.

The design and implementation using electronic data collection with adaptations to local contexts and translation into local languages was done using the World Bank's Survey Solutions Application which likely improved the overall data quality collected during this study. Pre-programmed skip patterns and filter variables to flag erroneous data e.g., wrong computation for ages and number of children or pregnancies born. Centralised data management and analyses were possible because of uniform protocols implemented in all five sites.[50,55]

This PhD study took advantage of emerging technological advances for example questionnaire design and programming, survey management, electronic data collection, and Paradata. Survey implementation, especially those conducted within households, usually have breaks/pauses or repeats the interview. Similarly, our study was not different from others in this regard. However, through the Survey Solution's Paradata, we were able to determine the actual length of the interviews. This audit trail took into consideration all parameters that happened during the interview (corrections, implementation delays, paused or restarted interviews days or weeks or months later). During analysis of time to completion

of survey modules, a decision to restrict the analysis to only interviews conducted within 0.5 minutes to 3 hours (180 minutes) was made to exclude any outliers with very quick or delayed interviews. Overall, 1.5% and 0.6% of women with implausible values less than 1.5 minutes and 180 minutes respectively were excluded.[50]

Methodological and contextual aspects

This PhD study sought to answer programmatically important questions about improving stillbirth and neonatal death capture in standardised surveys by closing identified gaps in lack of direct comparison between the FBH+ and FPH. A robust design with triple blinding (researchers, interviewers, and respondents) and randomisation protocol of FBH+ and FPH with similar implementation across all countries was used in this study.

During the data collection and analysis, methodological appraisal of data quality through bi-weekly data monitoring, and hierarchical supervision and using select data quality metrics was done. The methodological appraisal at analysis consisted of completeness, accuracy, consistency, validity, and timeliness.

Data integrity and interoperability using multiple data systems such as survey and HDSS data, and Paradata was done. We believe this increases the usefulness and comparability of our results. Lastly, the population-level linkage between HDSS and survey was critical in highlighting some issues with data validity, therefore, laying a foundation for future research of individual-level linkage of survey data to other sources (i.e., HDSS and health facility), which will consequently improve the overall stillbirths and neonatal deaths data quality.

Paradata analysis was proof of concept of its usefulness in assessing survey data quality and length of time to completion. In this PhD, besides the assessments of timing to complete survey modules, classification of answer correction types based on structure and nature of questions and their impact on the overall time taken, we also assessed the correction patterns.

8.5. Limitations of the PhD

Similar to other studies, this PhD has some limitations in its scope, design, implementation, and methodological and contextual factors.

PhD scope of study

First, this PhD only focuses on two pregnancy outcomes (stillbirths and neonatal deaths), leaving out other pregnancy-related outcomes, notably abortions, miscarriages, preterm births, newborn infections and congenital abnormalities. Overall stillbirth and neonatal death measurement and data capture in surveys could be improved with early access and contact with health services such as antenatal care and skilled birth attendance, and gestational age (GA) data, provided health facility data can be linked to individual women/children or women and health cards which they retain and can produce during the household interview. Additionally, birth and death notification, registration and certification are important aspects for improved stillbirth and neonatal deaths within a country's CRVS system. With fully functional CRVS systems, the aim is to do away with the need for large surveys to estimate mortality; however, since CRVS systems are not fully developed, surveys conducted in LMICs can be complemented with information/ data captured by a birth or death certificate which may be reviewed by an interviewer during a household visit/ interview, although this is currently rarely done in DHS for stillbirths or neonatal deaths. However, although these aspects may generally improve survey data and quality, this PhD does not assess them.

Stillbirth and neonatal death data quality assessment for DHS data

Secondly, in view of time constraints during this PhD's scoping phase, the data quality assessments and analyses for the DHS programme data were based on aggregate data downloaded via STATcompiler and not microdata, which limits conclusions and assessments. Furthermore, our data quality assessment criterion "stillbirth rate-to-neonatal death rate ratio" that was used for assessing stillbirth data quality is simplistic and may not be robust especially when SBR and NMR vary [10,62]. Other data quality criterion stillbirths such as SBR: ENMR and relative male risk (RMR) for stillbirths were not used because sex of the stillbirth was only collected in the FPH. Similarly for neonatal deaths, data quality criteria such as neonatal mortality rate to infant mortality rate ratio (NMR:IMR), sex ratio at birth, sex ratio at death and heaping index for neonatal deaths were not utilised in this PhD because the primary focus is stillbirths. Additionally, for NMR were similar between the two survey modules and national-level estimates, therefore, we did not conduct a data quality assessment for neonatal deaths.

Implementation factors

Although robust designs with blinding and randomisation protocols and similar implementation plan were used during this PhD study across sites, in Matlab a different interviewer training protocol from the one prescribed in the EN-INDEPTH interviewer manual was used, which potentially contributed to differences in the results for the FBH+ versus FPH comparison in Matlab. Differences in interviewer training, methods, and lengths between Matlab and other sites could have influenced interviewers' understanding/comprehension of the survey modules especially in the FPH in Matlab.

Most LMICs rely on retrospective surveys for data on stillbirths and neonatal deaths. In these retrospective surveys, maternal recall of GA in months is recorded and used to determine whether a fetal death is a stillbirth or not. Similarly, in this PhD study, the FBH+ and FPH approaches captured data on GA in months rather than in weeks as is defined in the International Classification of Diseases eleven (ICD-11). This proxy measurement (GA in months) is characterised by heaping and biases[76], thus reducing the quality and accuracy of data for stillbirths generated from surveys. Furthermore, GA data from HDSS surveillance was based on maternal reporting of pregnancies and no other methods like the first ultrasound scan in the first trimester or Ballard score or last menstrual period [77,78] except in Matlab where a pregnancy test is implemented when a menstrual period is missed. Regular, early pregnancy testing in Matlab improves the probability of early pregnancy detection and gestational age assessment at the time of birth. Other sites (Dabat, IgangaMayuge and Kintampo) which capture GA retrospectively also have infrequent update and pregnancy surveillance rounds which may affect the overall quality of GA data. Subsequently, the SBR and NMR estimates because of the surveillance women may know the GA or have cards with GA data which impacts on recall in the survey and hence quality of GA survey data. Although completeness of the GA in months was high for FBH+/FPH questions, the completeness of information on GA in weeks for stillbirths was low[58] which was a finding from this study that it is difficult to get women to recall/estimate GA in weeks.

Another related limitation from this study is validity aspects of GA estimation in both survey and HDSS data were not assessed. This study could not validate the GA information gathered in either survey or HDSS because of the lack of a high-quality standard dataset to compare with.

Another implementation related limitation is that programming for sub-sections 2.1 (summary of lifetime births or pregnancies) and 2.2 (detailed record of lifetime births or pregnancies) was different for both FBH+ and FPH survey modules in this study, for example, the FBH+ module's programming validated the number of births between the two sub-sections (i.e., a summation of entries between the two sub-sections is reviewed at the end of the two sections); however, a similar protocol was not implemented for the FPH module. Although this PhD study was implemented based on DHS methodology and protocols (i.e., interviewer manuals and instructions), we note that this is a major limitation to data completeness, as indicated in Chapter 5.[50] The validation in the FBH+ between the two sub-sections may have contributed to consistency in number of live births reported by the two modules which was different for the FPH.

Software aspects

During the period of data collection July 2017 to August 2018, the World Bank released several updates of the Survey Solutions which may have affected the data quality processes, programming, training, and structural changes to the database.

First, the software updates led to delays and changes in the survey implementation because software updates needed reprogramming of some sections of the EN-INDEPTH questionnaire which sometimes required stopping the data collection to ensure that all data collection tablets are updated to the most current version. Additionally, incompatibility between data from different software versions (i.e., dates) led to data transformations at data cleaning and analysis stages of this PhD work which may have affected the overall data quality. From our experiences, the newer Survey Solutions versions had improved features, although they came at a cost such as delays due to reprogramming, incompatibility between similar variables in old and newer versions.

On reflection, an important lesson is that it would have been better to stick to one software version to avoid these associated delays and changes in data structure because in the end, data transformation will be required, which may reduce the data quality for example, data from the same variable may be assigned to a different scale in a different version (timestamps with minutes and seconds vs timestamps with only hour). Secondly, frequent software updates led to delays in data collection because subsequent data collection was halted in

some sites (e.g., Bandim and Matlab) and interviewers required refresher training to introduce the new changes. These software updates also necessitated concurrent updates of interviewer user manuals which required back and forth translations. If I were to repeat this study, I would prefer to use a later software version but would make a decision to stick to one version.

Our review of the different software platforms (Census and Survey Processing software (CSPro), Open Data Kit, Qualtrics, RedCap and Survey Solutions) before the survey implementation revealed that the World Bank's Survey Solutions platform was the optimal choice for data collection at the study design stage because it had the ability to link different questions to the roster file, collect Paradata, include validation rules, open-source software with both online and offline administration, hierarchical survey administration between interviewers, supervisors and data managers and end-user technical support provided (Table 8.1).[36] Survey Solutions' hierarchical survey administration allows supervisors and survey administrators to review interviewer submissions daily and give feedback to the interviewers instantly. Additionally, supervisors/administrators can accept or reject the submitted interviews. Rejected interviews return to the interviewers to make any necessary changes to the submitted data. I think DHS and MICS, which implement CSPro, should consider switching to Survey Solutions because of its additional abilities and advantages during survey implementation and management. CSPro has improved in recent years to include the capture of Paradata; however, it does not offer the hierarchical mode of data collection, and it shows one question at a time if the android platform is used (Table 8.1).

Table 8. 1: EN-INDEPTH Data collection software/mobile data collection platform review

	PROS	**CONS**	**Selected Software**
CSPro – windows tablets	• No licence required, free. • Used for data collection at 75-80% of all DHS countries. • It allows to display birth history/ fertility history questions in one table on a screen	• Requires tablets with Windows OS; • INDEPTH sites use Android OS based tablets; • We need to buy Windows OS operated tablets for all sites • Adaptation of the questionnaire for our study requires knowledge of ASCII language and the adaptation is not straightforward • Limited user support	No

CSPro – android tablets	• No licence required, free. • Sites have Android OS based tablets • Allow to ask birth history/ fertility history questions one by one	• Limited user support • Adaptation of the questionnaire for our study requires knowledge of ASCII language and the adaptation is not straightforward • Does not allow to display birth history/ fertility history questions in one table on a screen	No
Survey Solution	• No licence required, free • Design of the questionnaire is done online • The app is similar to ODK • The birth history/fertility history questions are asked one by one but the app allows to create a roster similar to the table • Data collection and data management is done integrated in the app • App collect detailed audit of the survey called Paradata	• Does not allow to display birth history/ fertility history questions in one table on a screen	Yes
Qualtrics	• The software allows to display the birth history/fertility history questions in one table • Adaptation is more straightforward	• One licence costs - £5000 per year • We need to create the questionnaire	No
ODK	• No licence required, free • INDEPTH sites have experience in collecting data using this app • LSHTM provides support with setting up the server • Birth history/ fertility history questions are asked one by one	• To create the questionnaire, we need to programme in ODK, time consuming • Does not allow to display birth history/ fertility history questions in one table on a screen • LSHTM does not provide support in creating the questionnaire	No

Methodological and contextual factors

The first methodological limitation about this PhD study is that when assessing the length of time to completing the maternity history modules, we only focused on a subset of the

Paradata parameters notably the Answer Set (final answers or data retained in the database) and Corrections made by the interviewer. We exclude process-related activities for example, interviewer and supervisor comments, time lags between assignment of interviews and starting, pausing, completing, and restarting interviews and system calculations. Although these were accounted for in our computations which consisted of the first and last timestamps for the first and last questions in the maternity history modules, the effects of other survey process-related parameters (such as error messages and notification, comments from the interviewer or supervisors, static calculations, interview assignments and reassignments and revisits) were not examined in this PhD. Alternative analysis and approaches that were beyond the scope of this PhD but can be assessed using paradata include: understanding response rates, relationships between correction rates and correction type and being supervised. Individual and team productivity, contact attempts without success, response propensity, and workload impact for countries with additional questions. I believe, appraisal of individual and teams' productivity using digital, or visualisation approaches contributes to our overall understanding of survey data quality because potential errors or inconsistencies in data collection may be identified earlier on during the survey implementation rather than, as in this study, at analysis.

Although we relied on local data collection teams familiar with the different contexts and settings, we invested additional time into training, field testing, and development of field user manuals and guides.

Thirdly, Paradata is voluminous and very complicated to manage; however, recent advancements in statistical software such as dashboards and dynamic documents using Markdown in both Stata and R programming software can make Paradata analysis seamless.[79,80] During this PhD, I faced several challenges (methodological and analytical) with the survey Paradata. Therefore, all paradata analyses were focused on only the maternity histories. The overall timing and corrections of/in sections of the survey questionnaires (i.e., background information, pregnancy and postnatal, fertility preference, SES) was not examined. Therefore, their overall impact on the timing and number of corrections for maternity history sections could not be ascertained. Other software platforms that we reviewed do not produce a detailed record of the interview/survey implementation beyond the timestamps. I think performing routine paradata analysis with automated processes

during or after survey implementation for purposes of monitoring shall improve the overall data quality.

Lastly, although study sites were selected based on predefined criteria described in Chapter 4, in Chapter 7, we found that HDSS pregnancy surveillance was limited in most sites and could not be used as a reference standard as we had initially hoped. Whilst we could compare the survey and HDSS data by arms of the study, we could not compare the true prevalence of these adverse pregnancy outcomes in the population because the HDSS surveillance were limited. Furthermore, differences in contexts and pregnancy surveillance (i.e., Bandim HDSS) did not allow for comparisons between HDSS and survey data in all sites. A related limitation to this PhD study is that the HDSS and survey data comparisons were restricted to women who were continuously resident within the HDSS since 1st January 2012, these women may differ from their mobile counterparts whom we excluded; therefore, these results may not be generalisable to all populations or even the HDSS population.

8.6. Evidence synthesis - Measurement changes and implications now

Considering all the evidence generated by the work in this thesis together; the FPH module's performance was better producing a 21% higher SBR than that of the FBH+ with not much differences in implementation time. Further evidence that the choice of maternity history module and approaches (forward vs backward) is strongly influenced by the DHS's core questionnaire (Chapter 1). Hence this PhD presents a case of the importance of generating robust evidence to inform the DHS questionnaire selection procedure and engagements with the DHS in the process as was done by the EN-INDEPTH group for the phase eight revisions of the DHS questionnaire.

Within the DHS programme choice between implementing a survey questionnaire with a FBH+ or a FPH was left to the country DHS coordinating teams; however, the FBH was the core module, so most countries opted to use it by default. Similarly, the varying approaches used to capture data on stillbirths and neonatal deaths are also left to the country teams to decide, although the core DHS model questionnaire utilises the forward approach.[46,63] Furthermore, the variations in data from the FPH module across the DHS programme surveys could be attributed to variations in implementation and trainings across the countries. Recently, in August 2017 the DHS programme produced a methodological report number 19 on data quality entitled "An Assessment of the Quality and consistency of Age and Date Reporting in DHS Surveys, 2000-2015".[81] This report gives details about procedures for editing and data quality. Another methodological report has shown the different aspects and indicators for assessing data quality for stillbirths and neonatal deaths. Consequently, the data editing algorithm for identifying events within the period of inquiry (3-year or 5-year threshold) has also been reviewed and changed from Century Month Code (CMC) to Century Day Code (CDC)[82] in an effort to further improve data quality through reducing date displacement over the periods of inquiry and to also foster better gestational age assessments although our survey data revealed that completion of day of the event was generally missing, therefore, like the DHS programme, the CMC algorithm was implemented in this PhD.

The current digital age because of technological advancements, recommends electronic data collection for all future surveys to improve data capture quality. The World Bank's Survey Solutions also captures Paradata which can be examined to provide further insights into survey implementation and management. However, it is also important for data analysts to

have the expertise and physical resources like computers with enough random-access memory (RAM) to conduct these analyses because they are quite cumbersome based on our findings and experiences. Dashboards and automation of analyses using the recent improvements in statistical software can also improve the usability of Paradata from surveys.

From this PhD work, we established that the FPH has potential to increase the capture of stillbirths over the FBH+ if utilised in a household survey.[50] Based on these results, the DHS programme made a decision to switch from a FBH+ to a FPH in its core modal questionnaire in an effort to further improve capture of pregnancy losses in its eighth phase.[39] However, although improvements in data capture are observed when a FPH module is used, surveys need to have a data quality assessment protocol for stillbirths and neonatal deaths through using existing metrics e.g., SBR:NMR ratio and including the SBR:NMR ratio in the standard DHS analysis tables or develop new metrics that can be used for data quality assessments.

Changes from paper-based to electronic data collection through developing and rolling out an electronic data collection protocol that uses CSPro software in DHS programme countries is timely [83]; however, internal inconsistencies between sub-sections 2.1 (summary) and 2.2 (detailed record) for the FPH module (In Matlab, the sub-section 2.2 missed at least 14% of events recorded in sub-section 2.1 compared on less than 3% in other sites) should be corrected and should account for automatic reconciliation between the summary and detailed records as it was implemented in the FBH+ during phase 8 and in this PhD study.[55,84] Furthermore, Paradata, which includes survey the audit trail and timestamps, can reveal problematic questions, and interviewer performance during survey implementation in addition to the preprogramed error notifications which could reduce data missingness and encourage instant data correction, ultimately leading to improved data quality. Recently CSPro introduced a Paradata function in its version 7.5.[85] However, there is need to examine how it can be used to improve survey implementation and data capture quality.

This PhD study's assessments in Chapter 6 revealed that approximately eight out of ten survey interviews have at least one correction, especially questions with free-text and multiple-choice type questions had the most corrections. In addition, this PhD revealed that methodologies to conduct Paradata analysis exist although, not enough evidence of its uses is available in Maternal and Newborn research especially for stillbirths and neonatal deaths.

Chapters 3, 5, 6 and 7 showed that data quality for stillbirths and neonatal deaths may be affected by various factors (such as missing data due to loss when migrating from one system to another, inability to link survey data to other data sources because of lack of a high-quality standard dataset to reference).[50,62,71,86] A qualitative assessment by Kwesiga found that contextual factors, deliberate deletion or concealing of pregnancy outcomes from the interviewers or respondents respectively, had an influence on the data quality.[20] Increasing coverage of community and facility births and deaths registration and recording, HMIS (e.g., DHIS-2) and CRVS[87] – more data available – to improve women's need to know this information which will improve survey reporting. Another aspect is that unwanted pregnancies may not be recorded.[88]

Overall, findings from this work differed between the African and Asian sites although pooled estimates suggested improved capture of stillbirth using the FPH. Further assessments into the data quality shall reveal further strengths or weakness in using either the FBH+ or FPH. Improved data quality and interoperability shall improve overall measurement of stillbirths and neonatal deaths therefore, reducing potential misclassification, omission and date displacement of when the birth events occurred.

8.7. Future research

This PhD revealed that there were no previous studies that directly compared two maternity histories either a FBH+ or a FPH for capture of stillbirths and neonatal deaths using "robust" randomised controlled trial study design.[89] These findings are timely as all countries have made commitments to reduce stillbirth rate and neonatal mortality rates to less than 12 per 1,000 total births and 12 per 1,000 livebirths respectively by the year 2030.[61] Further assessments of data quality and accuracy were done through subjecting survey data to data quality criterion or comparing survey and HDSS population-level estimates. However, further assessments need to conduct individual-level linkage analyses to conduct more detailed data quality assessments and comparisons of data collected using different platforms. This appraisal of both HDSS and survey data is important to assessing the impact of non-sampling errors on the SBR and NMR.

The recent addition of the day of death/birth to the roster in the DHS phase-7 questionnaire was aimed to improve data quality, additional changes to the data editing algorithm for

identifying events within the threshold from CMC to CDC, however, additional research in how these changes have impacted data quality is needed.

This PhD examines Paradata to answer the question “On average, how long does the FBH+ or FPH survey module take to complete?”; however, it does not answer the question “How can Paradata information be used to improve prospective and future survey implementation and monitoring?” I recommend further assessments and automation of Paradata analyses through using dashboards and dynamic documents to improve timely identification of errors, feedback and monitoring, in an effort to improve data quality. Future surveys should optimise the use of technological advances during implementation through using automation, local data collection feedback loops and dashboards to track data quality, interviewer performance and survey management.

Results from this PhD influenced the switch between a FBH+ and FPH in phase-8 of the DHS programme, and also compared our survey data to HDSS; however, no data quality assessments between surveys and routine data were conducted because we lack a “high-quality reference standard dataset” in view of challenges with creating or obtaining a “gold standard dataset”. As digital and technological innovations and advances are increasing in LMICs, this will allow for interoperability and data linkage hence improving the prospects of creating “high-quality standard datasets” that more accurately reflect the truth. Future studies should consider examining the data quality aspects between surveys and routine data (i.e., HDSS, health information systems among others). The notion of “high-quality standard datasets”, may not be entirely achievable, except, with improved data quality and interoperability backed with the technological advancement. These advancements shall improve capture of stillbirths, neonatal deaths and reductions in potential misclassification, omission and date displacements that could be seen in future standardised surveys. However, understanding the different contexts and methodological variation is important when interpreting results from HDSS data.

Improvements in question structure and software design, and additional efforts to optimise survey implementation of the FPH approach, including training, interviewer, prompts, translations, and developing and supervision models, will require further research because they directly affect the data quality. Furthermore, qualitative assessments should be linked to research aimed at improving our understanding of question performance, question

structure, and software design using Paradata to inform survey design. In addition, production of DQR from all data collection platforms for data on stillbirth and neonatal deaths will increase the acceptability and trustworthiness of the available data on these indicators.

Since free-text and multiple-choice questions were identified to have the most corrections, future surveys should emphasise such questions during interviewers and data processors' trainings, in addition to data validation and cleaning. Furthermore, the impact of data quality reporting on subsequent survey estimates reported from standardised surveys should be examined.

Individual-level data linkage between surveys and HDSS would further enhance our understanding of differences in data capture and quality for improving stillbirth and neonatal mortality estimates. Data accuracy and quality assessments through health cards and data linkage for pregnancy outcome reporting would reduce omission and misclassification of SB in surveys and HDSS data.[90]

8.8. Recommendations

Findings in this PhD highlight the importance of improved data capture and quality for stillbirths and neonatal deaths. Standardised surveys notably the DHS remains the major source of data on stillbirths and neonatal deaths in many LMICs.[46] Therefore, six recommendations are suggested

1. **Training, supervision and considering contextual factors when interpreting results from analyses:** The recent change to the core module at the DHS programme is justified to improve data capture particularly for stillbirths, however, another dimension of "data quality" depends on the entirety of the survey implementation, therefore it is crucial that extra efforts are done before and during data collection (training and supervision) and during data cleaning and analysis. Furthermore, future studies need to consider the contextual factors when interpreting results from where these data are gathered.
2. **Data quality reports for stillbirth and neonatal deaths alongside main results reports:** Household surveys and longitudinal surveillance, especially those capturing data on pregnancy outcomes (stillbirths and neonatal deaths), should emphasise data monitoring, validation, cleaning, and production of timely data quality reports (DQR)

alongside the main estimates and reports. The DQR should show details of measures taken to improve the data relevance, accuracy, and reliability such as assumptions made, missing data analyses, methods of data imputation and how outliers were handled among others. The DQR should be produced promptly and be accessible to encourage coherence and comparability.

HDSS should subsequently produce their data quality reports alongside estimates reports and demographic profiles because it encourages transparency which will in turn improve the confidence with HDSS data on stillbirth and neonatal deaths.

3. **Investments to improve data quality**: Poor data quality will continue to lead to poor policy or programme choices or wrong decisions and interventions concerning maternal and newborn health, therefore taxpayers' or donor funds and money will not be well spent. However, good data quality, effective communication of estimates and decision-making are enabling conditions giving countries mileage as they make efforts towards realising their SDG and ENAP targets. HDSS sites particularly, should consider conducting more frequent update rounds because of the potential improvements in capturing events earlier on. Additionally, conducting early pregnancy tests for women within the reproductive age group can be used a quick and "somewhat cheaper" means to confirm pregnancy status of the women during home visits (as evidence from Matlab shows). Although these enhancements to HDSS systems are suggested, the sites should note that they also increase their associated running costs of the sites' business processes.
4. **Electronic data capture, paradata possibilities and visualisation:** Digital health and visualisation using digital and technological advances can improve the generation of DQR and final reports through implementation dashboards with visualisation capabilities to ensure timely checking and validation of data and data collection processes; this will subsequently improve reporting timelines and estimates generated.
5. **Increasing coverage of facility births, HMIS (e.g., DHIS-2) and CRVS:** Optimal data validation systems and processes are crucial for improving measurement and data quality for stillbirths and neonatal deaths. Provision of frameworks for data linkage across different platforms and data sources, notably hospital, national identification,

national registries cause of death, birth and death notification, registration, and certification systems are essential.

6. **Data use:** Investment in national data quality improvement strategies, governance, digital health, data visualisation and monitoring and evaluation has the potential to improve stillbirths and neonatal deaths data. This will consequently improve the early identification of implementation challenges and enable the provision of solutions that will guide them to achieve their SDG and ENAP targets.

8.9. Conclusion

This PhD demonstrated that capture of stillbirths was improved using a FPH maternity history compared to the FBH+. This evidence was used to inform DHS' switch to the FPH module in its eighth phase. However, limited evidence concerning stillbirths and neonatal data accuracy and quality in standardised surveys is available. Future studies should also consider examining the combined effect of survey processes and implementation, software and questionnaire designs and omission and misclassification on the overall stillbirth and neonatal deaths estimate due to potential measurement errors. Therefore, additional efforts towards improved survey implementation of the FPH approach, including training, interviewer prompts, translations, and developing and testing standard data quality criteria for stillbirths and neonatal deaths in surveys and routine data, are warranted. If I were starting this study now, I would do the following differently; 1) Improve programming to include a validation between summary and detailed records of the maternity history, 2) Emphasize the differences and approaches to both the FPH and FBH+ during training and survey implementation 3) Conduct an individual-level linkage and analyses for the HDSS and survey beyond the population-level comparison (Chapter 7), this would further generate insights into validating the results and show further evidence about the extent of the potential misclassification, time displacement and omission in absence of high-quality standard datasets.

References

1. Waiswa P, Akuze J, Moyer C, 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.
2. UNICEF, WHO, World Bank, UN. Levels and trends in child mortality: report 2020, 2020.
3. Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *Lancet* 2014; **384**(9938): 189-205.
4. World Health Organization. Every Newborn: An action plan to end preventable deaths. 2014.
5. United Nations Children's Fund, Cappa C, Petrowski N. Birth registration for every child by 2030: are we on track?: Unicef; 2019.
6. Bhatia A, Ferreira LZ, Barros AJ, Victora CG. Who and where are the uncounted children? Inequalities in birth certificate coverage among children under five years in 94 countries using nationally representative household surveys. *International journal for equity in health* 2017; **16**(1): 1-11.
7. Stanton CK. An Investigation into the Use of Demographic Survey Data for the Study of Perinatal Mortality and Its Determinants. Baltimore, Maryland: Johns Hopkins University; 1996.
8. de Bernis L, Kinney MV, Stones W, et al. Stillbirths: ending preventable deaths by 2030. *Lancet* 2016; **387**(10019): 703-16.
9. United Nations. Sustainable Development Goal 3: Ensure healthy lives and promote well-being for all at all ages. 2015. http://www.un.org/sustainabledevelopment/health/ (accessed 5/2/2016 2016).
10. Lawn JE, Blencowe H, Waiswa P, et al. Stillbirths: rates, risk factors, and acceleration towards 2030. *Lancet* 2016; **387**(10018): 587-603.
11. Cleland J. Demographic data collection in less developed countries 1946-1996. *Popul Stud (Camb)* 1996; **50**(3): 433-50.
12. Demographic and Health Surveys. DHS model questionnaires. 2018. https://dhsprogram.com/What-We-Do/Survey-Types/DHS-Questionnaires.cfm (accessed 08.02.2018 2018).
13. DHS Program. Demographic Health Surveys. 2017. https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm.
14. Rutstein SO, Rojas G. Guide to DHS statistics. *Calverton, MD: ORC Macro* 2006.
15. Biemer PP, Groves RM, Lyberg LE, Mathiowetz NA, Sudman S. Measurement errors in surveys: John Wiley & Sons; 2011.
16. Streatfield PK, Alam N, Compaore Y, et al. Pregnancy-related mortality in Africa and Asia: evidence from INDEPTH Health and Demographic Surveillance System sites. *Glob Health Action* 2014; **7**(1): 25368.
17. Espeut D, Becker S. The validity of birth and pregnancy histories in rural Bangladesh. *J Health Popul Nutr* 2015; **33**(1): 17.
18. Blacker J, Brass W. Experience of retrospective demographic enquiries to determine vital rates. *The Recall Method in Social Surveys, London University Institute of Education, London* 1979.
19. 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): 24011.
20. Kwesiga D, Tawiah C, Imam MA, et al. Barriers and enablers to reporting pregnancy and adverse pregnancy outcomes in population-based surveys: EN-INDEPTH study. *Popul Health Metr* 2021; **19**(Suppl 1): 15.
21. Moss L. The recall method in social surveys: William Heinemann Limited; 1979.
22. Rutstein SO. Infant and child mortality: levels trends and demographic differentials. Voorburg, Netherlands: International Statistical Institute., 1983.
23. Rutstein SO. Infant and child mortality: levels trends and demographic differentials. Voorburg, Netherlands: International Statistical Institute., 1984.

24. Silva RMA. Papers on Indirect Mortality Estimation & Analysis in Low-Resource Settings: UC Berkeley; 2013.
25. Curtis SL. Assessment of the quality of data used for direct estimation of infant and child mortality in DHS-II surveys. Calverton, Maryland: Marco Internation Inc, 1995.
26. Graesser AC, Karnavat AB, Daniel FK, Cooper E, Whitten SN, Louwerse M. A computer tool to improve questionnaire design in statistical working policy. Federal Committee on Statistical Methodology 2001; Washington, DC: Washington, DC: Bureau of Labor Statistics; 2001. p. 36-48.
27. Graesser AC, Wiemer-Hastings K, Kreuz R, Wiemer-Hastings P, Marquis K. QUAID: A questionnaire evaluation aid for survey methodologists. Behavior Research Methods, Instruments, and Computers; 2000; Washington, DC: Washington, DC: Bureau of Labor Statistics; 2000. p. 254-62.
28. Graesser AC, Wiemer-Hastings K, Wiemer-Hastings P, Kreuz R. The Gold Standard Of Question Quality On Surveys: Experts, Computer Tools, Versus Statistical Indices. American Statistical Association; 2000; Washington, DC: Washington, DC: Bureau of Labor Statistics; 2000.
29. Wiemer-Hastings K, Wiemer-Hastings P, Rajan S, Graesser AC, Kreuz RJ, Karnavat A. A detector for presuppositions in survey questions. the Joint Language Technology Conference; 2000; Washington, DC: ACL Press; 2000.
30. Pullum TW, Stokes SL. Identifying and adjusting for recall error with application to fertility surveys. *Survey Measurement and Process Quality* 1997: 711-32.
31. Pullum TW, Becker S. Evidence of Omission and Displacement in DHS birth histories: Methodological reports Rockville, Maryland, USA: ICF International, 2014.
32. DHS Program, USAID. Ghana Demographic Health Survey, 2008.
33. Sebastian-Coleman L. Measuring data quality for ongoing improvement: a data quality assessment framework: Newnes; 2012.
34. USAID. Ghana Maternal Health Survey 2007, 2007.
35. Pullum TW. An Assessment of Age and Date Reporting in the DHS Surveys, 1985-2003. Calverton, Maryland: Macro International Inc., 2006.
36. 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**(1): 010901.
37. Organization WH. Mother and Newborn Information for Tracking Outcomes and Results (MONITOR) technical advisory group. 2019.
38. Demographic and Health Surveys. DHS8 Questionnaire Summary: Revision process and new content: DHS, 2019.
39. Demographic and Health Surveys. DHS phase 8 model questionnaires. 2019. https://www.dhsprogram.com/publications/publication-DHSM11-DHS-Questionnaires-and-Manuals.cfm (accessed 30.03.2021 2019).
40. Akuze J, Cousens S, Lawn JE, et al. Four decades of measuring stillbirths and neonatal deaths in Demographic and Health Surveys: historical review. *Population health metrics* 2021; **19**(1): 1-14.
41. Akuze J, Blencowe H, Waiswa P, et al. Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in five countries: the Every Newborn-INDEPTH study. *The Lancet Global Health* 2020; **8**(4): e555-e66.
42. 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.
43. Mikkelsen L, Phillips DE, AbouZahr C, et al. A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet* 2015; **386**(10001): 1395-406.
44. UN. Sustainable Development Goal 3: Ensure healthy lives and promote well-being for all at all ages. 2015. http://www.un.org/sustainabledevelopment/health/ (accessed 5/2/2016 2016).
45. World Health Organization. Every Newborn: an action plan to end preventable deaths. 2014.
46. Cleland J. Demographic data collection in less developed countries 1946–1996. *Population Studies* 1996; **50**(3): 433-50.

47. Cousens S, Blencowe H, Stanton C, et al. National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. *The Lancet* 2011; **377**(9774): 1319-30.
48. Sankoh O, Byass P. The INDEPTH Network: filling vital gaps in global epidemiology. *Int J Epidemiol* 2012; **41**(3): 579-88.
49. Joseph Akuze, Simon Cousens, Joy E Lawn, et al. Thirty-five year history of Demographic and Health Surveys' capture of stillbirths and neonatal deaths. *BMC population metrics* 2019.
50. Akuze J, Blencowe H, Waiswa P, 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-e66.
51. INDEPTH. INDEPTH Network. 2016. http://www.indepth-network.org/about-us.
52. World Bank. Survey Solutions CAPI/CAWI platform: Release 5.26. . Washington DC,: The World Bank, 2018.
53. Akuze J, Blencowe H, Fisker AB, et al. Linking pregnancy outcomes between population-based surveys and prospective demographic surveillance: EN-INDEPTH study BMC Population Health Metrics; 2020.
54. Fisker AB, Rodrigues A, Helleringer S. Differences in barriers to birth and death registration in Guinea-Bissau: implications for monitoring national and global health objectives. *Trop Med Int Health* 2019; **24**(2): 166-74.
55. Thysen SM, Tawiah C, Blencowe H, et al. Electronic data collection: Experiences from a multi-site population survey, the EN-INDEPTH study. *BMC Population Health Metrics* 2020.
56. Nareeba T, Dzabeng F, Alam N, 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**(Suppl 1): 7.
57. Kwesiga D. Qualitative study of barriers and enablers to reporting of pregnancy and adverse pregnancy outcomes in population-based surveys: EN-INDEPTH multi-country population-based study. *BMC Population Health Metrics* 2019.
58. Blencowe H, Bottecchia M, Kwesiga D, et al. Stillbirth outcome capture and classification in population-based surveys: EN-INDEPTH study. *Population health metrics* 2021; **19**(1): 1-19.
59. Haider MM, Mahmud K, Blencowe H, et al. Gestational age data completeness, quality and validity in population-based surveys: EN-INDEPTH study. *Population health metrics* 2021; **19**(1): 1-18.
60. Biks GA, Blencowe H, Hardy VP, et al. Birthweight data completeness and quality in population-based surveys: EN-INDEPTH study. *Popul Health Metr* 2021; **19**(Suppl 1): 17.
61. Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *The Lancet* 2014; **384**(9938): 189-205.
62. Akuze J, Cousens S, Lawn JE, et al. Four decades of measuring stillbirths and neonatal deaths in Demographic and Health Surveys: historical review. *Popul Health Metr* 2021; **19**(Suppl 1): 8.
63. Becker S, Mahmud S. A validation study of backward and forward pregnancy histories in Matlab Bangladesh: International Statistical Institute, 1984.
64. DHS, USAID. Ghana Demographic Health Survey, 2008.
65. Espeut D, Becker S. The validity of birth and pregnancy histories in rural Bangladesh. *Journal of Health, Population and Nutrition* 2015; **33**(1): 17.
66. Taiwah C, Thuysen S. Choices and learnings from large scale electronic data collection: EN-INDEPTH multi-country population-based study. *BMC Population Health Metrics* 2019.
67. World Bank. Survey Solutions Computer-Assisted Personal Interviewing (CAPI)/Computer Assisted Web Interviewing (CAWI) platform: Release 5.26. 2018. http://surveys.worldbank.org/capi (accessed 12/3/2019.
68. Kreuter F. Improving surveys with paradata: Analytic uses of process information: John Wiley & Sons; 2013.
69. Nicolaas G. Survey paradata: a review. 2011.

70. Choumert-Nkolo J, Cust H, Taylor C. Using paradata to collect better survey data: evidence from a household survey in Tanzania. *Review of Development Economics* 2019; **23**(2): 598-618.
71. Gordeev VS, Akuze J, Baschieri A, et al. Paradata analyses to inform population-based survey capture of pregnancy outcomes: EN-INDEPTH study. *Population health metrics* 2021; **19**(1): 1-14.
72. Akuze J, Baschieri A, Kerber K, et al. Do different HDSS surveillance systems result in different quality of pregnancy outcome data? International Population Conference 2017 (IPC/IUSSP 2017). Cape Town, South Africa: Statistics South Africa; 2017.
73. Kadobera D, Waiswa P, Peterson S, 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. 2017; **10**(1): 1356641.
74. Rerimoi AJ. Development and assessment of systems for population-based estimation of neonatal and perinatal mortality in The Gambia: London School of Hygiene & Tropical Medicine; 2019.
75. Newcombe. Howard B., Kennedy. James M, Axford. SJ, James. Allison P. Automatic linkage of vital records. *J Science* 1959; **130**(3381): 954-9.
76. Pullum TW, Becker S. Evidence of omission and displacement in DHS birth histories. DHS Methodological Reports No. 11. Rockville, Maryland, USA: ICF International, 2014.
77. Stevenson A, Joolay Y, Levetan C, Price C, Tooke L. A Comparison of the Accuracy of Various Methods of Postnatal Gestational Age Estimation; Including Ballard Score, Foot Length, Vascularity of the Anterior Lens, Last Menstrual Period and Also a Clinician's Non-Structured Assessment. *Journal of Tropical Pediatrics* 2021; **67**(1): fmaa113.
78. Miller L, Wanduru P, Santos N, et al. Working with what you have: How the East Africa Preterm Birth Initiative used gestational age data from facility maternity registers. *PloS one* 2020; **15**(8): e0237656.
79. Hemken D. STMD: Stata module to convert dynamic Markdown to HTML format, using Stata dyndoc. 2019.
80. Roca-PhD T. Web programming and dataviz with Stata. 2014.
81. AGE CO. DHS METHODOLOGICAL REPORTS 19. 2017.
82. Croft T, Aileen M. J. Marshall, Courtney K. Allen. Guide to DHS Statistics. Rockville, Maryland, USA: ICF. : D, 2018.
83. United States Census Bureau. Census and Survey Processing System (CSPro). 2019. https://www.census.gov/data/software/cspro.html (accessed 30/12/2019.
84. Akuze J, Blencowe H, Waiswa P, et al. Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in 69,176 pregnancies in five countries: Every Newborn-INDEPTH study. *Lancet Global Health* 2019.
85. Fischer OP. Innovative approaches to support the global 2020 Round of Population and Housing Censuses. *Statistical Journal of the IAOS* 2020; **36**(1): 51-4.
86. Joseph Akuze, Hannah Blencowe, Ane B Fisker, et al. Comparing population-level estimates of pregnancy outcomes between population-based surveys and prospective demographic surveillance systems: an EN-INDEPTH study 2021.
87. Kasasa S, Natukwatsa D, Galiwango E, et al. Birth, stillbirth and death registration data completeness, quality and utility in population-based surveys: EN-INDEPTH study. *Popul Health Metr* 2021; **19**(Suppl 1): 14.
88. Yargawa J, Machiyama K, Hardy VP, et al. Pregnancy intention data completeness, quality and utility in population-based surveys: EN-INDEPTH study. *Population health metrics* 2021; **19**(1): 1-18.
89. Parab S, Bhalerao S. Study designs. *International journal of Ayurveda research* 2010; **1**(2): 128.
90. Nareeba T, Dzabeng F, Alam N, et al. Neonatal and child mortality data in retrospective population-based surveys compared with prospective demographic surveillance: EN-INDEPTH study. *Population health metrics* 2021; **19**(1): 1-17.

Appendices

A1: Summary of the role of the candidate in the work presented in this thesis

Chapter	Components (or associated paper)	Activity	Responsibility	Additional input
Chapter 1	Background	Conceptualisation and writing	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa,
Chapter 2	Thesis aim, objectives, design and structure	Conceptualisation and writing	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa, Vladimir Gordeev,
Chapter 3	Historical review of stillbirth and neonatal deaths measurement in HDSS	Conceptualisation of paper	Joseph Akuze Waiswa, Hannah Blencowe, Vladimir Gordeev, Angela Baschieri, Joy Lawn,	Reviewed by Simon Cousens, Peter Waiswa
		Literature review	Joseph Akuze Waiswa	
		Undertaking systematic reviews of DHS questionnaire modules	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens and Joy Lawn
		Data analysis and development of geo-spatial maps	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens and Joy Lawn
		Drafting of manuscript	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa, Vladimir Gordeev, Angela

				Baschieri, Trevor Croft, Fred Arnold
		Reviewing drafts and approval of the final manuscript	All authors	
Chapter 4	EN-INDEPTH protocol paper	Conceptualisation of paper	Angela Baschieri, Vladimir Gordeev, Joseph Akuze Waiswa, Hannah Blencowe, Joy Lawn,	Reviewed by Doris Kwesiga, Peter Waiswa
		Literature review	Angela Baschieri, Vladimir Gordeev, Joseph Akuze Waiswa	Reviewed by Doris Kwesiga, Peter Waiswa, Joy Lawn, Hannah Blencowe
		Development and compiling of study questionnaires and survey solutions user manuals, ethical clearance applications and standard operating procedural documents	Joseph Akuze Waiswa	Angela Baschieri, Vladimir Gordeev, Joy Lawn, Hannah Blencowe
		Drafting of manuscript	Angela Baschieri, Vladimir Gordeev, Joseph Akuze Waiswa	Reviewed by EN-INDEPTH collaborative group authors
		Reviewing drafts and approval of the final manuscript	All authors	
Chapter 5	A randomised comparison of the measurement of stillbirths and neonatal death	Conceptualisation of paper	Joseph Akuze Waiswa, Hannah Blencowe, Joy Lawn, Peter Waiswa, Angela Baschieri, Vladimir Gordeev	Reviewed by Doris Kwesiga, Peter Waiswa
		Literature review	Joseph Akuze Waiswa	Reviewed by Doris Kwesiga, Peter Waiswa, Joy Lawn, Hannah Blencowe, Simon Cousens

		Data management	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens
		Data analysis	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn
		Drafting of manuscript	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa, Vladimir Gordeev, Angela Baschieri and collaborative group authors
		Reviewing drafts and approval of the final manuscript	All authors	
Chapter 6	Paradata analysis	Data management	Joseph Akuze Waiswa	Reviewed by Vladimir Gordeev
		Conceptualisation	Vladimir Gordeev, Joseph Akuze Waiswa, Hannah Blencowe, Joy Lawn	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa, Angela Baschieri and collaborative group authors
		Data analysis	Vladimir Gordeev	Reviewed by Joseph Akuze Waiswa, Hannah Blencowe, Joy Lawn, Angela Baschieri and collaborative group authors
		Drafting of manuscript	Vladimir Gordeev Joseph Akuze Waiswa	Reviewed by

				Hannah Blencowe, Joy Lawn, Angela Baschieri and collaborative group authors
		Reviewing drafts and approval of the final manuscript	All authors	
Chapter 7	Population -level comparison of EN-INDEPTH Survey and HDSS data	Conceptualisation and writing	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa, Vladimir Gordeev, Angela Baschieri
		Literature review	Joseph Akuze Waiswa	Reviewed by Peter Waiswa, Joy Lawn, Hannah Blencowe, Simon Cousens
		Data management	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens
		Data analysis	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn
		Drafting of manuscript	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa, Vladimir Gordeev, Angela Baschieri, Ane Fisker, Moinuddin M Haider, Francis Dzabeng and collaborative group authors

		Reviewing drafts and approval of the final manuscript	All authors	
Chapter 8	Discussion, conclusions, and recommendations	Conceptualisation and writing	Joseph Akuze Waiswa	Reviewed by Hannah Blencowe, Simon Cousens, Joy Lawn, Peter Waiswa, Vladimir Gordeev, Angela Baschieri

A2: Supplementary methods, analyses and results to Chapter 7

Table S 1: Definition of independent variables to be used in this analysis

	Variable	*Definition*	*Source*	*Scale*
1	*Survey module*	*Survey module that the respondent was randomised to during the interview*	*Household survey*	*Categorical (Full Birth History, Full Pregnancy History)*
2	*HDSS site*	*Health and demographic surveillance site to which the woman belongs to*	*Household survey*	*Categorical (Bandim, Dabat, Iganga-Mayuge, Matlab and Kintampo)*
3	*Age*	*Age in years on the date of the survey*	*Household survey*	*Continuous and Categorical (15-19, 20-24, 25-29, 30-34, and 35+)*
4	*Education level*	*Level of education attained by respondent*	*Household survey*	*Categorical (No education, Primary only, Primary & Secondary, Higher and Missing)*
5	*Socioeconomic status*	*Socioeconomic status or classification of respondent based on recorded household assets*	*Household survey*	*Categorical (Poorest, 2, 3, 4, Richest)*
6	*Parity*	*The number of times that a respondent has given birth to a foetus with a gestational age of 7 months or more, regardless of whether the child was born alive or was stillborn*	*Household survey*	*Categorical (0,1,2,3,4, and 5+)*
7	*Gender*	*Pregnancy outcome's gender*	*Household survey*	*Categorical (Male, Female)*

8	*Gender*	*Interviewer's gender*	*Interviewer survey*	*Categorical (Male, Female)*
9	*Age*	*Interviewer's age in years*	*Interviewer survey*	*Continuous and Categorical (15-19, 20-24, 25-29, 30-34, and 35+)*
10	*DHS experience*	*Previous experience with DHS survey*	*Interviewer survey*	*Dichotomous (Yes, No)*
11	*Non-DHS experience*	*Previous experience with other non-DHS survey*	*Interviewer survey*	*Dichotomous (Yes, No)*

Table S 2: Questions for assessing the HDSS system for pregnancy surveillance

Questions	**Answers**
1. Give an overview of your HDSS	
a. Population	
b. Households	
c. Live births	
d. Stillbirths	
e. Stillbirth Rate	
f. Neonatal deaths	
g. Neonatal Mortality Rate	
h. Total Fertility Rate	
i. In Migration rate	
j. Out migration rate	
2. Who carries out routine surveillance visits?	
3. Who is allowed to be the primary respondent during visits?	
4. How often are routine surveillance visits made?	
5. How are vital events notified and followed up?	
6. What happens when a birth captured when the pregnancy was not previously recorded?	
7. How can pregnancy surveillance and outcome (births/vital events) data be linked?	

8. What percentage of births is in a facility?	
9. Is facility birth data linked to HDSS?	
10. Do you think your HDSS is missing pregnancies? If so, which ones?	
11. How are induced vs spontaneous abortions defined?	
12. How are stillbirths defined? Any difference between MSB/FSB?	
13. Are there probes about whether a baby that died moved/cried/ breathed at all (not just in VA)?	
14. What are the specific local barriers around reporting pregnancy loss, if any, that may impede reporting to surveyors/enumerators	
15. Does your site reach all women of childbearing age with each surveillance round? If not, how often?	
16. What is the percentage of out-migration? Is there birth-related migration out of the HDSS?	
1. Measuring birth weight a. Any data on birth weight? b. Is the data from facility records only or is weight taken at home by HDSS staff? c. Is it captured in routine surveillance visits? d. Are women asked about perceived size of baby at birth? e. Is the data available as a subset of the routine surveillance or sub studies?	
2. Measuring Gestational Age a. Is there data on gestational age at delivery? b. If so, what data exist? c. Is it captured in routine surveillance visits or linked to facility data? d. Are the data available as a subset of the routine surveillance or sub studies?	
17. What are possible enhancements to the pregnancy surveillance system and reporting?	
18. Challenges expected in harmonisation of data in your site?	

Table S 3: Reasons for exclusion of data by year from the supplementary analysis

Reasons for exclusion	Sites and years whose data was excluded for supplementary analysis			
	Dabat	IgangaMayuge	Matlab	Kintampo
Data loss associated with migration from OpenHDS system to Open Data Kit	ü(2017 and 2018)			

Errors in data which hindered linkage of the various HDSS database tables	(2012, 2013, 2017 and 2018)			
Funding challenges which led to changes in pregnancy surveillance		(2017 and 2018)		(2016, 2017 and 2018)
Administrative, Logistical bureaucracies and delays which led to changes in pregnancy surveillance		(2017 and 2018)		
Incomplete data years	(2017 and 2018)	(2017 and 2018)	(2017 and 2018)	(2016, 2017 and 2018)

Table S 4: Overview of HDSS in five EN-INDEPTH study sites

Availability of outcome indicators					
	Bandim	**Dabat**	**IgangaMayuge**	**Matlab**	**Kintampo**
Stillbirths	Yes	Yes	Yes	Yes	Yes
Neonatal deaths	Yes	Yes	Yes	Yes	Yes
Birthweight	Yes - Urban only	Yes	Yes	Yes	Yes – Since 2015
Source of birthweight data	Recorded from health cards during household visits	Recorded from health cards during household visits and from facilities where pilot is happening	Recorded from health cards during household visits	Recorded at facilities and in few cases, it is recorded from health cards in households	Recorded from health cards during household visits
Perceived size of baby at birth	No	Yes since 2014	Yes	No	Yes
Gestation Age (GA)	Yes – at pregnancy registration	Yes	Yes	Yes	Yes
Birth certificate	Yes – On a subset around 400 (16%) women[2]	No	No	Yes – asked for all births in the HDSS	No
Death certificate	Yes – asked with verbal autopsy	No	Yes – asked with verbal autopsy	Yes – asked with verbal autopsy	Yes – asked with verbal autopsy

Causes of death	Yes – using verbal autopsy	Yes – using verbal autopsy	Yes – using verbal autopsy	Yes – using verbal autopsy ICD 10 codes are assigned	Yes – using verbal autopsy
Miscarriages/ Abortions	Yes - Miscarriages only	Yes - Abortions differentiated from miscarriages	Yes - Miscarriages only	Yes - Abortions differentiated from miscarriages	Yes - Abortions differentiated from miscarriages, however, no data on miscarriages is collected
Enhancements and challenges affecting reporting and data quality of pregnancies faced by HDSS sites					
Enhancements towards improving capture of pregnancies and outcomes	**Urban:** Training of field workers to improve pregnancy registration; Linking of HDSS data with facility data **Rural:** Improving outcome classification through calculating GA based on data captured during pregnancy registration	Improving capture of birth outcomes through pregnancy observations Adapting tested and standardised tools to improve measurement of GA Using standardised and consistent definitions e.g. stillbirths	Using health campaigns, camps and other incentives to pregnant women to encourage them to report pregnancies and pregnancy outcomes	Matlab has a well-established system to capture early pregnancies and outcomes, therefore, no enhancements towards improving capture of pregnancies and outcomes were suggested.	Early identification and recording of pregnancies during first trimester. Using tested and standardised tools for improved measurement of GA Overcoming socio-cultural hindrances for stillbirth reporting
Challenges in harmonisation and comparison of data on pregnancies	Misclassification between stillbirths and neonatal deaths Missing data when a pregnancy is not registered and assigned a pregnancy ID	Misclassification between stillbirths and neonatal deaths Missing data and information on some indicators	Misclassification between stillbirths and neonatal deaths Missing data and information on some indicators	Misclassification between stillbirths (since 2012) and neonatal deaths (since 2003)	Misclassification between stillbirths and neonatal deaths Missing data and information on some indicators

Neonatal deaths (Excluding problematic years)

Additional supplementary analysis excluding pregnancy events that occurred in site-years with identified data collection issues included 31,618 and 36,350 livebirths recorded in the HDSS and survey, respectively, with 658 neonatal deaths reported in HDSS and 899 in the survey (Supplementary Tables 5a and 5b). The estimated NMR ranged from 24 to 29.7 per 1,000 livebirths for survey data and 18.2 to 22.2 per 1,000 livebirths for HDSS data among women in the FBH+ arm (Figure S3a). Among women in the FPH arm, the NMR ranged from 22.8 to 24.9 per 1000 livebirths for survey data and ranged from 11.8 to 22.1 per 1000 livebirths for HDSS data (Figure 3a). Similar to the primary analyses, we found a higher NMR in the survey compared to the HDSS with a difference of 23% (95% CI (4% – 45%)) among women in the FBH+ arm and 20% (95% CI (-5% – 50%)) among women in the FPH arm. There was no evidence of heterogeneity between sites for the HDSS: survey NMR ratio in either the FBH+ (p=0.297, I^2 =18.6%) or FPH group (P=0.135, I^2= 46.1%) (Figure S3a).

Stillbirths (Excluding problematic years)

Additional supplementary analysis excluding pregnancy events that occurred in site-years with identified data collection comparability issues, we found 32,130 and 36,902 total births recorded in the HDSS and survey, respectively, with 512 stillbirths reported in HDSS and 552 in the survey (Tables S5a and S5b). The estimated SBR ranged from 5.9 to 17.3 per 1,000 total births for survey data and from 3.5 to 21.4 per 1,000 total births for HDSS data among women interviewed with FBH+ in the survey (Figure S3b). Among women interviewed with FPH, the SBR ranged from 11.4 to 20.8 per 1000 total births for survey data and ranged from 5.9 to 25.5 per 1000 total births for HDSS data (Figure 3b). Similar to the primary analyses, we found little difference in reported SBR between survey and HDSS 2% (95% CI (-29% – 45%)) among women in the FBH+ group and 5% (95% CI (-36% – 70%)) among women who belong to the FPH group. There was some evidence of heterogeneity between sites for the HDSS: survey SBR ratio in either the FBH+ (p=0.080, I^2 =55.7%) and strong evidence of between site heterogeneity FPH group (p<=0.0001, I^2= 83.6%) (Figure S3b).

Table S 5a: Comparing pregnancy outcome between EN-INDEPTH survey and HDSS by module and year

	Dabat				IgangaMayuge				Matlab				Kintampo			
	FBH+		FPH		FBH+		FPH		FBH+		FPH		FBH+		FPH	
No. of women interviewed in EN-INDEPTH	6,327		6,266		6,788		6,649		10,809		10,653		6,049		6,143	
No. women who were continuously resident in the HDSS*	5,752		5,686		4,725		4,677		10,796		10,636		3,089		3,103	
% of women interviewed who were continuously resident	91		91		70		70		100		100		51		50.5	
	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey
No. women with 1> pregnancies ending after 1st January 2012*	**2,322**	**3,083**	**2,282**	**3,103**	**1,494**	**1,882**	**1,527**	**1,904**	**10,796**	**10,559**	**10,363**	**10,067**	**2,976**	**3,068**	**2,983**	**3,077**
Total pregnancies recorded*	**2,619**	**4,769**	**2,582**	**4,770**	**2,341**	**3,260**	**2,380**	**3,340**	**14,092**	**13,713**	**13,880**	**12,495**	**3,604**	**5,391**	**3,607**	**5,453**
2012	*447*	*744*	*433*	*737*	*518*	*591*	*506*	*588*	*1,889*	*1,890*	*1,840*	*1,737*	*650*	*787*	*638*	*740*
2013	*417*	*729*	*411*	*711*	*400*	*544*	*437*	*571*	*2,181*	*2,236*	*2,262*	*2,196*	*661*	*913*	*688*	*927*
2014	*417*	*718*	*429*	*733*	*372*	*514*	*361*	*562*	*2,349*	*2,363*	*2,280*	*2,113*	*786*	*866*	*816*	*930*
2015	*610*	*724*	*640*	*758*	*378*	*509*	*395*	*545*	*2,556*	*2,567*	*2,428*	*2,265*	*862*	*991*	*840*	*956*
2016	*641*	*770*	*608*	*730*	*348*	*524*	*352*	*507*	*2,663*	*2,598*	*2,628*	*2,345*	*585*	*980*	*561*	*978*
2017	*87*	*808*	*61*	*805*	*292*	*507*	*293*	*499*	*2,137*	*2,011*	*2,131*	*1,802*	*60*	*767*	*64*	*835*
2018	*-*	*276*	*-*	*296*	*33*	*71*	*36*	*68*	*317*	*48*	*311*	*37*	*-*	*87*	*-*	*87*
Total live births recorded*#	**2,411**	**4,648**	**2,375**	**4,644**	**2,330**	**3,051**	**2,363**	**3,039**	**12,046**	**12,068**	**11,866**	**11,862**	**3,492**	**5,001**	**3,487**	**4,966**
2012	*433*	*735*	*421*	*723*	*513*	*564*	*504*	*540*	*1,527*	*1,677*	*1,558*	*1,639*	*335*	*736*	*333*	*661*
2013	*389*	*717*	*371*	*694*	*398*	*525*	*431*	*529*	*1,862*	*1,980*	*1,943*	*2,094*	*640*	*848*	*652*	*852*
2014	*289*	*702*	*313*	*719*	*371*	*492*	*360*	*520*	*2,015*	*2,084*	*1,976*	*2,019*	*682*	*801*	*687*	*855*
2015	*594*	*711*	*618*	*741*	*377*	*478*	*391*	*496*	*2,190*	*2,238*	*2,075*	*2,140*	*810*	*927*	*770*	*881*
2016	*625*	*738*	*592*	*706*	*348*	*470*	*348*	*456*	*2,272*	*2,305*	*2,219*	*2,226*	*753*	*927*	*750*	*896*

	Dabat				IgangaMayuge				Matlab				Kintampo			
	FBH+		FPH		FBH+		FPH		FBH+		FPH		FBH+		FPH	
2017	*81*	*782*	*60*	*779*	*290*	*459*	*292*	*435*	*1,855*	*1,741*	*1,838*	*1,709*	*283*	*689*	*300*	*750*
2018	*-*	*263*	*-*	*282*	*32*	*63*	*36*	*63*	*281*	*44*	*257*	*35*	*-*	*73*	*-*	*71*
Number of post-neonatal survivors	**2,344**	**4,524**	**2,321**	**4,525**	**2,277**	**2,956**	**2,315**	**2,968**	**11,801**	**11,801**	**11,621**	**11,586**	**3,435**	**4,878**	**3,422**	**4,853**
2012	*419*	*713*	*412*	*705*	*504*	*544*	*496*	*520*	*1,483*	*1,623*	*1,521*	*1,593*	*321*	*712*	*322*	*647*
2013	*375*	*698*	*358*	*671*	*393*	*513*	*420*	*518*	*1,829*	*1,937*	*1,913*	*2,051*	*627*	*826*	*635*	*832*
2014	*284*	*684*	*310*	*701*	*359*	*477*	*353*	*509*	*1,974*	*2,038*	*1,923*	*1,962*	*673*	*775*	*675*	*828*
2015	*587*	*694*	*610*	*727*	*369*	*466*	*383*	*488*	*2,133*	*2,186*	*2,035*	*2,093*	*801*	*906*	*757*	*868*
2016	*608*	*720*	*585*	*684*	*338*	*454*	*337*	*448*	*2,228*	*2,253*	*2,174*	*2,174*	*749*	*914*	*741*	*877*
2017	*77*	*763*	*55*	*763*	*282*	*442*	*290*	*424*	*1,834*	*1,721*	*1,804*	*1,679*	*280*	*676*	*297*	*732*
2018	*-*	*252*	*-*	*274*	*31*	*60*	*35*	*61*	*276*	*44*	*251*	*34*	*-*	*69*	*-*	*69*
Number of stillbirths recorded*	**11**	**43**	**27**	**48**	**9**	**22**	**13**	**62**	**211**	**209**	**192**	**168**	**67**	**73**	**77**	**104**
2012	*-*	*5*	*-*	*4*	*4*	*4*	*2*	*12*	*37*	*35*	*27*	*32*	*9*	*12*	*12*	*20*
2013	*-*	*4*	*-*	*6*	*2*	*-*	*5*	*11*	*32*	*33*	*32*	*24*	*20*	*17*	*20*	*23*
2014	*1*	*3*	*4*	*7*	*1*	*4*	*-*	*8*	*38*	*41*	*33*	*25*	*17*	*10*	*18*	*12*
2015	*7*	*5*	*13*	*8*	*-*	*5*	*1*	*8*	*36*	*39*	*35*	*32*	*8*	*11*	*14*	*14*
2016	*-*	*13*	*10*	*10*	*-*	*2*	*4*	*11*	*36*	*33*	*34*	*28*	*7*	*7*	*11*	*20*
2017	*-*	*12*	*-*	*8*	*2*	*5*	*1*	*11*	*28*	*27*	*21*	*26*	*6*	*15*	*2*	*13*
2018	*-*	*1*	*-*	*5*	*-*	*2*	*-*	*1*	*4*	*1*	*10*	*1*	*-*	*1*	*-*	*2*
Number of early neonatal deaths (0-6 days) recorded*	**37**	**80**	**29**	**92**	**48**	**78**	**40**	**60**	**204**	**227**	**198**	**223**	**44**	**95**	**49**	**88**
2012	*10*	*12*	*6*	*13*	*9*	*18*	*7*	*18*	*38*	*46*	*33*	*38*	*12*	*17*	*8*	*14*
2013	*10*	*12*	*9*	*18*	*3*	*9*	*10*	*9*	*27*	*35*	*21*	*36*	*13*	*16*	*13*	*17*

	Dabat				IgangaMayuge				Matlab				Kintampo			
	FBH+		FPH		FBH+		FPH		FBH+		FPH		FBH+		FPH	
2014	*2*	*11*	*1*	*14*	*11*	*12*	*5*	*8*	*33*	*40*	*46*	*51*	*8*	*23*	*11*	*22*
2015	*4*	*12*	*5*	*10*	*7*	*11*	*6*	*7*	*52*	*47*	*32*	*37*	*4*	*14*	*10*	*7*
2016	*8*	*14*	*4*	*18*	*10*	*13*	*9*	*6*	*35*	*44*	*36*	*40*	*4*	*12*	*4*	*12*
2017	*3*	*13*	*4*	*14*	*7*	*14*	*2*	*10*	*16*	*15*	*24*	*20*	*3*	*10*	*3*	*14*
2018	*-*	*6*		*5*	*1*	*1*	*1*	*2*	*3*	*-*	*6*	*1*	*-*	*3*	*-*	*2*
Number of late neonatal deaths (7-27 days) recorded*	**24**	**44**	**16**	**27**	**5**	**17**	**8**	**11**	**41**	**40**	**47**	**53**	**8**	**28**	**16**	**25**
2012	*4*	*10*	*3*	*5*	*-*	*2*	*1*	*2*	*6*	*8*	*4*	*8*	*2*	*7*	*3*	*-*
2013	*4*	*7*	*4*	*5*	*2*	*3*	*1*	*2*	*6*	*8*	*9*	*7*	*-*	*6*	*4*	*3*
2014	*3*	*7*	*2*	*4*	*1*	*3*	*2*	*3*	*8*	*6*	*7*	*6*	*1*	*3*	*1*	*5*
2015	*3*	*5*	*3*	*4*	*1*	*1*	*2*	*1*	*5*	*5*	*8*	*10*	*5*	*7*	*3*	*6*
2016	*9*	*4*	*3*	*4*	*-*	*3*	*2*	*2*	*9*	*8*	*9*	*12*	*-*	*1*	*5*	*7*
2017	*1*	*6*	*1*	*2*	*1*	*3*	*-*	*1*	*5*	*5*	*10*	*10*	*-*	*3*	*-*	*4*
2018		*5*	*-*	*3*	*-*	*2*	*-*	*-*	*2*	*-*	*-*	*-*	*-*	*1*	*-*	*-*
Number of TOPs recorded*	**NA**	**NA**	**NA**	**1**	**NA**	**NA**	**NA**	**23**	**517**	**NA**	**522**	**97**	**61**	**NA**	**65**	**56**
2012	*NA*	*NA*	*NA*	*-*	*NA*	*NA*	*NA*	*1*	*70*	*NA*	*71*	*12*	*3*	*NA*	*11*	*16*
2013	*NA*	*NA*	*NA*	*-*	*NA*	*NA*	*NA*	*-*	*97*	*NA*	*86*	*20*	*13*	*NA*	*20*	*12*
2014	*NA*	*NA*	*NA*	*-*	*NA*	*NA*	*NA*	*2*	*77*	*NA*	*87*	*12*	*22*	*NA*	*19*	*6*
2015	*NA*	*NA*	*NA*	*-*	*NA*	*NA*	*NA*	*6*	*98*	*NA*	*96*	*21*	*11*	*NA*	*9*	*5*
2016	*NA*	*NA*	*NA*	*1*	*NA*	*NA*	*NA*	*4*	*99*	*NA*	*102*	*15*	*10*	*NA*	*2*	*10*
2017	*NA*	*NA*	*NA*	*-*	*NA*	*NA*	*NA*	*9*	*65*	*NA*	*66*	*17*	*2*	*NA*	*4*	*6*
2018	*NA*	*NA*	*NA*	*-*	*NA*	*NA*	*NA*	*1*	*11*	*NA*	*14*	*-*	*-*	*NA*	*-*	*1*

	Dabat				IgangaMayuge				Matlab				Kintampo			
	FBH+		FPH		FBH+		FPH		FBH+		FPH		FBH+		FPH	
Number of miscarriages recorded*	**4**	**78**	**1**	**77**	**3**	**187**	**5**	**216**	**1,318**	**1,436**	**1,300**	**368**	**NA**	**317**	**NA**	**327**
2012	*-*	*4*	*-*	*10*	*1*	*23*	*-*	*35*	*211*	*178*	*184*	*54*	*NA*	*39*	*NA*	*43*
2013	*-*	*8*	*-*	*11*	*-*	*19*	*1*	*31*	*190*	*223*	*201*	*58*	*NA*	*48*	*NA*	*40*
2014	*1*	*13*	*-*	*7*	*-*	*18*	*1*	*32*	*219*	*283*	*184*	*57*	*NA*	*55*	*NA*	*57*
2015	*1*	*8*	*1*	*9*	*1*	*26*	*3*	*35*	*232*	*290*	*222*	*72*	*NA*	*53*	*NA*	*56*
2016	*-*	*19*	*-*	*13*	*-*	*52*	*-*	*36*	*256*	*260*	*273*	*76*	*NA*	*46*	*NA*	*52*
2017	*2*	*14*	*-*	*18*	*-*	*43*	*-*	*44*	*189*	*244*	*206*	*50*	*NA*	*63*	*NA*	*66*
2018	*-*	*12*	*-*	*9*	*1*	*6*	*-*	*3*	*21*	*3*	*30*	*1*	*NA*	*13*	*NA*	*13*

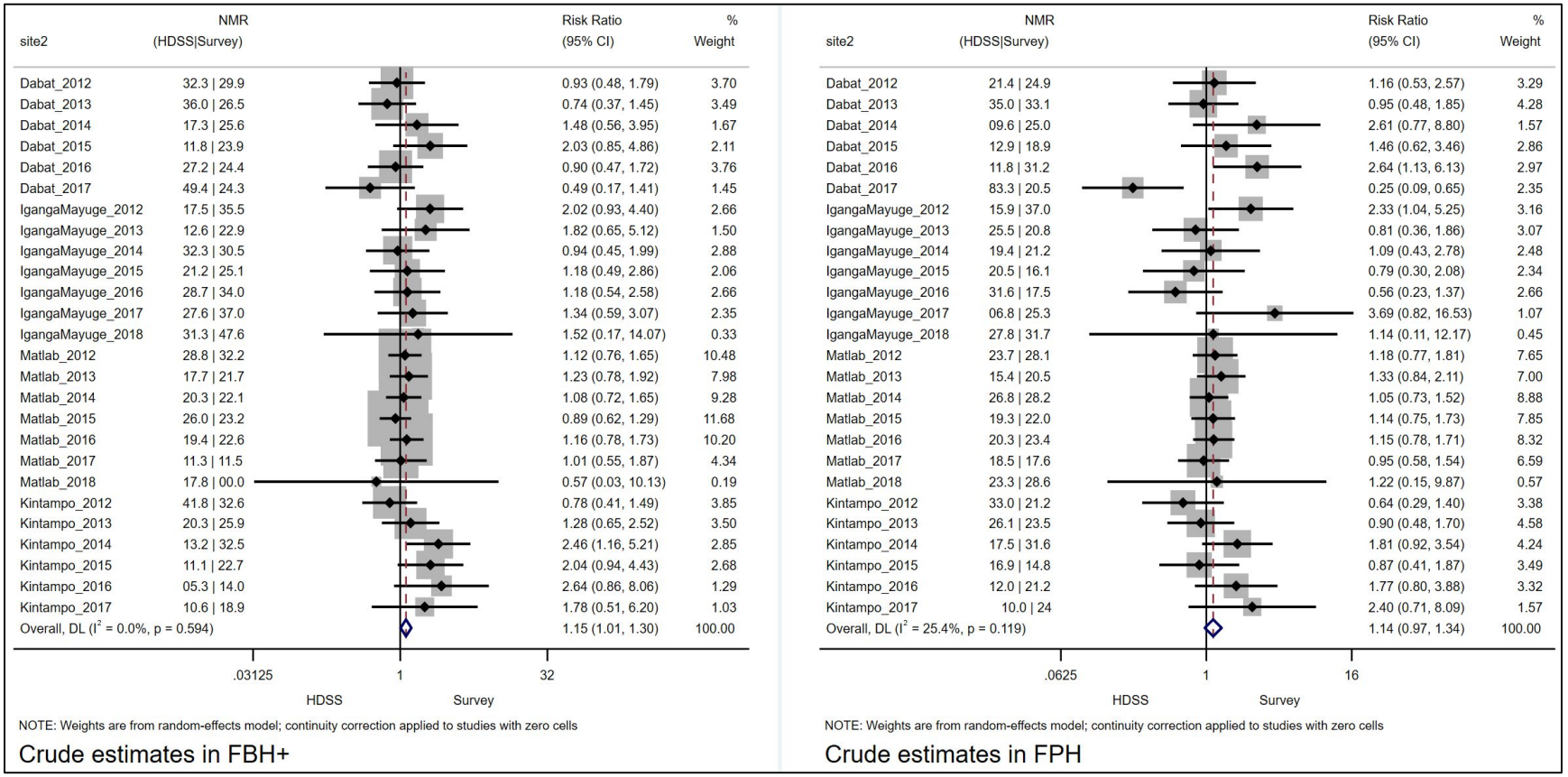
** Among women who were continuously resident in the HDSS between 1 January 2012 and the survey interview date; # Number of live births includes (Number of post-neonatal survivors plus number of neonatal deaths); The highlighted years indicate the problematic years in all sites, that were excluded in supplementary analysis i.e., occurred 2012,2013, 2017 and 2018 in Dabat, occurred 2017 and 2018 in IgangaMayuge, occurred 2017 and 2018 in Matlab and occurred 2016, 2017 and 2018 in Kintampo*

Table S 5b: Comparison of pregnancy outcome indicators between EN-INDEPTH survey and HDSS by module (Truncated)

	Dabat				IgangaMayuge				Matlab				Kintampo			
	FBH+		FPH		FBH+		FPH		FBH+		FPH		FBH+		FPH	
No. of women interviewed in EN-INDEPTH	6327		6266		6788		6649		10809		10653		6049		6143	
No. women who were continuously resident in the HDSS*	5752		5686		4725		4677		10796		10636		3089		3103	
% of women interviewed who were continuously resident	90.9		90.7		69.6		70.3		99.9		99.8		51.1		50.5	
	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey	HDSS	Survey
No. women with 1> pregnancies ending after 1st January 2012*	2322	3083	2282	3103	1494	1882	1527	1904	10796	10559	10363	10067	2976	3068	2983	3077
Total pregnancies recorded*	1668	2212	1677	2221	2016	2682	2051	2773	11638	11654	11438	10656	2959	3557	2982	3553
Total live births recorded*#	1508 (90.4)	2151 (97.2)	1523 (90.8)	2166 (97.5)	2007 (99.6)	2529 (94.3)	2034 (99.2)	2541 (91.6)	9866 (84.8)	10284 (88.2)	9771 (85.4)	10118 (95)	2467 (83.4)	3312 (93.1)	2442 (81.9)	3249 (91.4)
Number of post-neonatal survivors	1479	2098	1505	2112	1963	2454	1989	2483	9647	10037	9566	9873	2422	3219	2389	3175
Number of stillbirths recorded*	8 (0.5)	21 (0.9)	27 (1.6)	25 (1.1)	7 (0.3)	15 (0.6)	12 (0.6)	50 (1.8)	179 (1.5)	181 (1.6)	161 (1.4)	141 (1.3)	54 (1.8)	50 (1.4)	64 (2.1)	69 (1.9)
Number of early neonatal deaths (0-6 days) recorded*	14	37	10	42	40	63	37	48	185	212	168	202	37	70	42	60
Number of late neonatal deaths (7-27 days) recorded*	15	16	8	12	4	12	8	10	34	35	37	43	8	23	11	14
Number of TOPs recorded*	NA	NA	NA	1 (0)	NA	NA	NA	13 (0.5)	441 (3.8)	NA	442 (3.9)	80 (0.8)	49 (1.7)	NA	59 (2)	39 (1.1)
Number of miscarriages recorded*	2 (0.1)	40 (1.8)	1 (0.1)	29 (1.3)	2 (0.1)	138 (5.1)	5 (0.2)	169 (6.1)	1108 (9.5)	1234 (10.6)	1064 (9.3)	317 (2.9)	NA	195 (8.8)	NA	196 (5.5)
Stillbirth Rate (per 1000 total births)	5.3	9.7	17.4	11.4	3.5	5.9	5.9	19.3	17.8	17.3	16.2	13.7	21.4	14.9	25.5	20.8
Neonatal mortality rate (per 1000 live births)	19.2	24.6	11.8	24.9	21.9	29.7	22.1	22.8	22.2	24	21	24.2	18.2	28.1	21.7	22.8

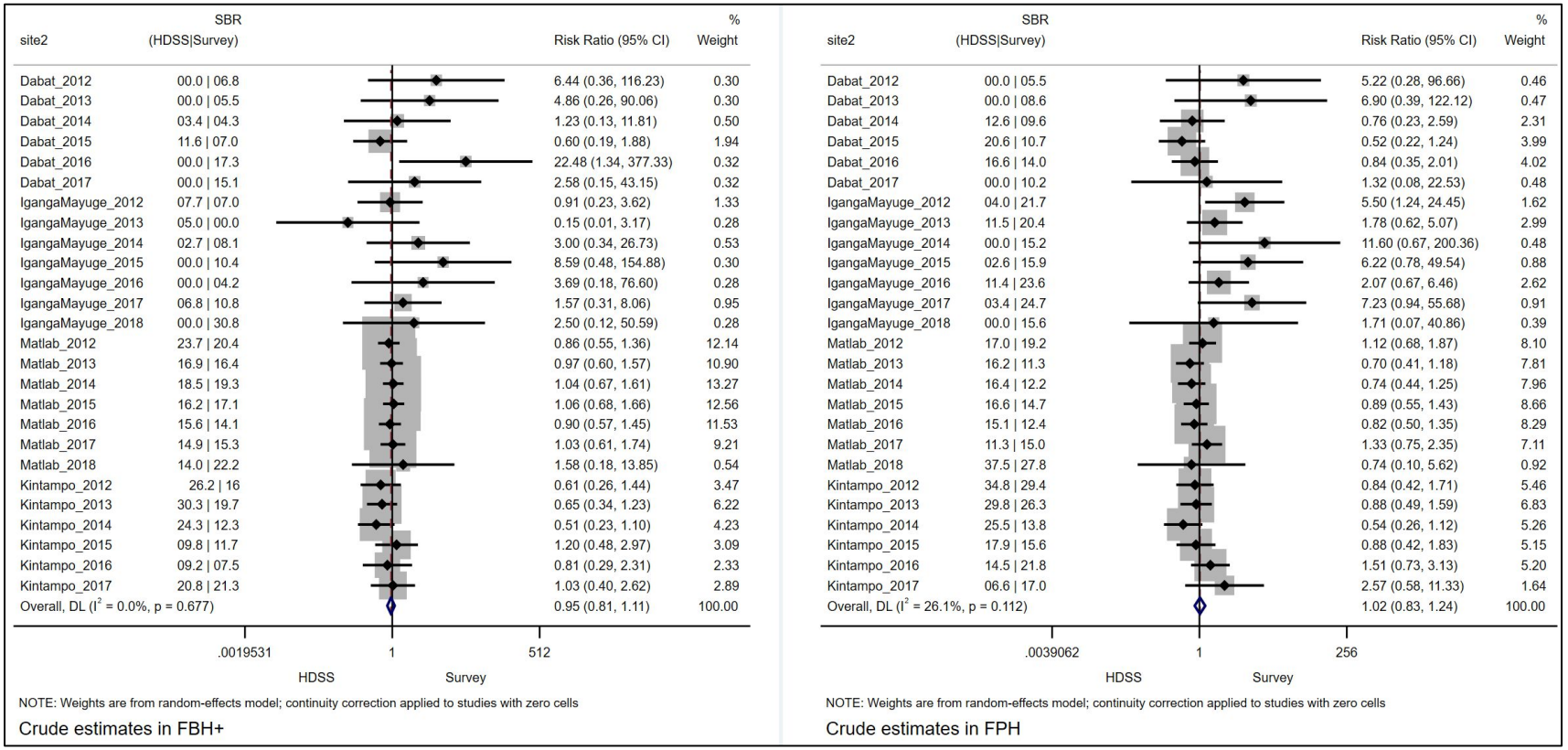
** Among women who were continuously resident in the HDSS between 1 January 2012 and the survey interview date; # Number of live births includes (Number of post-neonatal survivors plus number of neonatal deaths) ; In this analysis, we excluded pregnancies from some years in all sites, i.e., pregnancies that occurred 2012,2013, 2017 and 2018 in Dabat, occurred 2017 and 2018 in IgangaMayuge, occurred 2017 and 2018 in Matlab and occurred 2016, 2017 and 2018 in Kintampo*

Figure S 1a: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to HDSS by year



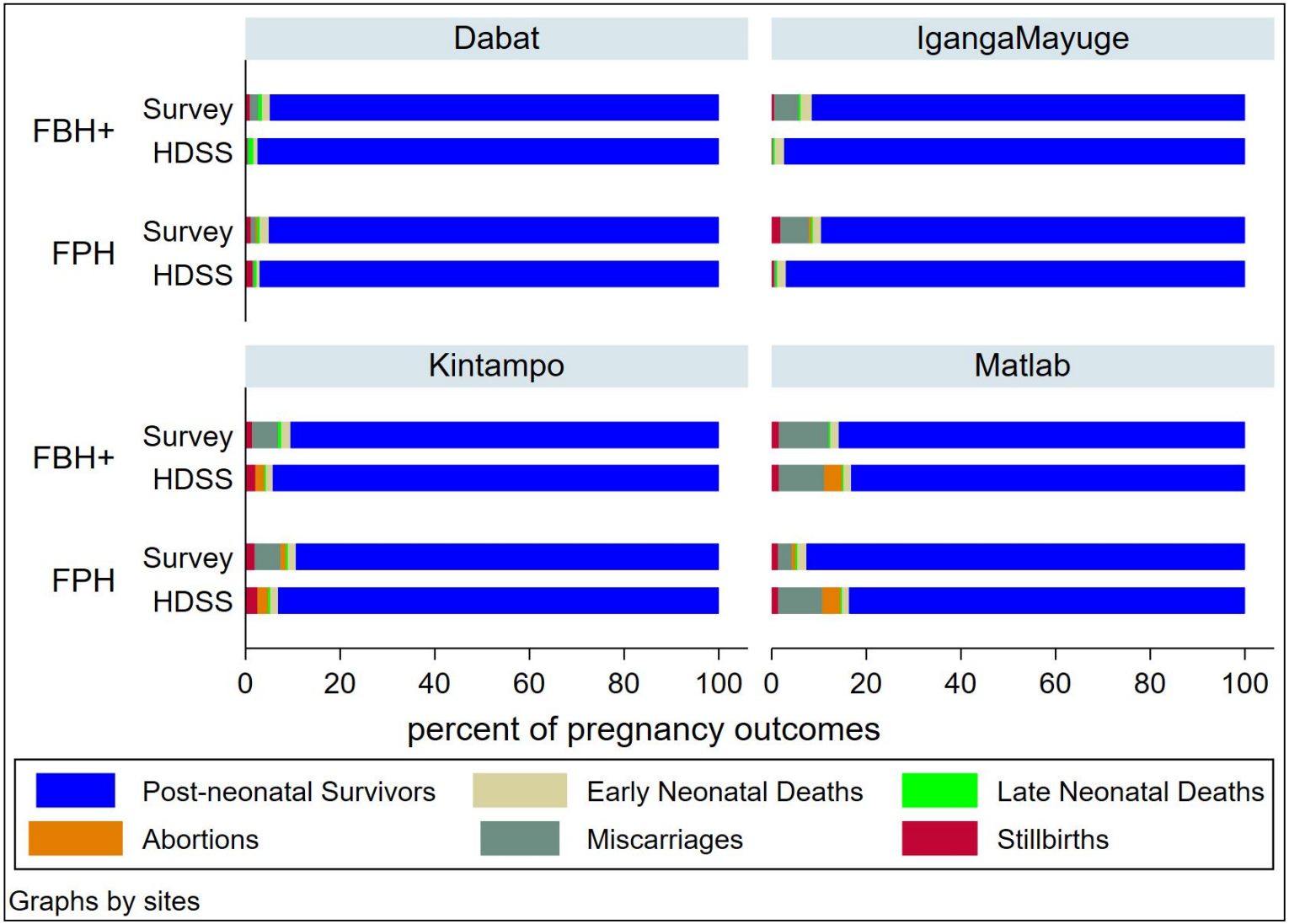
++ Used a simplified assumption that the HDSS and survey women are independent. Analysis was among women who are continuous residents since 1st January 2012

Figure S1b: Ratio of Stillbirth Rate in EN-INDEPTH survey compared to HDSS by year



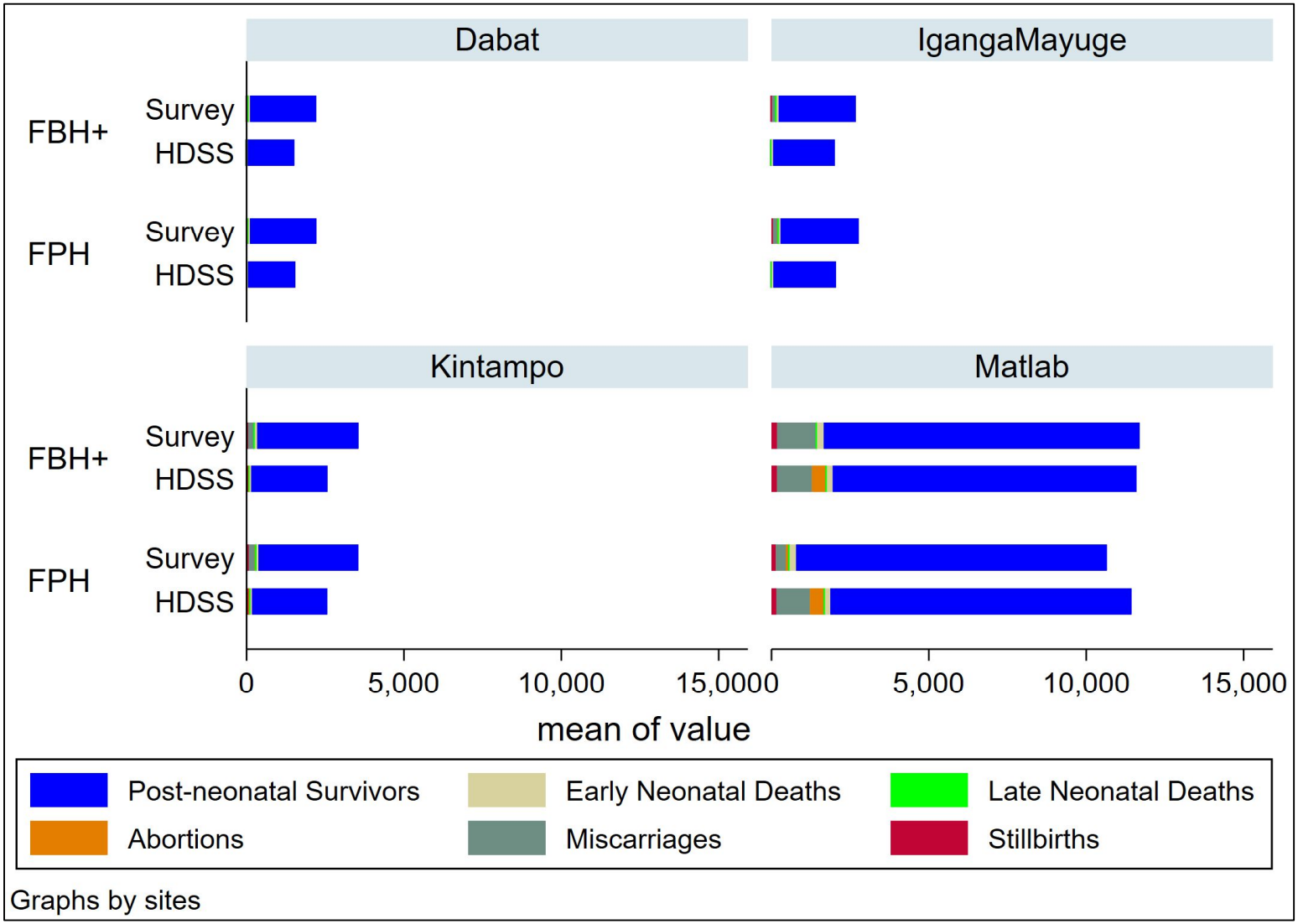
++ Used a simplified assumption that the HDSS and survey women are independent. Analysis was among women who are continuous residents since 1st January 2012

Figure S 2a: Distribution of pregnancy outcomes between EN- INDEPTH survey and HDSS by module



In Dabat, data from 2012-2013 and 2017-2018 are excluded, IgangaMayuge and Matlab 2017-2018 data are excluded and Kintampo 2016-2018 data are excluded.

Figure S2b: Distribution of pregnancy outcomes between EN- INDEPTH survey and HDSS by module

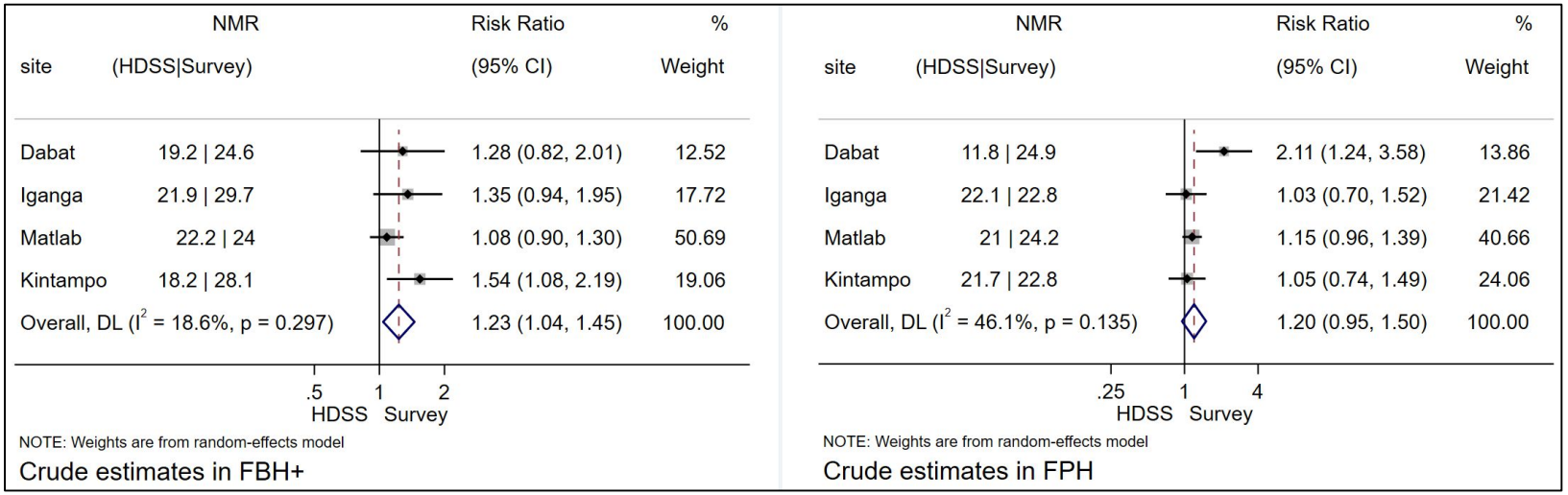


Annual estimates

Annual estimates for NMR were generally higher in the survey compared to the HDSS. The survey recorded higher NMR compared to the HDSS with a difference of 15% (95% CI (1% – 30%)) among women in the FBH+ group and 14% (95% CI (-3% – 34%)) among women who belong to the FPH group. There was no evidence of heterogeneity between sites for the HDSS: survey NMR ratio in either the FBH+ ((p=0.594) I^2 =0.0%) or FPH group ((P=0.119) I^2= 25.4%) (Figure S2a).

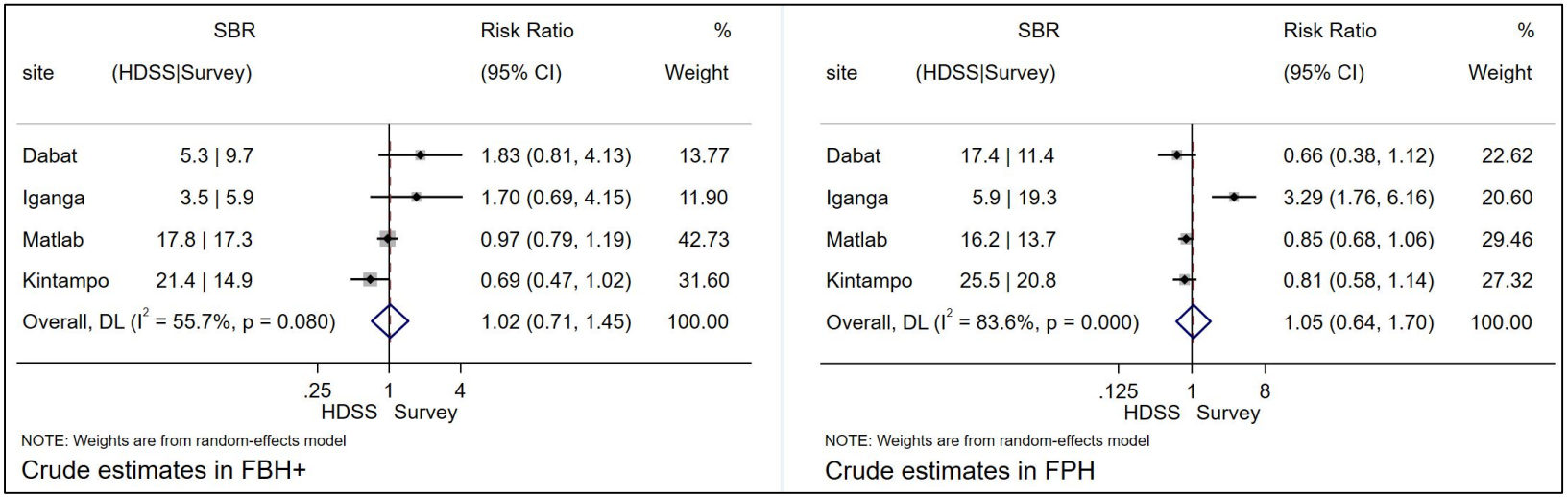
Annual estimates for SBR were generally similar between the survey the HDSS data. The difference in recorded SBR between HDSS and survey was -5% (95% CI (-19% – 11%)) among women in the FBH+ group and 2% (95% CI (-13% – 24%)) among women who belong to the FPH group. There was no evidence of heterogeneity between sites for the HDSS: survey SBR ratio in either the FBH+ ((p=0.677) I^2 =0.0%) or FPH group ((P=0.112) I^2= 26.1%) (Figure S2b).

Figure S 3a: Ratio of Neonatal Mortality Rate in EN-INDEPTH survey compared to HDSS++



++ Used a simplified assumption that the HDSS and survey women are independent. Analysis was among women who are continuous residents since 1st January 2012; In this analysis, we excluded pregnancies from some years in all sites, i.e., occurred 2012,2013, 2017 and 2018 in Dabat, occurred 2017 and 2018 in IgangaMayuge, occurred 2017 and 2018 in Matlab and occurred 2016, 2017 and 2018 in Kintampo

Figure S 3b: Ratio of stillbirth rate in EN-INDEPTH survey compared to HDSS++



++ Used a simplified assumption that the HDSS and survey women are independent. Analysis was among women who are continuous residents since 1st January 2012; In this analysis, we excluded pregnancies from some years in all sites, i.e., occurred 2012,2013, 2017 and 2018 in Dabat, occurred 2017 and 2018 in IgangaMayuge, occurred 2017 and 2018 in Matlab and occurred 2016, 2017 and 2018 in Kintampo.

Comparison with national estimates

We compared SBR and NMR estimates from these analyses to national estimates. Although estimates from the sites are not representative of the entire country, the national estimate can provide a benchmark to compare our estimates, to understand variation which may be true epidemiological variation or due to data factors.

National NMR estimates were generally higher than the estimates obtained from with HDSS or survey by module group except in Uganda Ghana (Figures S4a and S4b).

National SBR estimates for the countries were generally higher than the estimates obtained from the HDSS or either survey module except in Uganda and Ghana (Figures 5a and 5b).

Although sites are not nationally representative of the countries' situation, we found that the SBR estimates from the survey were much lower than the respective country SBR estimates

A3: Ethical approvals

London School of Hygiene and Tropical Medicine: For Project and PhD study

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

Observational / Interventions Research Ethics Committee

Prof Joy Lawn
Professor & Director of MARCH
Department of Infectious Disease Epidemiology (IDE)
Epidemiology and Population Health (EPH)
LSHTM

24 May 2017

Dear Joy

Study Title: A randomized comparison of household survey modules for measuring pregnancy outcomes in five INDEPTH HDSS sites.

LSHTM Ethics Ref: 12218

Thank you for responding to the Observational Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document Type	File Name	Date	Version
Investigator CV	CV_Gordeev 06 2016 copy	28/02/2017	v1
Investigator CV	cv_blencowe	28/02/2017	v1
Investigator CV	Dr Angela Baschieri CV_January_2017	28/02/2017	v1
Investigator CV	JOSEPH CV 1-Jan2017	28/02/2017	v1
Investigator CV	DORIS KWESIGA CV 2017	28/02/2017	v1
Investigator CV	Waiswa Biosketch_NIH_updated Sept 2016	28/02/2017	v1
Investigator CV	CV Joy Lawn june 2016	28/02/2017	v1
Protocol / Proposal	DHS7-Womans-QRE-EN_INDEPTH_survey	28/02/2017	v1
Protocol / Proposal	DHS-7_Fieldworker_QRE_EN_20Oct2015_DHSQ7	28/02/2017	v1
Covering Letter	05172017_ENAP_INDEPTH_IRB_response	18/05/2017	v1
Information Sheet	17052017_ Consent forms	18/05/2017	v1
Protocol / Proposal	17052017_INDEPTH_ENAP_ResearchProtocol	18/05/2017	v1
Protocol / Proposal	16052017_CN_Objective 4_Concept_Note	18/05/2017	v1

After ethical review

The Chief Investigator (CI) or delegate is responsible for informing the ethics committee of any subsequent changes to the application. These must be submitted to the Committee for review using an Amendment form. Amendments must not be initiated before receipt of written favourable opinion from the committee.

The CI or delegate is also required to notify the ethics committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reactions (SUSARs) which occur during the project by submitting a Serious Adverse Event form.

At the end of the study, the CI or delegate must notify the committee using an End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at: http://leo.lshtm.ac.uk

Additional information is available at: www.lshtm.ac.uk/ethics

Yours sincerely,

[redacted]

Professor John DH Porter
Chair

ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

London School of Hygiene and Tropical Medicine: Confirmation for PhD study

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
+44 (0)20 7636 8636 • www.lshtm.ac.uk

Joseph Waiswa
RD Student
LSHTM

13 August 2018

Dear Joseph,

Re: Research Degree Project

Thank you for submitting further information following the audit initiated by Professor Della Freeth, former Pro-Director (Learning & Teaching) in 2017. The aim of the audit was to assure the School that all Research Degree (RD) candidates had obtained the appropriate approvals before they start their projects.

The Research Governance Committee considered the initial results of this audit and recommended further action. As a result, a sub-group of the Research Governance Committee reviewed RD projects without valid LSHTM ethics approval in their name, as this has been a requirement since 2014.

As members of this sub-group of the Research Governance Committee, we have reviewed your project and supporting documents. We are satisfied that the aims and analyses are sufficiently detailed in your supervisor(s)' project, LSHTM ref 12218.

Should other data collection or analytical methods not detailed in the existing ethics approval be required for your project, please submit an application to the LSHTM ethics committee prior to any further data collection or analysis of the data/tissue.

Should you have any queries, please contact Patricia Henley in the first instance.

Yours sincerely,

[redacted]

Prof Sasha Shepperd
Pro-Director Learning

[redacted]

Prof Audrey Prost
Head of the Doctoral College

[redacted]

Ms Patricia Henley
Head of Research Governance & Integrity

Cc Angela Baschieri, EPH

REPÚBLICA DA GUINÉ-BISSAU
Ministério da Saúde Púbica

Instituto Nacional de Saúde Púbica

Comité Nacional de Ética na Saúde

Nº Refª 072/CNES/INASA/2017 Bissau, 12 de junho de 2017

À
Drª. Amabélia Rodrigues, Epidemiologista, PhD
Projecto de Saúde Bandim
Instituto Nacional de Saúde Pública, Guiné-Bissau
Apartado 861, Bissau Codex
TM: (245) 96 607 8659/95 527 8134
a.rodrigues@bandim.org

BISSAU

ASSUNTO: Aprovação do Protocolo de Pesquisa

Com os melhores cumprimentos.

O Comité Nacional de Ética na Saúde reunido na sua primeira sessão ordinária no dia 09 de junho de 2017, ao analisar um protocolo do estudo que lhe foi submetido para apreciação e aprovação sob o título "***Plano de ação para Todos os Recém-nascidos ENAP: para a melhoria da medição dos resultados da gravidez. Uma comparação randomizada de módulos de inquérito domiciliar para medir resultados da gravidez***".

Compulsado o Protocolo, o coletivo verificou que todos os requisitos foram cumpridos na integralidade, pelo que se decidiu em consenso por sua aprovação.

Aceite os protestos da nossa elevada consideração.

O Presidente

[redacted]

Dr. Cunhate Na Bangna

ጎንደር ዩኒቨርስቲ
ምርምርና ህብረተሰብ አገልግሎት ምክትል ፕሬዚዳንት
ጎንደር፣ ኢትዮጵያ

University of Gondar
V/P for Research & Community Service
Gondar, Ethiopia

Rfe. VP/RCS/05/1074 /2016

Date 19th April /2017

To Dr. Gashaw Andarge
University of Gondar

Subject: - Ethical Clearance

Your research project proposal entitled **"Improving household survey modules for measuring pregnancy outcomes every newborn action plan partnering with ending preventable maternal mortality."** has been reviewed by the Institutional Ethical Review Board of University of Gondar for its Ethical soundness, and it is found to be ethically acceptable.

Thus, the Research and Community Service Vice President Office has awarded this Ethical Clearance for the above stated study to be carried out by **Dr. Gashaw Andarge** as Principal Investigator and **Dr. Abebaw Gebeyehu, Dr. Solomon Mekonnen, Dr. Kassahun Alemu, Dr. Mezgebu Yitayal,** as Co-investigator as of 13th April, 2017.

These investigators are expected to submit their research progress report to the Vice President for Research and Community Service Office of the University of Gondar.

Best Regards

[redacted]

ዶ/ር ዘውዱ ተሾመ አውላቸው
D/r Zewdu Teshome Awlacew
Delegate, Vice President for
Research & Community Service

ጎንደር ዩኒቨርሲቲ
ምርምርና ህብረተሰብ አገልግሎት ምክትል ፕሬዚዳንት ጽ/ቤት
V/President for Research & Community Service Office
University of Gondar

C.C.
Office of Research and Publication Directorate
Office of institutional Research Review Board
University of Gondar

የፖ.ሣቁ
P.O. Box 196
ጎንደር ኢትዮጵያ
Gondar, Ethiopia

ቴሌግራም ጠ.ኮ
Cable A.A.U. PH.
Fax - 251-058-114 1240
251-058 -114- 12-33
251- 058-114 -1236

Telephone PBX
President office
V/P/for A/ & Research
Research & Com/Service C/Process

ስልክ
058-111 -01 -74
058 -114 -12-31
058-114-12-36
058-231- 11-30
058-811-90-69

URL Addres:- WWW.ugondar.edu.et

Uganda National Council for Science and Technology

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

Our Ref: SS 4244 **11th October 2017**

Dr. Simon Kasasa
Makerere University School of Public Health
Kampala

Re: Research Approval: **A Randomized Comparison of Household Survey Modules for measuring Pregnancy Outcomes in Iganga-Mayuge Health and Demographic Surveillance Site (IMHDSS)**

I am pleased to inform you that on **28/08/2017**, the Uganda National Council for Science and Technology (UNCST) approved the above referenced research project. The Approval of the research project is for the period of **28/08/2017** to **28/08/2018**.

Your research registration number with the UNCST is **SS 4244**. Please, cite this number in all your future correspondences with UNCST in respect of the above research project.

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

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

Below is a list of documents approved with this application:

	Document Title	Language	Version	Version Date
1.	Research proposal	English	1.1	July 2017
2.	Informed Consent Forms	English, Lusoga	1.1	July 2017
3.	Interview guide	English, Lusoga	1.1	July 2017
4.	Assent Forms	English, Lusoga	1.1	July 2017

Yours sincerely,

[redacted]
Beth Mutumba
for: Executive Secretary
UGANDA NATIONAL COUNCIL FOR SCIENCE AND TECHNOLOGY

Copied to: Chair, Mildmay Uganda, Research Ethics Committee

MILDMAY UGANDA

Research Ethics Committee (MUREC)

26 June 2017

Dr. Simon Kasasa,
Makerere University School of Public Health,
&Makerere University Center for Health & Population Research
Iganga-Mayuge Health & Demographic Surveillance Site.

Dear Dr. Kasasa:

Re: Initial approval of your Research protocol titled: # REC REF 0305-2017 "A randomized comparison of household survey modules for measuring pregnancy outcomes in Iganga-Mayuge Health and Demographic Surveillance Site (IMHDSS)".

Thank you for submitting an application for approval of the above referenced protocol to MUREC.

I am glad to inform you that approval is hereby given to conduct the study; this approval is given following your exhaustive responses to initial comments raised by MUREC. This approval is for one year, effective 26th June 2017 and will expire on 26th June 2018. Extension beyond this expiry date and changes to the protocol including data collection tools must be brought to the attention of MUREC.

However, before you proceed you are required to submit the protocol to Uganda National Council for Science and Technology (UNCST) for registration.

You are also required to provide progress reports at an annual interval, to notify Mildmay Uganda Research Committee on completion, as well as when publishing results.

Please do not hesitate to contact us if you have any questions.

I wish you success in this endeavor.

Yours Sincerely,

[redacted]

Mary Odiit
Vice Chairperson
Mildmay Uganda Research Ethics Committee (MUREC)

MILDMAY UGANDA RESEARCH ETHICS COMMITTEE (MUREC)
APPROVED
VALID UNTIL
DATE: 26th June 2018
P.O. BOX 24985, KAMPALA, UGANDA

Location/ Correspondence
Mildmay Uganda
Plot 27, Lweza
P.O.Box 24985, Kampala

Communication
Tel: 0392174236
Email: murec@mildmay.or.ug
www.mildmay.or.ug

Guiding Research For the Future

GHANA HEALTH SERVICE ETHICS REVIEW COMMITTEE

In case of reply the number and date of this Letter should be quoted.

Research & Development Division
Ghana Health Service
P. O. Box MB 190
Accra
Tel: +233-302-681109
Fax + 233-302-685424
Email: ghserc@gmail.com

MyRef. GHS/RDD/ERC/Admin/App/702
Your Ref. No.

Dr. Yeetey Enuameh
Kintampo Health Research Centre
Ghana Health Service
Kintampo

The Ghana Health Service Ethics Review Committee has reviewed and given approval for the implementation of your Study Protocol.

GHS-ERC Number	**GHS-ERC: 19/06/17**
Project Title	A Randomized Comparison of Household Survey Modules for Measuring Pregnancy Outcomes in Five INDEPTH HDSS Sites
Approval Date	26th July, 2017
Expiry Date	25th July, 2018
GHS-ERC Decision	**Approved**

This approval requires the following from the Principal Investigator

- Submission of yearly progress report of the study to the Ethics Review Committee (ERC)
- Renewal of ethical approval if the study lasts for more than 12 months,
- Reporting of all serious adverse events related to this study to the ERC within three days verbally and seven days in writing.
- Submission of a final report **after completion** of the study
- Informing ERC if study cannot be implemented or is discontinued and reasons why
- Informing the ERC and your sponsor (where applicable) before any publication of the research findings.

Please note that any modification of the study without ERC approval of the amendment is invalid.

The ERC may observe or cause to be observed procedures and records of the study during and after implementation.

Kindly quote the protocol identification number in all future correspondence in relation to this approved protocol

SIGNED..... [redacted] .....
DR. CYNTHIA BANNERMAN
(GHS-ERC CHAIRPERSON)

Cc: The Director, Research & Development Division, Ghana Health Service, Accra

Kintampo Health Research Centre (KHRC) Institutional Ethics Committee (IEC)
P.O Box 200
Kintampo, B/A
Ghana, West Africa

Tel: +233(3520)92037/+233504270501
E-mail: fred.kanyoke@kintampo-hrc.org
ethics@kintampo-hrc.org

FULL ETHICAL APPROVAL CERTIFICATE

Dr. Enuameh Yeetey
Kintampo Health Research Centre
Box 200
Kintampo, B/A
Ghana, West Africa

Date: 9th August, 2017

Study File Number: KHRCIEC/2017-14

Title of study: A randomized comparison of household survey modules for measuring pregnancy outcomes in five INDEPTH HDSS sites.

Principal Investigator: Yeetey Enuameh, Joy E Lawn

Co-Investigator(s): Charlotte Tawiah, Sam Newton, Alex Manu, Francis Djabeng, Ernest Nettey, Kwaku Poku Asante, Seth Owusu-Agyei, Peter Waiswa, Akuze Joseph Waiswa, Doris Kwesiga, Angela Baschieri, Hannah Blencowe, Vladimir Gordeev

Type of Review: Full Board Review

Approval Date: 9th August, 2017

Expiration Date: 9th August, 2018

1. The Kintampo Health Research Centre Institutional Ethics Committee (IEC) is constituted and operates in conformance with requirements of 45 CFR 46, 21 CFR 50, 21 CFR 56 and section 3 of the International Council on Harmonization Guidelines, as well as all applicable regulatory, legal, and other ethical requirements governing human subject research in Ghana. The OHRP Federal Wide Assurance number for the committee is 00011103; the IRB registration number is 0004854.

2. The above study in title was reviewed by the IEC on 20th June, 2017 and given conditional approval.

3. The Committee after carefully going through your revised study protocol has granted you full ethical approval for implementation of the study.

4. The following documents were reviewed and approved;
   4.1 A randomized comparison of household survey modules for measuring pregnancy outcomes in five INDEPTH HDSS sites. Version 2 dated July 20, 2017
   4.2 Adult Informed Consent form. Version 2, dated July 20, 2017
   4.3 Parental/LAR Informed Consent form. Version 2, dated July 20, 2017

THE CHAIRMAN, KINTAMPO
HEALTH RESEARCH CENTRE
INSTITUTIONAL ETHICS
COMMITTEE

P.O Box 200
Kintampo, B/A
Ghana, West Africa

Tel: +233(3520)92037/+233504270501
E-mail: fred.kanyoke@kintampo-hrc.org
ethics@kintampo-hrc.org

4.4 Child Assent form. Version 2, dated July 20, 2017
4.5 Data collection tools
4.6 Study Budget
4.7 Curriculum Vitae of study Investigators

5. During study implementation, the IEC must be informed within 72 hours by the principal investigator (PI) of learning of any (a) unexpected, serious, study related adverse events; (b) disclosed adverse events, or (c) unanticipated problems with the study which may pose risk to study participants or others, if applicable.

6. All safety monitory reports, including DSMB summaries and reports, must be submitted to the IEC as soon as they become available to PI(s).

7. Changes or modifications to this research activity must be submitted and approved by the IEC before they are implemented.

8. PI(s) would be required to submit application for renewal of this approval certificate (if necessary) plus a progress report.

9. PI(s) is required to notify the IEC of study completion (end of data collection/last follow-up) or early termination of the research project.

10. Submit final report of the study three months after approval certificate expires (study closure).

11. Before conduct of the study, submit original/final copy of your informed consent/assent forms for **authentication stamp** before making photocopies for your consent process.

12. Regulated study records, including IEC approvals and signed consent forms, must be securely maintained by PI(s) and available for audits for three years after the study is closed with the IEC.

Sincerely,

[redacted]

THE CHAIRMAN KINTAMPO HEALTH RESEARCH CENTRE INSTITUTIONAL ETHICS COMMITTEE

Nana Franklin Fei
Vice Chair
Institutional Ethics Committee
Kintampo Health Research Centre

Study File number: 2017-14

Kintampo Health Research Centre

Ghana Health service

P.O.Box 200 , Kintampo
BrongAhafo Region

Tel: +233-3520 92037 Fax:+233-38872

In case of reply the Number and date of this Letter should be quoted.

My Ref.: SRC/130617

Your Ref. No.

14 June 2017

The Principal Investigator,
ENAP Study
Kintampo Health Research Centre
P.O. Box 200
Kintampo

Dear Principal Investigator,

Decision on your protocol

Your protocol entitled **"A randomized comparison of household survey modules for measuring pregnancy outcomes in five INDEPTH HDSS sites"** has been reviewed.

It is my pleasure to inform you that the protocol has been given full approval by the Scientific Review Committee (SRC). Please submit a soft copy of the final protocol to the SRC for records keeping.

You are also required to submit a copy of the final protocol to the Kintampo Health Research Centre Institutional Ethics Committee for ethical approval before commencement of the study.

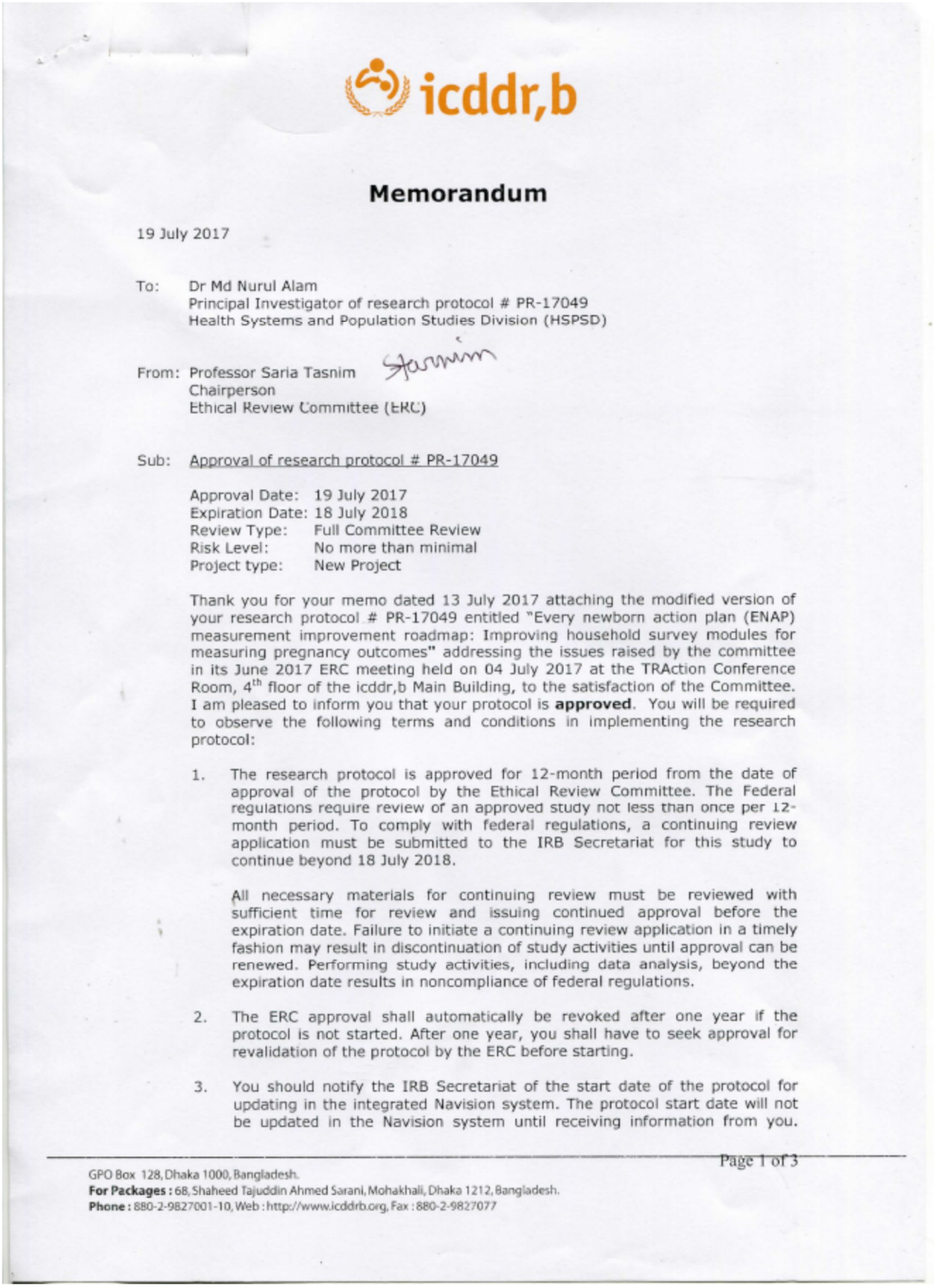
Accept my congratulations.

Sincerely,

[redacted]

Kenneth Ae-Ngibise
(For Chairman)

Matlab HDSS:



icddr,b

Memorandum

19 July 2017

To: Dr Md Nurul Alam
Principal Investigator of research protocol # PR-17049
Health Systems and Population Studies Division (HSPSD)

From: Professor Saria Tasnim
Chairperson
Ethical Review Committee (ERC)

Sub: Approval of research protocol # PR-17049

Approval Date: 19 July 2017
Expiration Date: 18 July 2018
Review Type: Full Committee Review
Risk Level: No more than minimal
Project type: New Project

Thank you for your memo dated 13 July 2017 attaching the modified version of your research protocol # PR-17049 entitled "Every newborn action plan (ENAP) measurement improvement roadmap: Improving household survey modules for measuring pregnancy outcomes" addressing the issues raised by the committee in its June 2017 ERC meeting held on 04 July 2017 at the TRAction Conference Room, 4th floor of the icddr,b Main Building, to the satisfaction of the Committee. I am pleased to inform you that your protocol is **approved**. You will be required to observe the following terms and conditions in implementing the research protocol:

1. The research protocol is approved for 12-month period from the date of approval of the protocol by the Ethical Review Committee. The Federal regulations require review of an approved study not less than once per 12-month period. To comply with federal regulations, a continuing review application must be submitted to the IRB Secretariat for this study to continue beyond 18 July 2018.

   All necessary materials for continuing review must be reviewed with sufficient time for review and issuing continued approval before the expiration date. Failure to initiate a continuing review application in a timely fashion may result in discontinuation of study activities until approval can be renewed. Performing study activities, including data analysis, beyond the expiration date results in noncompliance of federal regulations.

2. The ERC approval shall automatically be revoked after one year if the protocol is not started. After one year, you shall have to seek approval for revalidation of the protocol by the ERC before starting.

3. You should notify the IRB Secretariat of the start date of the protocol for updating in the integrated Navision system. The protocol start date will not be updated in the Navision system until receiving information from you.

GPO Box 128, Dhaka 1000, Bangladesh.
For Packages : 68, Shaheed Tajuddin Ahmed Sarani, Mohakhali, Dhaka 1212, Bangladesh.
Phone : 880-2-9827001-10, Web : http://www.icddrb.org, Fax : 880-2-9827077

A4: Data sharing agreements

EVERY NEWBORN ACTION PLAN
partnering with ENDING PREVENTABLE MATERNAL MORTALITY

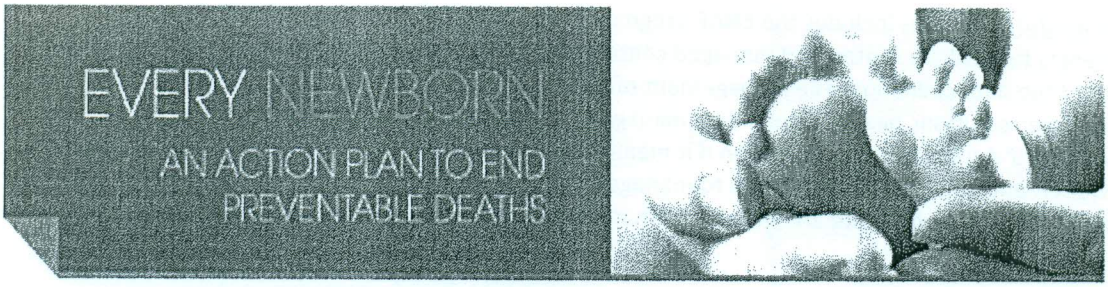
ENAP METRICS & INDEPTH NETWORK RESEARCH CENTERS' DATA SHARING AGREEMENT
Improving household survey and HDSS modules for measuring pregnancy outcomes

July 2017

INDEPTH Network
Better Health Information for Better Health Policy

MARCH
MATERNAL REPRODUCTIVE & CHILD HEALTH

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE



Information and data sharing agreement between the Bandim HDSS in Guinea-Bissau, Dabat HDSS in Ethiopia, Iganga-Mayuge HDSS in Uganda, Kintampo Health Research Centre in Ghana, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b) on behalf of Matlab HDSS, the INDEPTH Network Maternal Newborn and Child Health Working Group (MNCH-WG) technical secretariat which is hosted by the Makerere University School of Public Health (MakSPH), the London School of Hygiene and Tropical Medicine (LSHTM) and the INDEPTH Network, collectively hereafter referred as "the Parties". This is in relation to a collaborative research referred to as "Every Newborn Action Plan (ENAP) measurement improvement roadmap" during the period 2016 - 2018, hereafter referred as "the Project".

Background

It is critical to the success of the Project that the information to which this agreement relates is handled within the appropriate legislative and regulatory environment, with reference to any confidentiality and privacy regulation and legislation pertaining to the countries/HDSS sites where this data and information is collected and processed.

This agreement sets out the roles of each party to the agreement in relation to the information and data shared and their responsibilities therein. This agreement concerns the ENAP pregnancy history and birth history, household survey datasets, the qualitative data and key indicators from five-year period for the HDSS longitudinal dataset containing anonymized information and derivatives produced during the research for the Project, hereafter referred to as the "Project Data" (**See section A - Project Data for details**).

Parties to the Agreement

- LSHTM
- INDEPTH Network MNCH-WG technical secretariat hosted at MakSPH
- Institutions managing the following Health and Demographic Surveillance Sites (HDSS):
    - Bandim HDSS, Dabat HDSS, Iganga-Mayuge HDSS, Kintampo HDSS and icddr,b's Matlab HDSS
- INDEPTH Network operational secretariat

A. Aim and Purpose of the Agreement

This document is an agreement between all the Parties concerning the Project Data. All Parties mentioned in this document are members of the Project and are recipients of the LSHTM ENAP project EPIDZH6210 grant from the Children's Investment Fund Foundation (CIFF).

Project Data

The Project will collect three forms of datasets;

Survey dataset – This includes the ENAP pregnancy history and birth history household survey datasets that will be shared and managed centrally in a secure server located at the LSHTM at the end of the data collection. The management of the survey data collection is by each HDSS site using the survey solutions headquarters' data management platform[1]. The HDSS sites are free to use their own server during the data collection if it means the minimum security details, software and hardare requirements and has the personnel to manage the server, otherwise the LSHTM server is freely available and the HDSS site are still able to manage the data collection and quality control.

[1] http://support.mysurvey.solutions/customer/en/portal/topics/953694-headquarters/articles

Qualitative dataset – This includes data from the Focus Group discussions, Indepth interviews and transcripts created from these. These data will be anonymized.

HDSS dataset – This includes the key indictors from HDSS longitudinal datasets that will be kept at the individual HDSS sites and will be shared in the form of an agreed template for purposes of pooled analysis. In addition, the individual sites will lead on the analysis of individual level data with support from the technical secretariat or LSHTM where necessary or required.

For avoidance of doubt, it is acknowledged by all the parties that this Agreement shall only govern analysis of data collected in pursuit of the ENAP project only and shall not, except to the extent directly required by the project, extend to any background data of any of the parties.

A.1. **Why are Project Data being shared?**

The Parties are working to improve the HDSS and survey capture of pregnancies and outcomes through high quality, policy driven and multidisciplinary research. One benefit of the Project is the ability to draw findings from multiple-HDSS sites and the pregnancy history and birth history household surveys to inform other surveys like the Demographic and Health Surveys (DHS) and Multiple Indicators Surveys (MICS), the longitudinal surveillance systems like in the INDEPTH Network as well as global policies, practices and guidelines.

In addition, following the discussions during the ENAP & INDEPTH Research Protocol Design Workshop held from 15th – 17th June 2016 in Kampala, Uganda, the Parties agreed to make the Project Data as widely and freely available amongst the Parties involved as possible while safeguarding the privacy of the participants and protecting confidential and priority data in the HDSS sites. The Parties have agreed to promote global access (*according to guidelines of the institutions managing the HDSS, LSHTM, MakSPH and INDEPTH Network*) to the Project Data and information, through sharing the Project Data and information to ensure availability across the Parties involved in the study and to answer the questions asked in the Project's protocols, guided by this agreement.

The HDSS sites have confirmed their willingness to share the relevant data within the Parties as long as certain safeguards (**See section H**) are in place.

A.2. Timing of data sharing

The data sharing process will be done in the following steps.

The HDSS sites will supply their metadata/data dictionary and sample datasets based on an agreed template to the INDEPTH Network MNCH-WG technical secretariat at MakSPH and LSHTM teams before the implementation of the longitudinal HDSS surveillance. After sharing these, the Project technical secretariat with support from LSHTM will develop draft syntax for standardizing and linking HDSS and survey data and analysis with continued consultations with the HDSS statisticians.

Progress with specific activities within longitudinal HDSS data management and analysis processes is contingent on the future availability of data from the HDSS sites. To optimize the dissemination of results, all HDSS sites have agreed to make key indicators from their data available in the agreed standardized format within six months of completion of longitudinal surveillance data collection. However, given the remaining lifespan of the Project, it would be useful if the HDSS sites share the key indicators based on an agreed template as soon as possible after collection to maximize the projects impact.

Survey data will be collected electronically for all the sites and will be managed and stored on the survey solutions Headquarters platform during the data collection. Each site will manage and control its own "Headquarters" platform. During data collection, each site will give a monthly report on the status of the survey capture with technical secretariat. This data will then be shared immediately after quality checks and data collection completion by the HDSS sites data management teams.

A.3. Mode of data sharing

The HDSS site will nominate members who will join the data management and analysis working group that will take on the lead role in managing the data during and after data collection and developing the analysis plan for the Project Data. Once the teams are ready with their Project Data, the Project data manager/technical coordinator will deposit the Project Data into a secure central storage repository at the LSHTM.

1. HDSS surveillance dataset will be stored at the HDSS sites while Survey data from all the five HDSS sites after the data collection will be stored in a centralized secure environment at the LSHTM, with adequate access restrictions. The HDSS sites may at any time request for and access information on the storage and dissemination facilities in place.
2. The Project technical secretariat will provide support/templates/syntax for analysis for the HDSS longitudinal surveillance and survey data.

By making the survey Project Data available in a single location the process of providing and obtaining the required data is greatly simplified for all the Parties.

A.4. How will the data be used?

The usage of the Project Data is for scientific and academic purposes; inform the HDSS and DHS systems and to help governments, policy makers and programmers plan for health services better. All publications arising from the Project Data analyses will follow standard criteria of authorship (A separate authorship agreement will be in place as well) including joint publications for pooled datasets.

Findings and experiences from the Project will be disseminated through reports, meetings, workshops, conferences and in peer-reviewed journals. At the time of publication, fully anonymized data will be provided in accordance with the open access requirements of peer-reviewed journals.

B. Data and data quality sections

The Project Data provided to the repository will be fully amalgamated, and anonymized datasets including derived variables. It is important that no data provided will compromise the protection of participants' confidentiality. Qualitative transcripts should have participant and identifying place names removed, and these should be replaced with careful descriptions of the role or characteristics of the removed names to enable interpretation of the data.

The data provided must be the most recent, complete versions that have undergone all quality control and quality assurances steps as described in the Project's protocols and HDSS Standard Operating Procedures (SOPs). All quantitative data should be accompanied by a data dictionary and a data quality report showing if any inconsistencies or erroneous entries exist in the dataset. All qualitative transcripts should be accompanied by descriptive variables of respondent characteristics, the name of the researcher(s) and any field notes taken to describe the situational, institutional and cultural contexts of the interaction with the researcher.

C. Data Transfer and Storage

Data transfer

An SOP will govern the Project Data transfer process and individual arrangements agreed between the Project's data manager and the institutions managing the HDSS. These institutions will be responsible for providing the Project Data to the Project through any of four proposed routes:

1. Uploading via the web using the LSHTM Secure File Exchange service (Recommended)
2. Sending as an attachment by email using an encrypted ZIP file
3. Sending on a portable storage medium in an encrypted form

4. Providing it in person to the Project's data manager or technical coordinator

Encryption is essential. It is encouraged that encryption is implemented through the use of 7Zip software, an open source file archiving application (http://www.7-zip.org/) that provides a Strong AES-256 encryption capability. A user guide is provided for using 7Zip software.

The LSHTM Secure File Exchange service can be found at https://sfx.lshtm.ac.uk/ and a user guide is provided.

Project Data will be accepted in any suitable format. If there are data that use an uncommon file format (for example: Oracle Big/Little Endian datapump files, Lotus Notes database files, or portable database format) please discuss this with the Project's data manager before sharing them. Data will be converted by the Project's data manager into standard formats for storage within the repository. The formats to be made available in the repository will include, but may not be limited to CSV and PDF.

Data storage

The HDSS longitudinal datasets will be stored at each HDSS site

The Survey Data received from the HDSS site will be stored in a central server at the LSHTM following the discussions and recommendations at the ENAP & INDEPTH Research Protocol Design and study implementation workshops in Kampala-Uganda and Dhaka-Bangladesh respectively. The central server will be mirrored for each HDSS site's individual datasets using the Headquarters with access via a secure logon.

The server will be maintained by LSHTM IT Support, and backed up according to current best practice at the LSHTM. The repository will be managed and curated by the Project's data manager also known as the "Administrator". Direct access to the repository will only be available to authorized Project members via the "Headquarters". Services provided by LSHTM will include regular backup, maintenance of all security related patches for the server and regular review of hardware for faults.

The Project Data supplied to members of a working group for specific cross-HDSS activities will be stored on their computers in the short-term. Arrangements around such storage will be included in an end-user data transfer agreement to ensure appropriate safeguards are in place to avoid loss of, or unauthorised access to the data.

The Project Data will be stored for a minimum of 10 years following the Project completion in accordance with the LSHTM's Records Retention and Disposal Schedule and in accordance with the INDEPTH Network procedures.

D. Ownership, Management and Access

Data ownership

The HDSS sites will retain ownership of their longitudinal surveillance data as well as the pregnancy history and birth history household survey data. The shared pooled survey dataset will be co-owned by all parties.

Data management

The central server repository will be managed by the technical coordinator and the LSHTMs data manager, who will ensure that the repository is held in a secured environment, using resources that are fully supported (with regular maintenance and routine backups being made). The HDSS datasets will be managed by the HDSS statisticians and data managers.

Data access

All individuals who will have access to the Project datasets linked to the ENAP project will be required to sign a non-disclosure and confidentiality agreement (See Annex I).

Access to the repository will be via a web interface with a password-protected login. All users will be registered and verified before being able to access the repository in accordance with the provision set out in Section H of the agreement. These access arrangements will be managed by the Project's data administrator and technical coordinator and can be reviewed periodically. For example, initial access may only be granted to members of the working group, with access being extended to other members as cross-HDSS analysis activities develop. All HDSS sites will be able to access their own data immediately using the Headquarter platform. All HDSS sites will be able to review online and access their data.

Access arrangements for use of data in the repository from other HDSS sites will be outlined in the SOP for access and storage. Access to data will be guided by that stipulated in Section H of the relevant agreement. Arrangements around storing data on the end-user's computer will be included in a data access agreement to be signed by the end-user to ensure appropriate safeguards are in place to avoid loss of, or unauthorized access to the data. The SOP and data access agreement will be available once the web-based access component has been developed.

E. Potential future use

The Project's longer term ambition is to widen access to the repository to the external scientific community but this will require further discussion and agreement. A separate data transfer agreement will be developed for sharing metadata and raw data beyond the Project. Following the discussions with the HDSS member sites during the implementation workshop in Dhaka-Bangladesh, the external users will be required to write a short concept clearly elaborating what they intend to use the Project dataset. Subsequent analysis of data for publication or other purposes by external users should be in agreement and collaboration with the institutions managing the HDSS.

Required Information Dataset requests must include contact information, a research project title, and a description of the analysis proposed to be performed with the data. (This procedure is similar to what the DHS uses before sharing any DHS datasets).

F. Retention Period

The Repository will be preserved and made available until the Project, or parties acting legally in its place, deem that these data are to be destroyed. All reasonable efforts will be made to contact Providers with data in the repository before any such action.

G. Exit Strategy

If LSHTM are no longer in a position to support the repository using LSHTM facilities, then LSHTM will give the HDSS sites and the Project a period of six months to find a suitable location for this repository (possibly the INDEPTH Network data repository).

H. The Agreement

This agreement stipulates that the HDSS sites will:

1. Warrant that the necessary ethical and institutional approvals are in place for the sharing of data within the Project.
2. Warrant that agreement for the sharing of the Project Data has been granted by the other project Principal Investigator(s) (PIs).

3. Provide initial HDSS data structures based on agreed templates for creation of a standardized dataset from the routine surveillance with selected variables of interest to the Project.
4. Maintain ownership and responsibility for the upkeep of the Project Data provided.
5. Supply and maintain on the repository an acknowledgement of the original data providers and a citation for use in publications.

This agreement stipulates that the all Parties agree to:

1. Ensure these Project Data are preserved until all parties acting legally in its place, deem that these Project Data are to be destroyed or no longer required to be stored in the LSHTM repository.
2. Ensure these Data will be stored in a secured manner with appropriate safety and routine maintenance.
3. Ensure that a suitable mechanism is provided to enable members of the Project to engage actively with these Data for the purposes outlined in section A.4.
4. Ensure that access to these Data is controlled through secure, authenticated mechanisms to prevent any unauthorized access.
5. Manage the accreditation status of individuals accessing these Data.
6. Provide an audit trail for any changes introduced to these Data over time.
7. Ensure that subsequent analysis of data for publication or other purposes by external users are performed in collaboration with the institutions managing the HDSS. Where necessary, make available to the end-user a standard form of words, provided by the HDSS sites, for the purpose of acknowledgement of the HDSS sites.
8. Ensure study teams' participation in design, analysis and manuscript writing arising from all projects that use the data accrued.

I. Communications:

The Project's operational secretariat will appoint a member who will act as focal person for this agreement. Should the focal person be replaced, the Project team will communicate the name of the new contact person to the HDSS sites. Communication for administrative and procedural purposes may be made by email, fax or letter as follows:

Communications made by HDSS sites to Project team will be directed to:
Name of contact person:
Title of contact person:
Address of the contact person:
Email:
Tel:

Communications made by Project team to HDSS site will be directed to:

BANDIM
Name of contact person: Amabelia Rodrigues / Ane B. Fisker
Title of contact person: Director of Research / Senior researcher
Address of the recipient centre: Projecto Saude Bandim, Apertado 861, 1004 Bissau Cedex
Email: a.rodrigues@bandim.org / abf@ssi.dk
Tel: +245 966078659 / +45 32688366

Bandim

DABAT
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

IGANGA-MAYUGE
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

KINTAMPO HEALTH RESEARCH CENTRE
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

icddr,b's MATLAB HDSS
Name of contact person: Nurul Alam
Title of contact person: Scientist and Head, Matlab HDSS
Address of the recipient centre: Health Systems and Population Studies Division
Email:nalam@icddrb.org
Tel:+8801819438287

J. Signatories

The following signatories have read and agree with the Agreement as presented above:

Representatives of the 5 HDSS

BANDIM

Name: Prof. Peter Aaby, Site leader

Signature [redacted] Date 15/9-2017

DABAT

Name: _____

Signature _____ Date _____

IGANGA-MAYUGE

Name: _____

Signature _____ Date _____

KINTAMPO HEALTH RESEARCH CENTRE

Dabat

Tel:

DABAT
Name of contact person: Gashaw Andargie Biks
Title of contact person: Principal Investigator
Address of the recipient centre: Dabat HDSS, University of Gondar, Gondar Ethiopia
Email:gashawab@gmail.com
Tel:+251581116221/+251911385423

IGANGA-MAYUGE
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

KINTAMPO HEALTH RESEARCH CENTRE
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

MATLAB
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

J. Signatories

The following signatories have read and agree with the Agreement as presented above:

Representatives of the 5 HDSS

BANDIM

Name: _____

Signature _____ Date _____

DABAT

Name: Dr. Gashaw Andargie PhD. _____

Signature [redacted] Date July 28/2017

IGANGA-MAYUGE

Name: _____

Director, Institute of Public Health

University of Gondar College of Medicine & Health Sciences Institute of Public Health

IgangaMayuge

DABAT
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

IGANGA-MAYUGE
Name of contact person: **Edward Galiwango**
Title of contact person: **Site Operations Coordinator – Study Coordinator**
Address of the recipient centre: **MUCHAP - Iganga Mayuge HDSS**
Email: **edigali@yahoo.com**
Tel**: +256 772 458233**

KINTAMPO HEALTH RESEARCH CENTRE
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

icddr,b's MATLAB HDSS
Name of contact person: Nurul Alam
Title of contact person: Scientist and Head, Matlab HDSS
Address of the recipient centre: Health Systems and Population Studies Division
Email:nalam@icddrb.org
Tel:+8801819438287

J. Signatories

The following signatories have read and agree with the Agreement as presented above:

Representatives of the 5 HDSS

BANDIM

Name: _____

Signature _____ Date _____

DABAT

Name: _____

Signature _____ Date _____

MAKERERE UNIVERSITY CENTRE FOR HEALTH & POPULATION RESEARCH (MUCHAP) - IMHDSS
Name: DR. DAN KAJUNGU

Signature [redacted] Date 6th September, 2017

IGANGA-MAYUGE
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

KINTAMPO HEALTH RESEARCH CENTRE
Name of contact person: KWAKU POKU ASANTE
Title of contact person: DR
Address of the recipient centre: P.O. BOX 200, KINTAMPO NORTH MUNICIPALITY
Email: kwakupoku.a sante@kintampo-hrc.org
Tel: +233 208956598

MATLAB
Name of contact person:
Title of contact person:
Address of the recipient centre:
Email:
Tel:

J. Signatories

The following signatories have read and agree with the Agreement as presented above:

Representatives of the 5 HDSS/Centre

BANDIM

Name:

Signature _____ Date _____

DABAT

Name: _____

Signature _____ Date _____

IGANGA-MAYUGE

Name: _____

Signature _____ Date _____

KINTAMPO HRC

Name: DR KWAKU POKU ASANTE

Signature [redacted] Date 18TH SEPTEMBER, 2017

[redacted]

Matlab

IGANGA-MAYUGE

Name: _____

Signature _____ Date _____

KINTAMPO HEALTH RESEARCH CENTRE

Name: _____

Signature _____ Date _____

International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b)

Name: Prof John D. Clemens, Executive Director

Signature [redacted] Date [redacted]

Representatives of Makerere University technical secretariat

Name: _____

Signature _____ Date _____

Name: _____

Signature _____ Date _____

Representatives of the London School of Hygiene and Tropical Medicine

Name: _____

Signature _____ Date _____

Name: _____

Signature _____ Date _____

Representative of INDEPTH Network Secretariat

Name: Professor Osman Sankoh

Signature [redacted] Date 26/9/17

INDEPTH Network

Name: _____

Signature _____ Date _____

International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b) -

Name: Prof John D. Clemens, Executive Director _____

Signature _____ Date _____

Representatives of Makerere University technical secretariat

Name: _____

Signature _____ Date _____

Name: _____

Signature _____ Date _____

Representatives of the London School of Hygiene and Tropical Medicine

Name: _____

Signature _____ Date _____

Name: _____

Signature _____ Date _____

Representative of INDEPTH Network Secretariat

Name: Professor Osman Sankoh

Signature [redacted] Date 26/9/17

EN-INDEPTH metrics project team

KINTAMPO HEALTH RESEARCH CENTRE

Name: _____

Signature _____ Date _____

International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b) -

Name: Prof John D. Clemens, Executive Director _____

Signature _____ Date _____

Representatives of Makerere University technical secretariat

Name: Joseph Akuze. W

Signature [redacted] Date 2/10/17

Name: Peter Waiswa

Signature [redacted] Date 2/10/17

Representatives of the London School of Hygiene and Tropical Medicine

Name: JOY LAWN

Signature [redacted] Date 28/09/2017

Name: _____

Signature _____ Date _____

Representative of INDEPTH Network Secretariat

Name: Professor Osman Sankoh _____

Signature _____ Date 26/9/17

A5: HDSS Women's Listing Manual

LISTING MANUAL INTRODUCTION

This manual provides guidance on how to create user accounts in the Survey Solutions HQ under the teams and roles tab (tab number 1), as well as the listing process and the survey setup (creating listing, assignments, archiving interviews and deleting). It complements the guidance provided in two other manuals that were already shared with you.

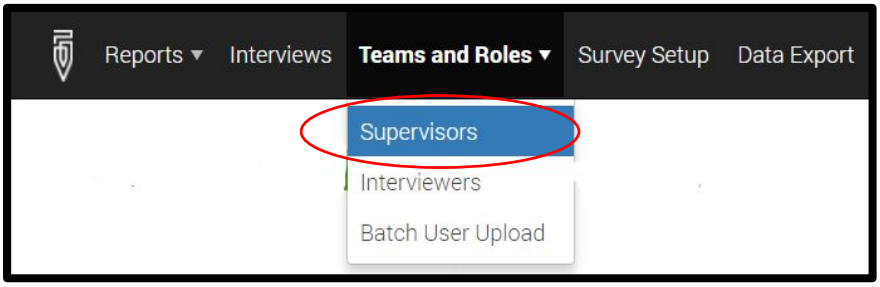
#1: TEAMS AND ROLES TAB

Under the Teams and Roles tab, the survey data manager or statistician or survey manager is able to create the supervisors and interviewers accounts for the individuals who will participate in the survey implementation.

A. HOW TO ADD/ARCHIVE SUPERVISORS

A1.1 Creating accounts for supervisors – ONE BY ONE

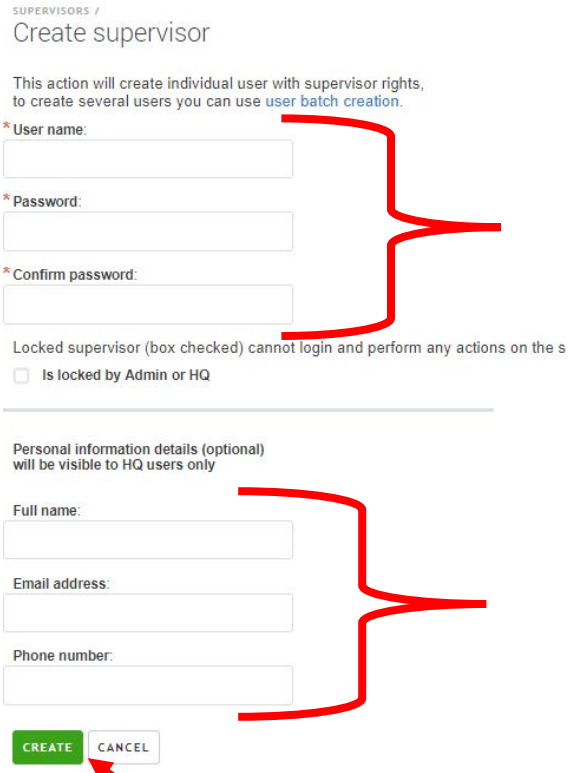
Ensure that you add supervisors for your HDSS. The supervisors responsible for assigning tasks to the interviewers and reviewing the data collectors completed questionnaires. To add a supervisor, click on the *Teams and Roles*



Then click on the green add supervisor tab

Then complete the form displayed below and click on ***CREATE*** to add the supervisor in the system.

Note: This option only allows you to create one supervisor account at a time.



Fill out this information for the user i.e. Username and Preferred password

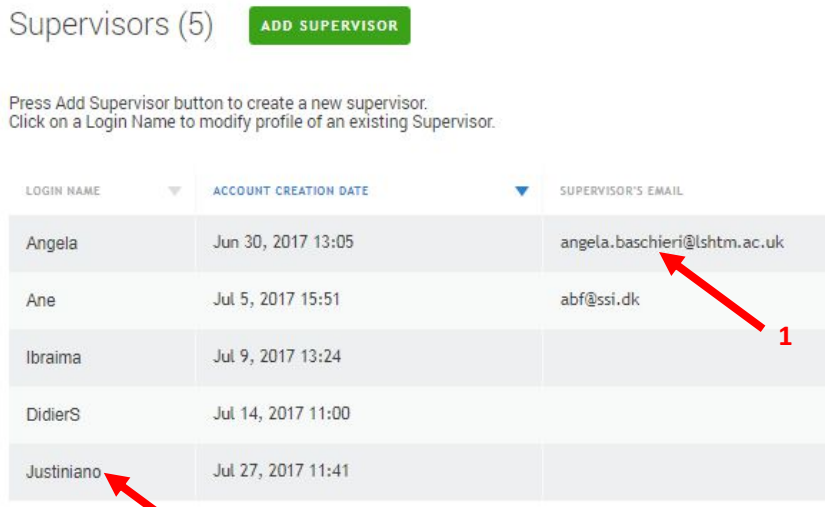
CAUTION!! Make sure the passwords are secure enough including numbers, letters and symbols.

Please check the box "Is locked by Admin or HQ" ONLY when you want to suspend the user from collecting the data. When checked, the user will not be able to collect any data or synchronize the tablet and transfer the collected data (see page 3).

If you fill out this part, the information will be only visible to the HQ.

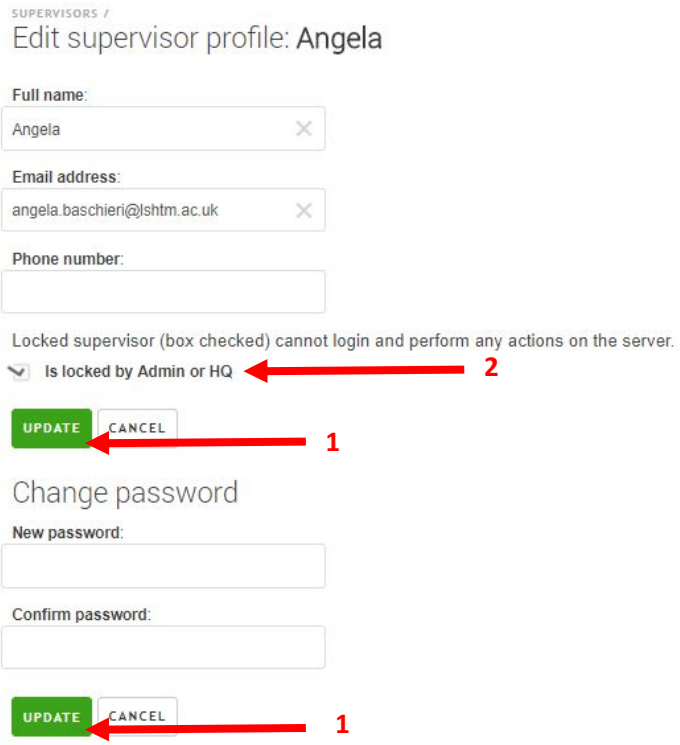
If you will need to send out emails to the users, ensure you fill out their emails.

Click create to add the supervisor else click on cancel



To send emails to the user/supervisor, click on the user's email address. Emails are not sent automatically, therefore, it is upon the HEADQUARTER user to compose the messages.

2. Click on the username to modify information



In order to modify any details of an existing supervisor account (including: Username, Passwords, Email address and Phone number), click on the current supervisor's username in the Name column (***See 2 on a previous page***).

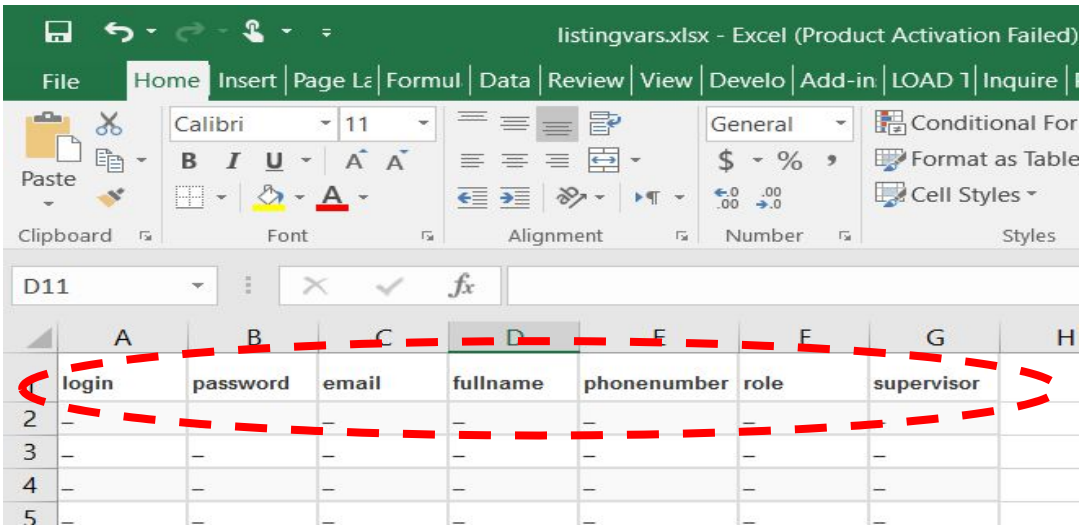
At this point you can do the following:

1. *Change/Update passwords and email, full name and phone number for the supervisor*
2. *Lock all access to the account (for example if the user of the account is discontinued or fired).*

A1.2 Creating accounts for supervisors – IN A BATCH

Before you start thinking about creating a batch upload of supervisors, you will need to do some background work. This includes opening a new MS excel workbook, to which you will add the following fields; (***login, password, email, fullname, phonenumber, role, supervisor***) as it is illustrated below.

NOTE: The fields you are creating are case sensitive and should appear exactly as they are written in this manual. Also the login, password and role <<Supervisor (case-sensitive) or Interviewer (case-sensitive)>> must be filled in for the process to be successful.

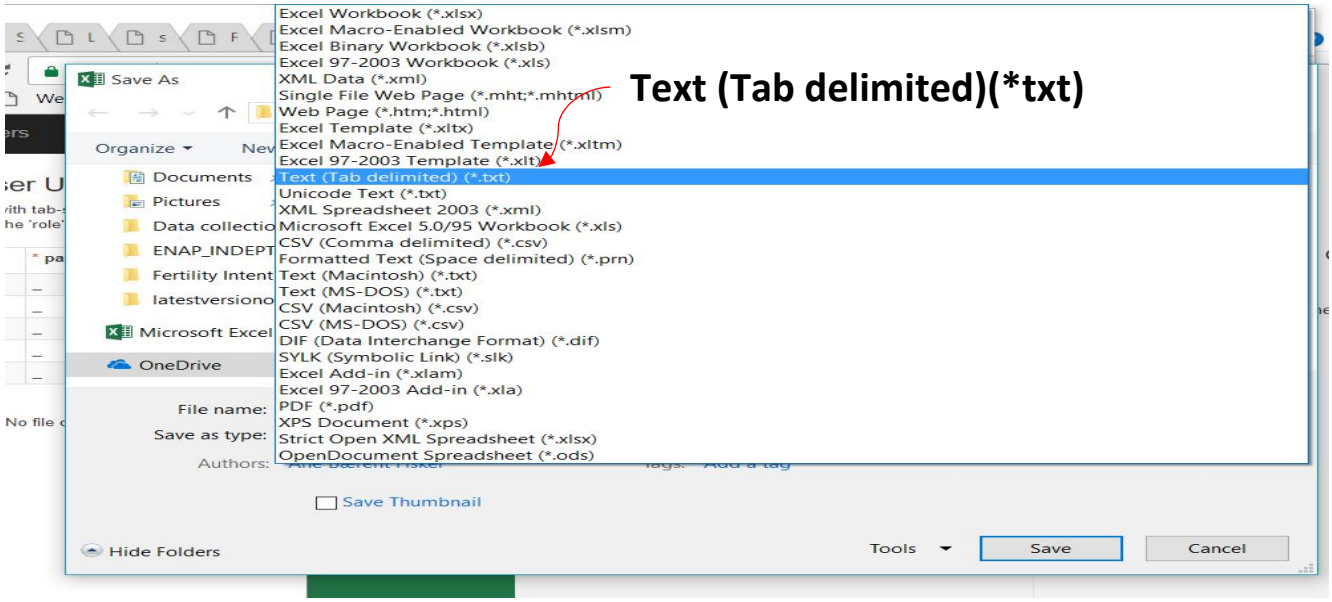


3

Once you have added the fields with the same exact wording, you should now save the workbook on your computer. Save as type "**Text (Tab-delimited) (*.txt)**" file.

Ensure that you save it in a location that you can find easily on your computer. For example in a folder on your Desktop.

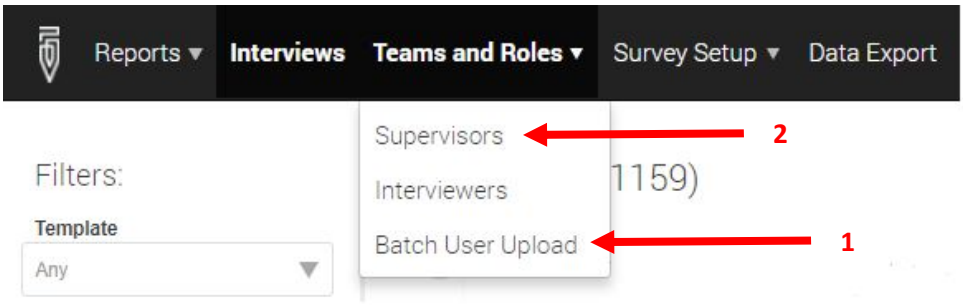
Note: Only 'Supervisor' OR 'Interviewer' are allowed in the 'role' column.



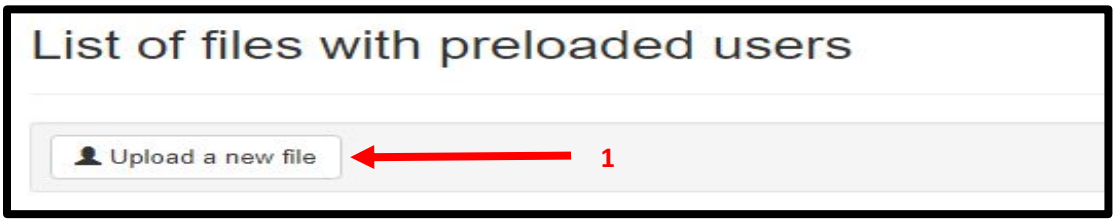
At this point you are now ready to commence with creating accounts of supervisors in a batch.

In order to upload more than one supervisor using the user batch upload option, two options are available to you:

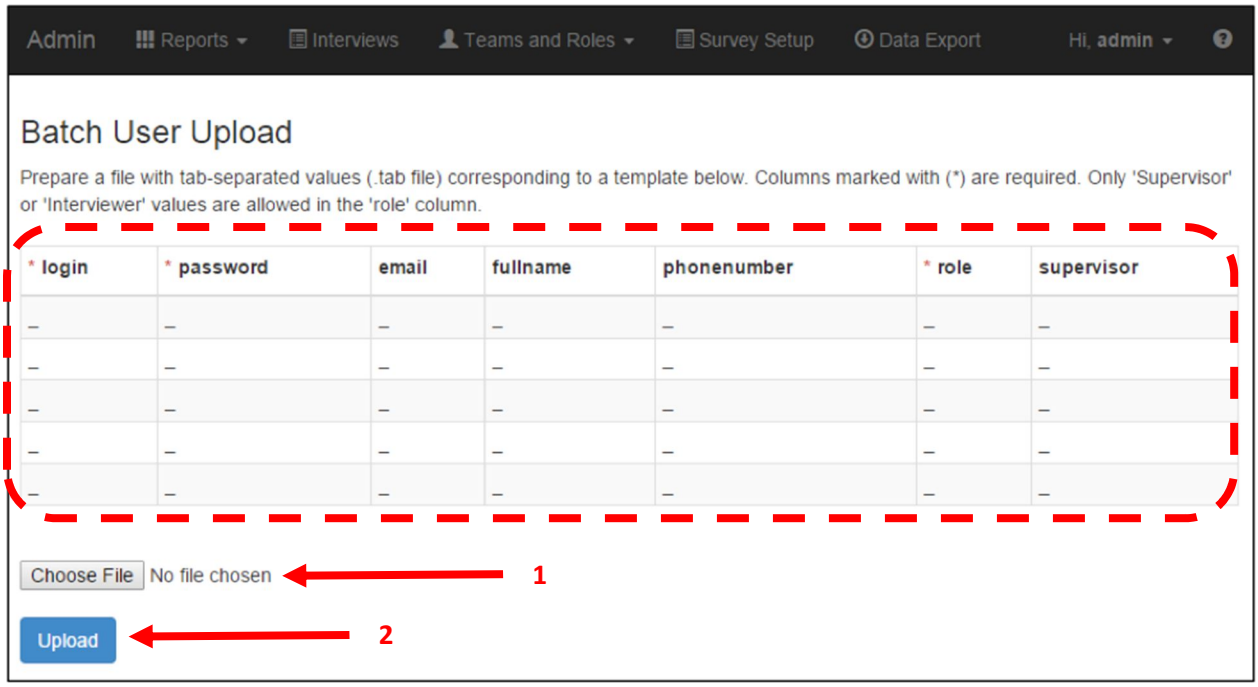
Option 1: Using the BATCH USER UPLOAD (**Option 1**) under the Teams and Roles tab on the HQ main tab



When you click option 1 (Batch User Upload) you will be directed to another page where you will be able to upload the text file that you have just created and saved.

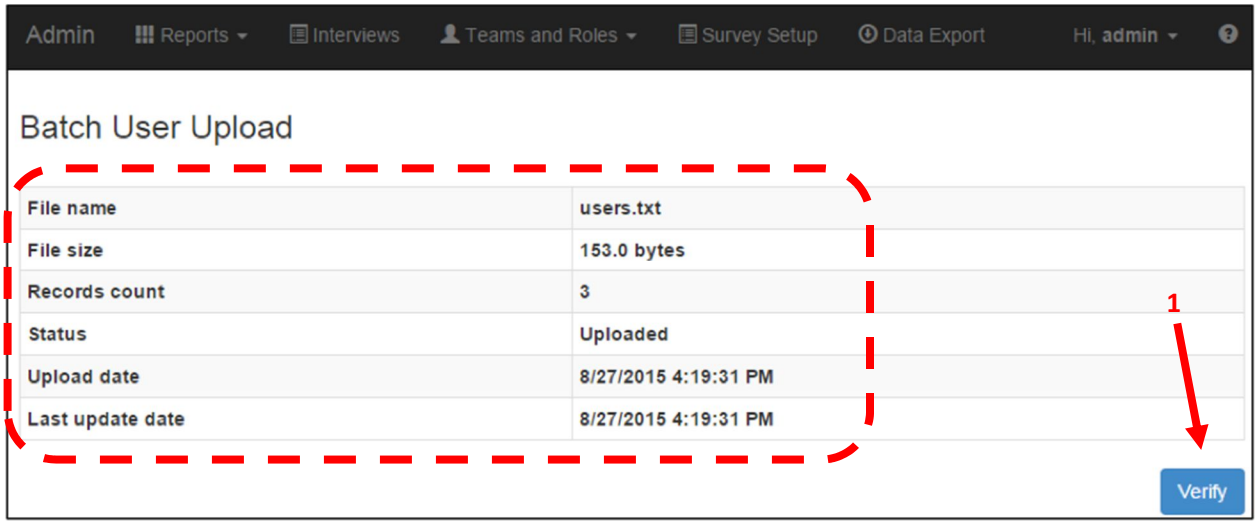


When clicking on the Upload a new file, survey solutions will present you with the layout of the file it expects. At this point you can find your file on your computer and upload it into the system.



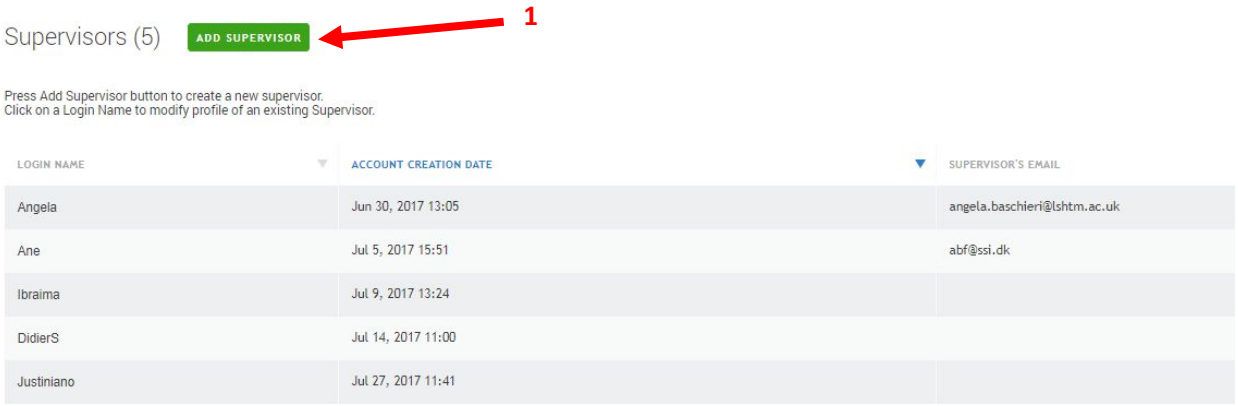
The survey solutions system will need to validate the uploaded list of users (by clicking on verify) report if any errors have been encountered you will be notified:

- Empty required fields (login, password, role)
- Password not conforming to the security requirements (too short, etc.)
- Duplicate user accounts
- Unknown supervisors (all interviewers created must become part of the team of a supervisor known to the system)
- Other checks

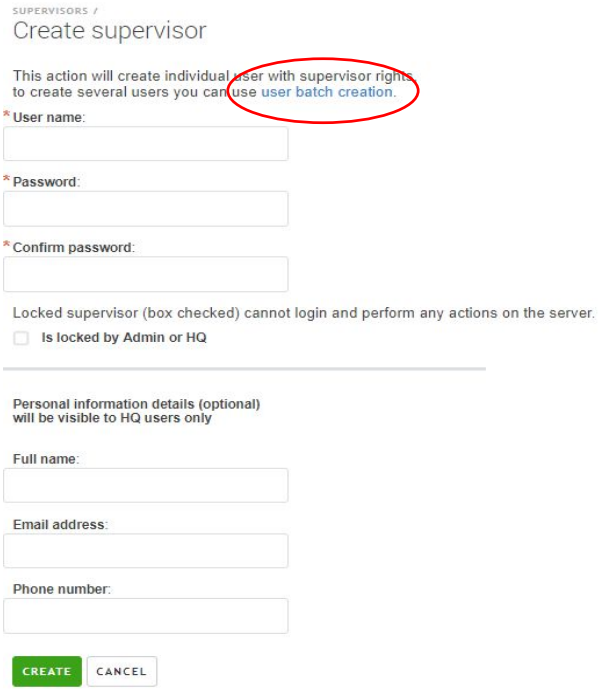


If the Survey Solutions program did not encounter any errors with the list of users, click the create users button to start creating accounts in the batch mode. This process will take a few minutes but is dependent on the number of users and the server performance.

Option 2: Using the same approach as before for adding a single user by clicking on the Supervisor option (**Option 2**). This will open up a new page with existing supervisors. To add the supervisors, click on the green add supervisor tab.



Then click on the user batch creation (Highlighted in Blue) under the Create supervisor dialog box



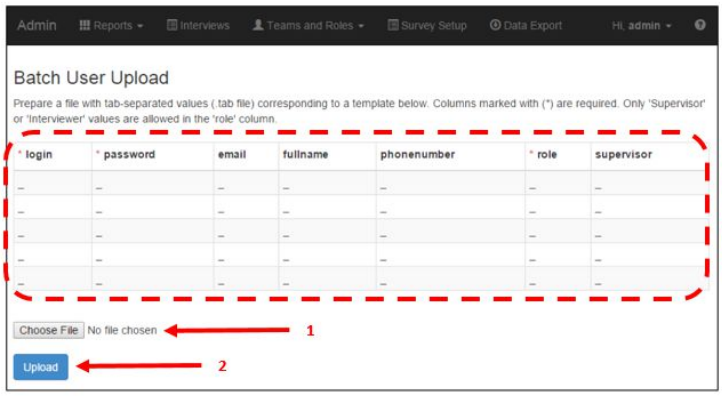
NOTE: When you click on the user batch creation, the same process of uploading, processing, system verification and creation of accounts already seen above under option 1 will be repeated.

For purposes of completeness we have repeated the process in the picture below.

BEGINNING OF REPEATED SECTION

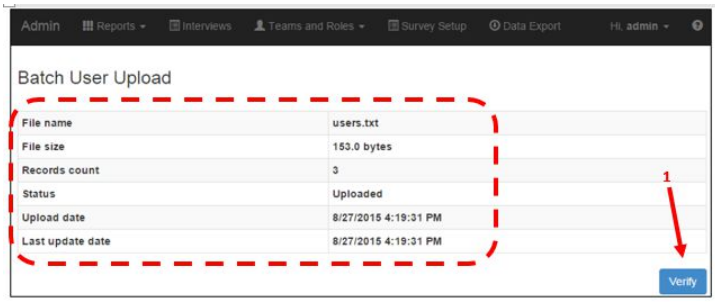
When you click option 1 (Batch User Upload) you will be directed to another page where you will be able to upload the text file that you have just created and saved.

When you click on the Upload a new file, survey solutions will present you with the layout of the file it expects. At this point you can find your file on your computer and upload it into the sytem.

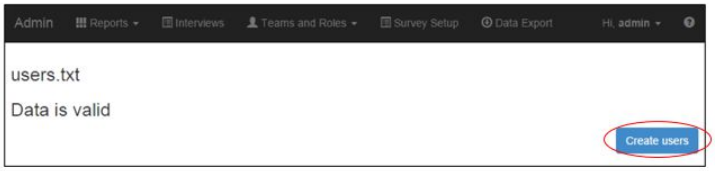


The survey solutions system will need to validate the uploaded list of users (by clicking on verify) report if any errors have been encountered you will be notified:

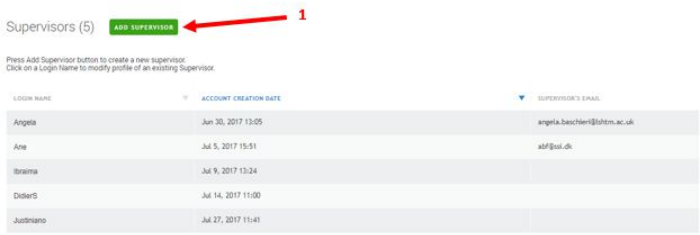
- Empty required fields (login, password, role)
- Password not conforming to the security requirements (too short, etc)
- Duplicate user accounts
- Unknown supervisors (all interviewers created must become part of the team of a supervisor known to the system)
- Other checks



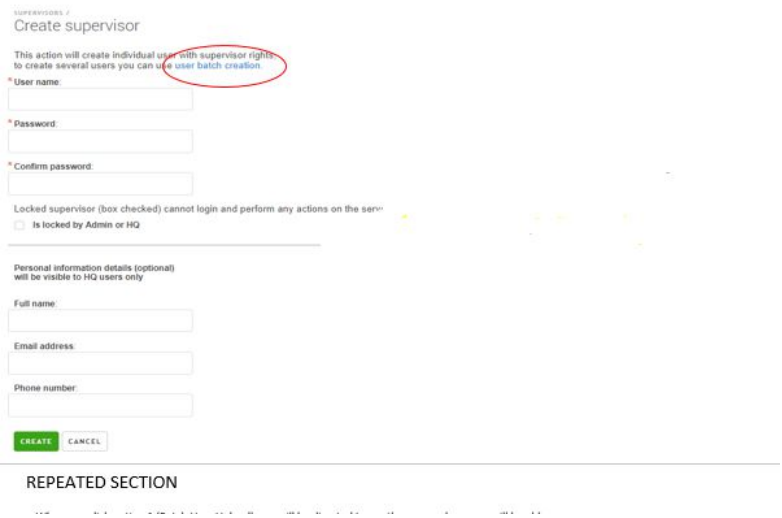
If the survey solutions program did not encounter any errors with the list of users, click the create users button to start creating accounts in the batch mode. This process will take a few minutes but is dependent on the number of users and the server performance.



Option 2: Using the same approach as before for adding a single user by clicking on the Supervisor option (**Option 2**). This will open up a new page with existing supervisors. To add the supervisors, click on the green add supervisor tab



Then click on the user batch creation (Highlighted in Blue) under the Create supervisor dialog box

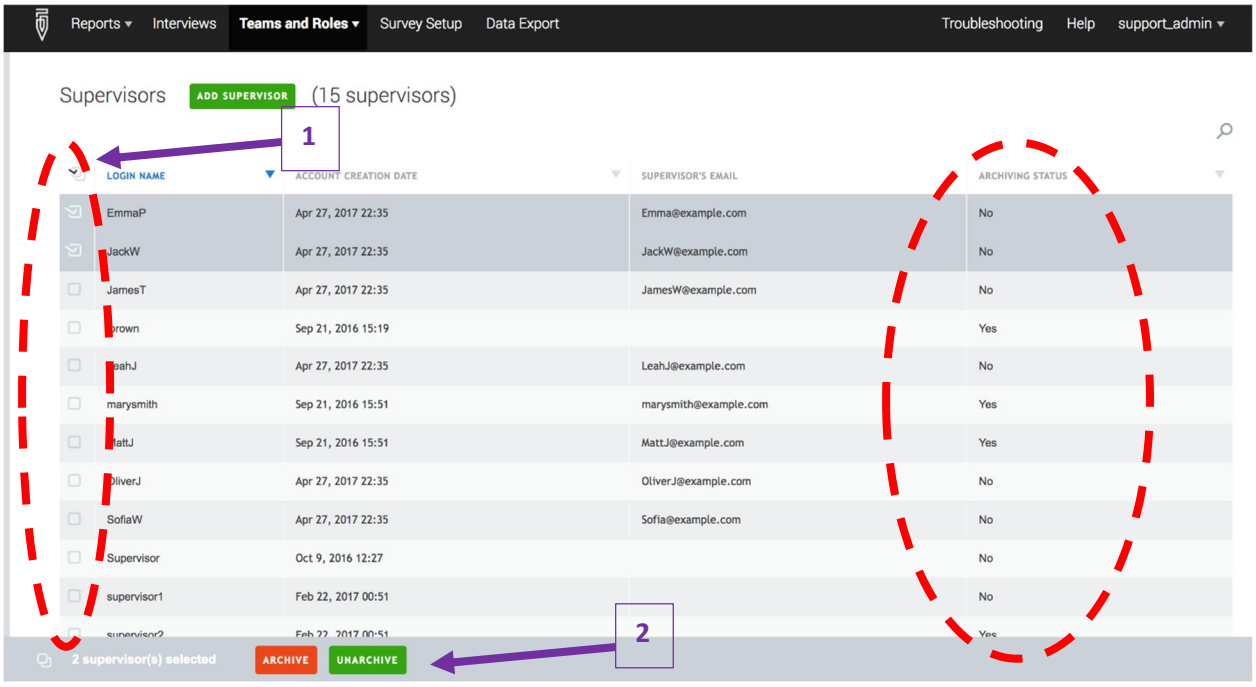


REPEATED SECTION

END OF A REPEATED SECTION

A2. Archiving supervisor accounts – ONE BY ONE AND BATCH

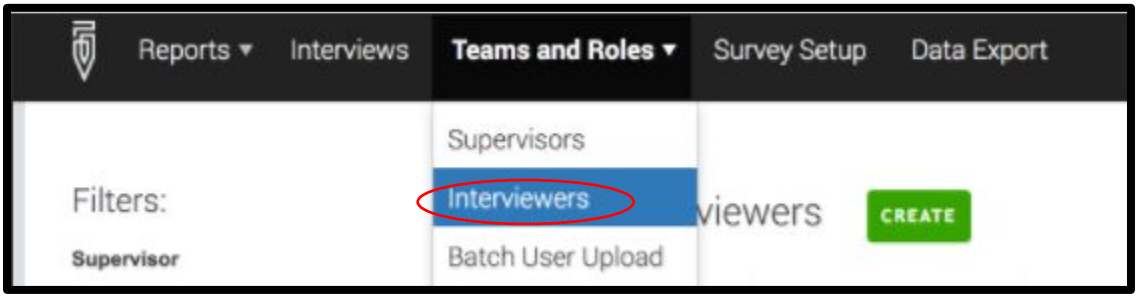
Administrators can archive supervisor accounts. By archiving a supervisor, you are also archiving all the interviewers on that supervisors' team. Supervisors and their data collection team will no longer be able to sync with headquarters. To archive a supervisor, select the check boxes for the supervisor(s) you wish to archive and click on the red archive button. Archived interviewer accounts can be restored. On the supervisors page, find the archived users using the Archive Status filter on the right most column. Then, select the check.

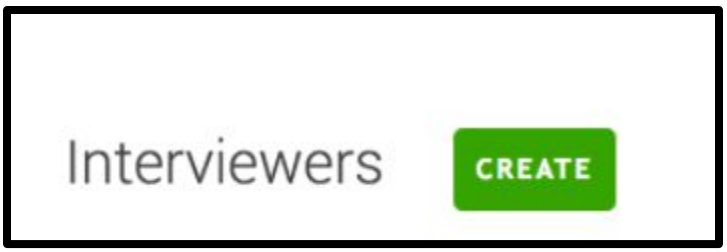


B. HOW TO ADD/ARCHIVE INTERVIEWERS

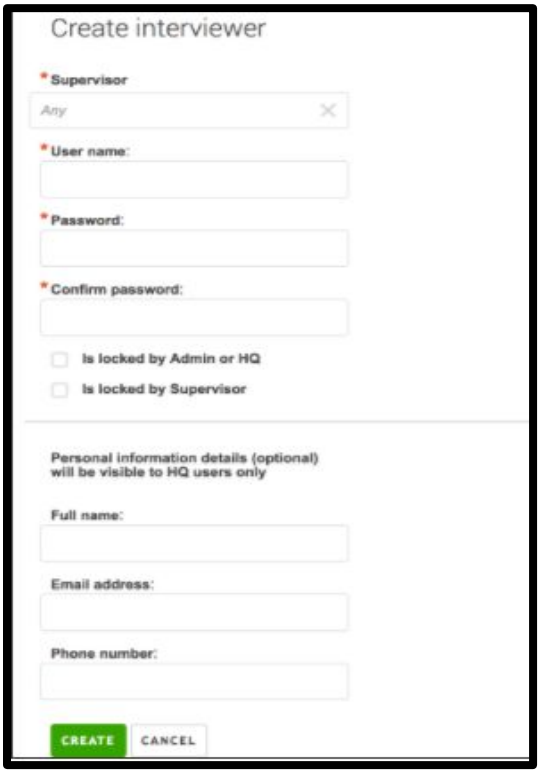
B1.1 Creating accounts for interviewers – ONE BY ONE

This process is similar to adding supervisors. You need to go to the **Teams and Roles** tab and then select Interviewers, then create Interviewers, then complete the form displayed.



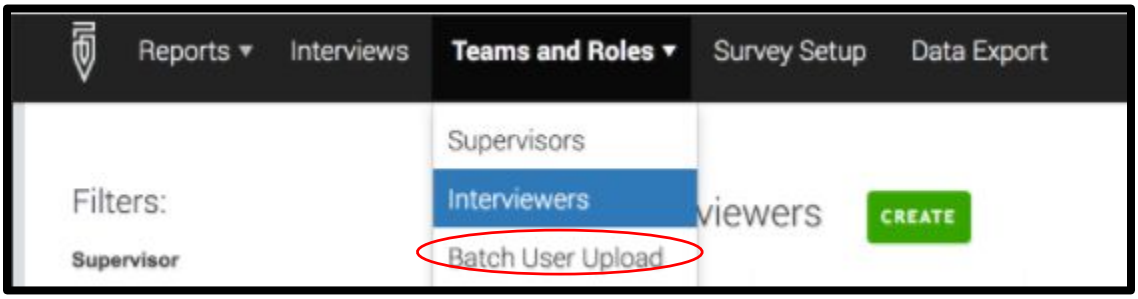


This approach allows you to upload only one Interviewer at a time.



B1.2 Creating accounts for interviewers – ONE BY ONE

To upload many Interviewers, you need to use the "Batch User Upload" option under the Teams and Roles tab. When you click on the Batch User Upload option under the Teams and Roles tab, you will be required to upload a text-delimited files.



The batch upload uses the same fields used when creating the supervisors' batch upload. Therefore, this means that you are able to great both the supervisors and the interviewers accounts in one click. This option saves you a lot of time.

NOTES:

The steps of creating the text-delimited file that have been described under the supervisor batch upload option above are the same steps you need. Up to 10,000 user accounts can be created from a single tab-delimited file, which must include the login, password, and role of the user, and may include additional attributes, like email or phone.

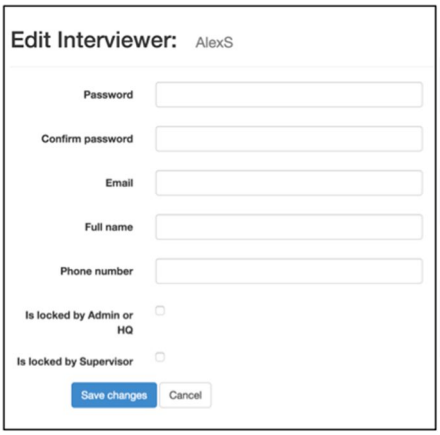
Batch creation of user accounts comes handy when: the same accounts need to be replicated on a different server; migrating from a different CAPI system with established accounts; the survey is so huge that manual creation of accounts via a web interface would take a significant amount of time.

Creating user accounts in batch mode is available to both the administrator and the HQ users. For new supervisor accounts the role must be specified as Supervisor (case-sensitive). For interviewer accounts the role must be specified as Interviewer (case-sensitive), and in addition the login of the supervisor must be specified, which determines the team where the interviewer account will be added. Interviewers may be added to existing supervisors, or to new supervisors mentioned anywhere in the users list. When creating new users with a batch upload feature, select the Batch user upload menu item of the Teams and Roles menu.

Modifying data related to interviewers and monitoring interviewers

To modify an interviewer's account—change the password or lock the account—click on the interviewer's name. To monitor an interviewer, check the activation status in the Connected to device column and the

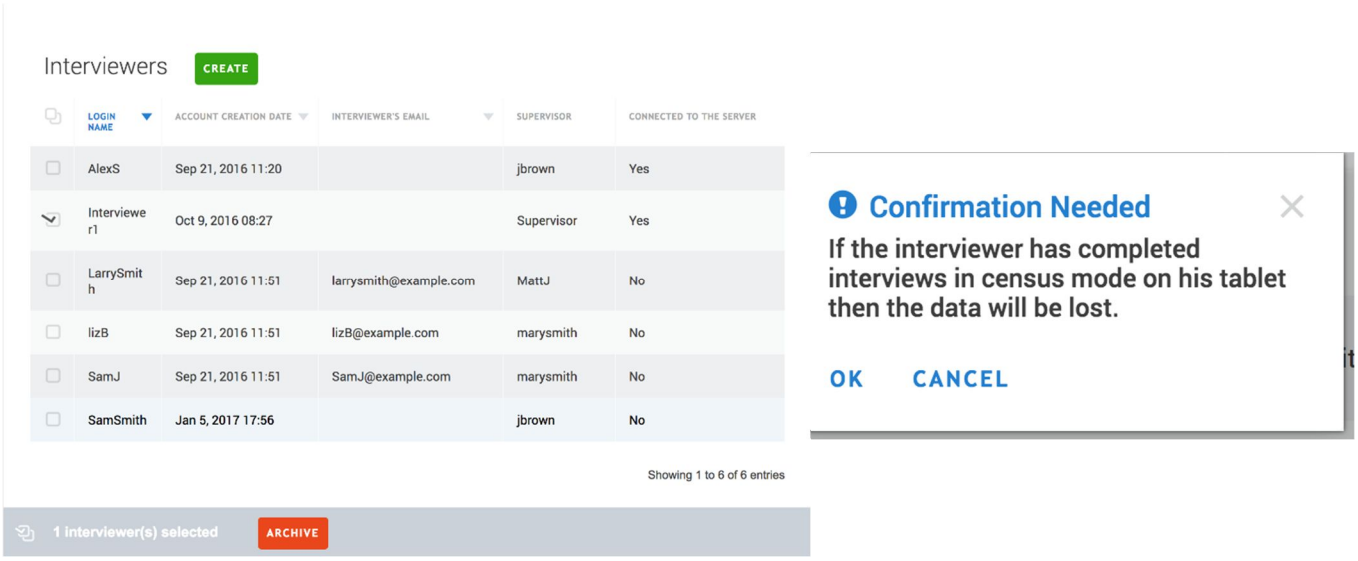
synchronization log. The Connected to device column displays the status of the tablet's activation or in other words it will show if an interviewer has successfully activated a tablet using their login credential



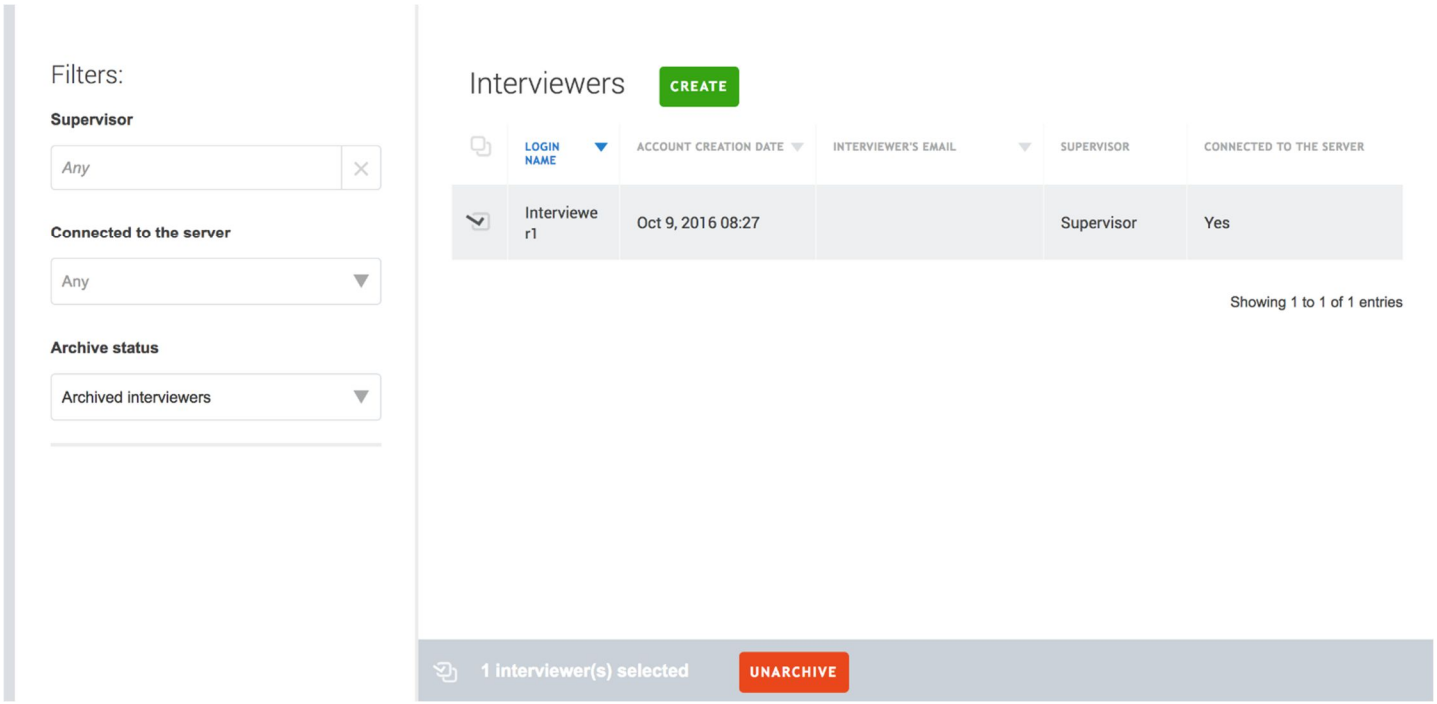
LOGIN NAME	ACCOUNT CREATION DATE	INTERVIEWER'S EMAIL	SUPERVISOR	CONNECTED TO THE SERVER
AlexS	Sep 21, 2016 11:20		jbrown	Yes
lizB	Sep 21, 2016 11:51	lizB@example.com	marysmith	No
SamJ	Sep 21, 2016 11:51	SamJ@example.com	marysmith	No
LarrySmith	Sep 21, 2016 11:51	larrysmith@example.com	MattJ	No
Interviewer 1	Oct 9, 2016 08:27		Supervisor	Yes
SamSmith	Jan 5, 2017 17:56		jbrown	No

B2. Archiving interviewer accounts

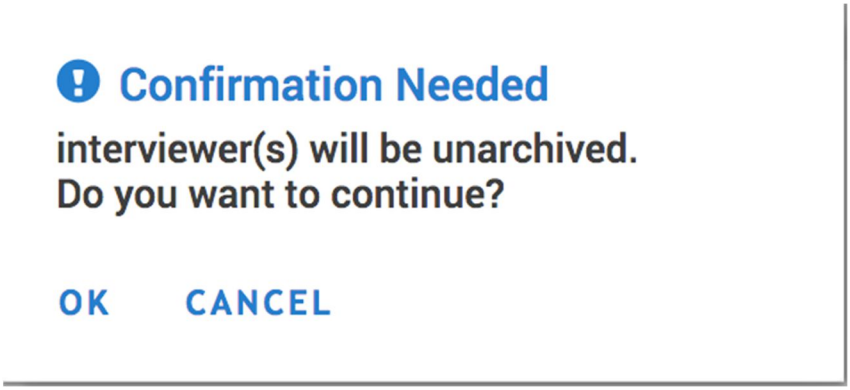
Administrators can also archive individual interviewer accounts. An interviewer will no longer be able to sync with headquarters. To archive an interviewer, select the check box associated with the interviewer(s), then click on the red archive button. Lastly, confirm the action.



Archived interviewer accounts can be restored. On the interviewers page, find the archived users by selecting *Archived users* on the menu on the left hand side. Then, select the check box associated with the interviewer(s), and click on the unarchive button at the bottom of the page. Lastly, confirm the action.



You will see the dialogue box below when you unarchive an interviewer



#2: SURVEY SETUP TAB

Under the **Survey Setup** tab, the survey manager will be able to run the survey through creating different assignments for the different team members created in Tab #1. The survey setup tab lets you do several things:

Questionnaires:

- Import template. Import questionnaires from the Designer site to the headquarters software.
- Create assignments: Headquarters provides two ways of creating assignments for a given questionnaire: one at a time, or in a many at a time (batch).
- Clone questionnaire template: Copy the questionnaire templates previously imported to the server. This feature is only available to the administrator.
- Delete questionnaire. This feature is only available to the administrator.

Assignments:

- Monitor assignments. Displays a tabular view of all assignment and enables you to edit the quantity of interviews for any assignment.

C. QUESTIONNAIRE

C1. Import questionnaire template

Before the questionnaire template is imported, it is important to test the questionnaire using the TESTER app prior, as well as during the training of interviewers. All translations and country specific modifications should be incorporated prior to the importing of the template and the beginning of the data collection.

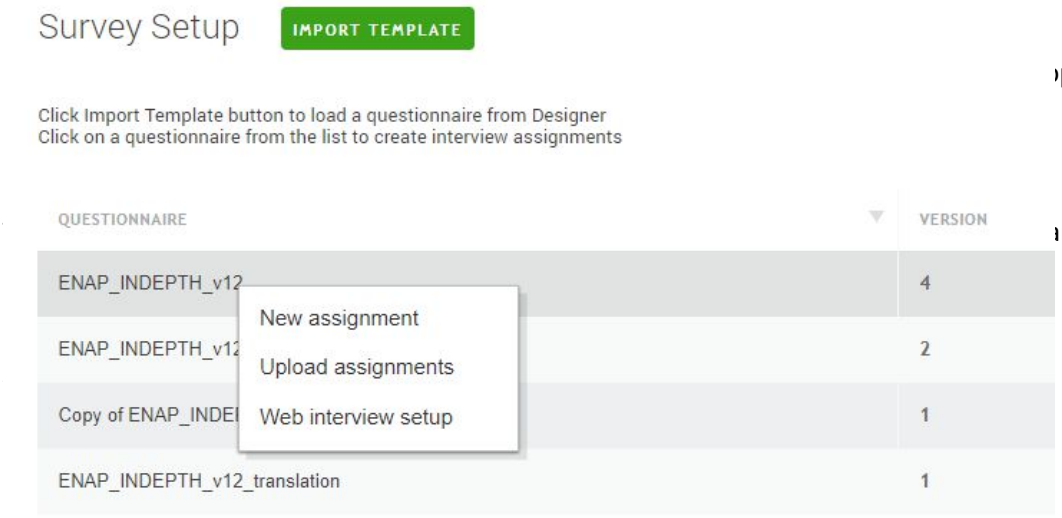
The final survey questionnaire will be uploaded to the HQ by the overall survey administrator (vladimir.gordeev@lshtm.ac.uk), upon request made by your country admin.

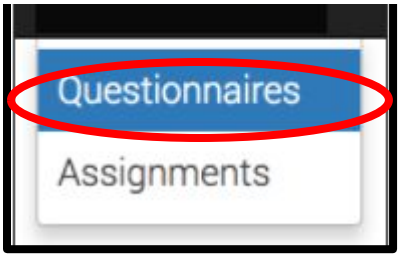
The major cumulative updates to the questionnaire would be applied once in 4-6 weeks following the feedback from all countries, unless errors are very critical. Please note that any corrections, whether minor or big, would require re-importing of the questionnaire into HQ which will also create a new separate database, hence, frequent re-uploading is not recommended if it can be avoided.

C2. Create assignments (TO INTERVIEWERS AND SUPERVISORS)

Once the questionnaire template has been uploaded into the survey headquarters questionnaire setup system. It is now possible to start distributing assignments to the different interviewers or supervisors.

There are two options available for creating assignments. One at a time and many assignments at a time.





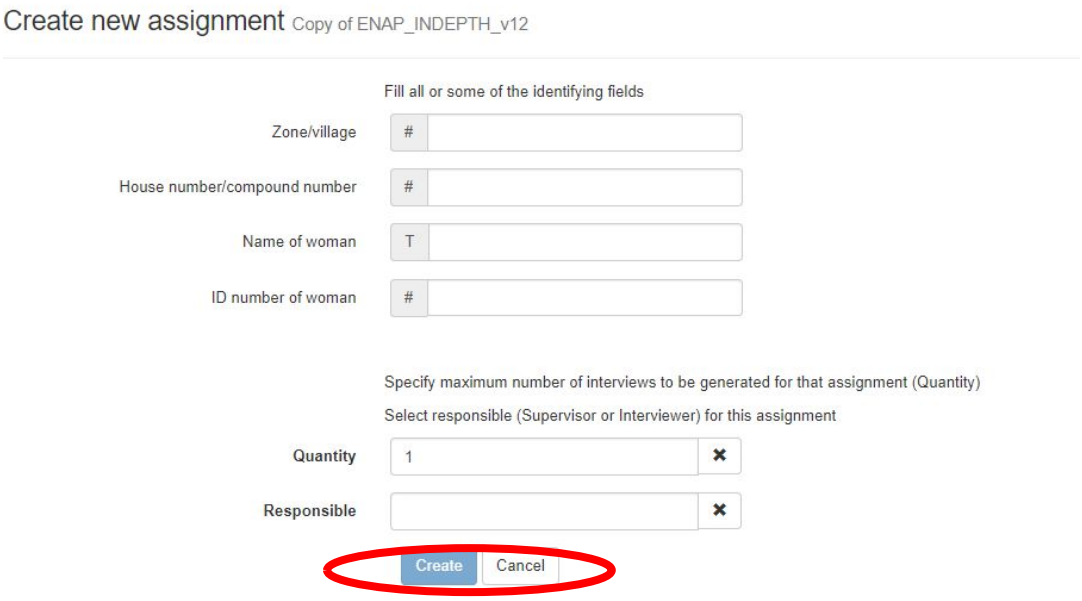
In order to create a new assignment, we need to click on the questionnaire template that you want to create an assignment for and then click on the *New Assignment* from the pop-up menu that you want to send out.

NOTE: *In the diagram below, the option for importing a questionnaire template is highlighted in green. However, as already mentioned under the uploading questionnaire template above, only the survey administrator should do this task as he has the sole access to the final questionnaire template from the designer account. In case any changes are made to the template, the INDEPTH-ENAP technical secretariat will communicate to the HDSS teams and will also guide them on how to update or upload the new questionnaire.*

STEP 3

After clicking on the new assignment option, a new window will open up. This is where you should fill in the identifier variables for the interviewer to be able to identify the respondent correctly.

After specifying the identifier information, the quantity and the Interviewer responsible, you can now complete the assignment creation by clicking on the create button.



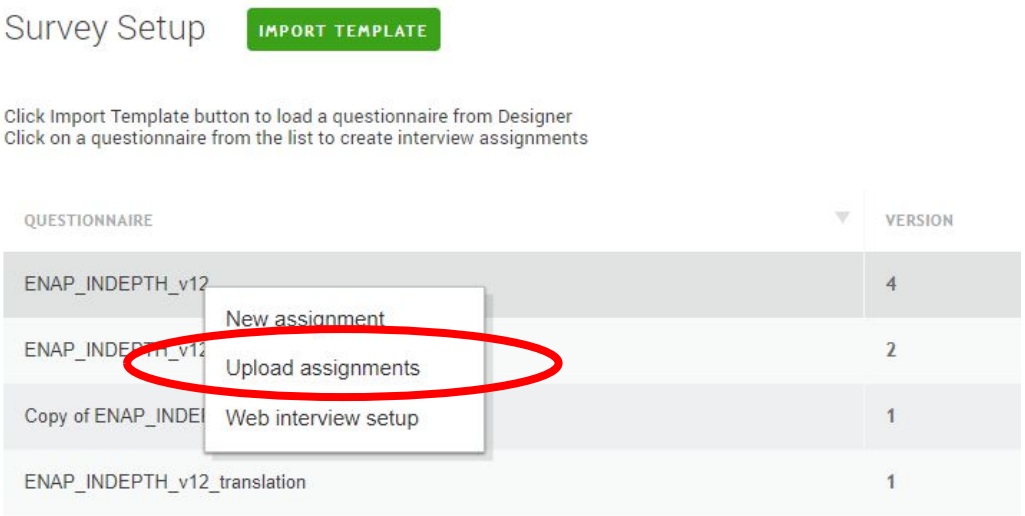
The only downside of this option is that you to repeat this process over and over again in case you have many new assignments to the interviewers.

OPTION TWO: creating several assignments at a time

Before choosing this option you need to first create a tab-delimited text file with the unique identifiers for the woman who has had a pregnancy in the last five years within the HDSS setting. Examples of these identifiers may include, study ID, woman's ID, name, age, residence etc. These set of variables need to be compiled by the HDSS statistician or data manager.

Repeat STEP 1 and 2 for option one

When using this option, of creating assignments, you will need to repeat the same procedures for steps one and two above for creating one assignment at time. But this time you should select upload multiple assignments.

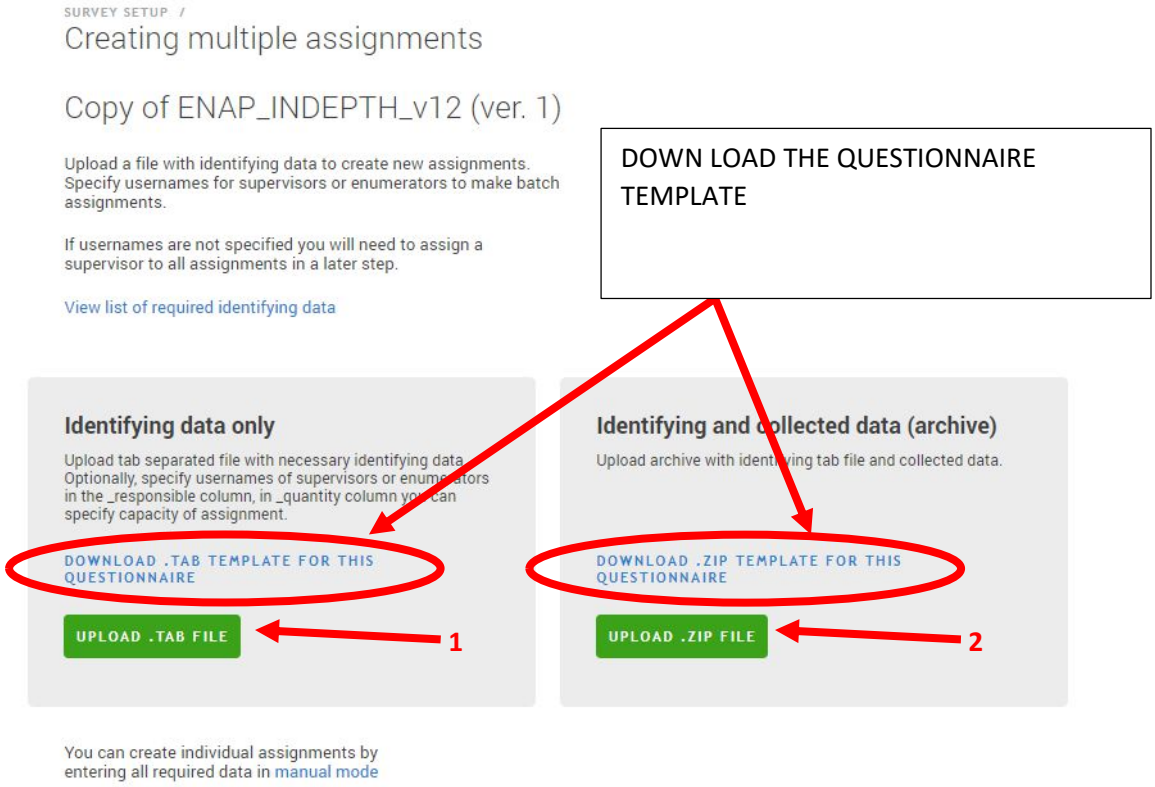


After selecting the Upload assignments, you will be re-directed to another page on which you will be able to upload the assignments from an external tab delimited file that contains the same information for each case to be created. Two options are allowed while uploading the external file in order to pre-fill the survey LISTING cases in the batch upload.

Option 1 - Identifying data only: This option is typically used for cross-sectional surveys, which will load only respondents address information (e.g., region, enumeration area, village, name of household head)

NOTE: FOR THE INDEPTH-ENAP SURVEY, WE WILL ALWAYS SELECT THIS OPTION

Option 2 – Identifying and collected data: This is an advanced mode which involves preloading information. This option is typically used fir panel surveys, which not only information on the respondent's address, but also information for any other field of the questionnaire (e.g., name and age of al household members from last wave of data collection)



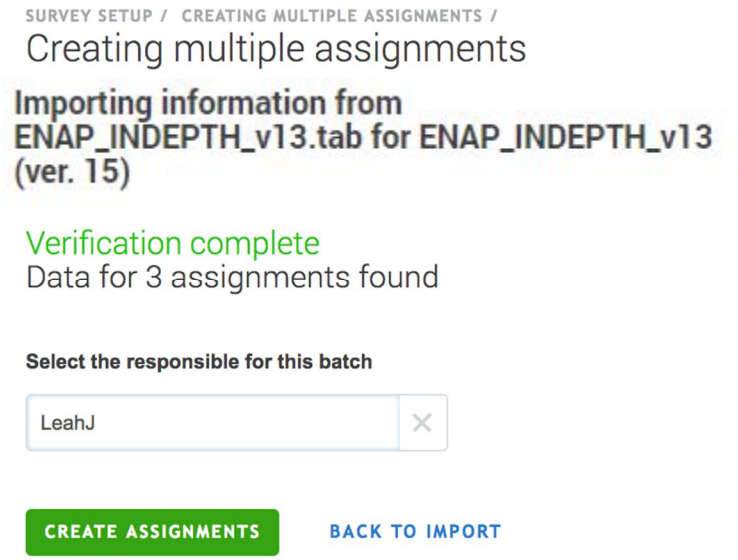
NOTE: To upload a .TAB FILE, you can either create a file with variable names of identifying questions or download the template

STEP 4

Two options are available, OPTION ONE (You upload a batch of assignments for every individual (person responsible at ago and OPTION TWO (You upload all assignments to everyone supervisors or interviewers in one text file).

Upload the prepared listing or identifier data in text-delimited version. Select identifying data only if you have the prepared file on your PC. This is the only option where you can specify your interviewers and supervisors a priori. If your upload is successful, the system will tell you that it is successful. If it is unsuccessful, the system will also notify you.

UNSUCCESSFUL ATTEMPT



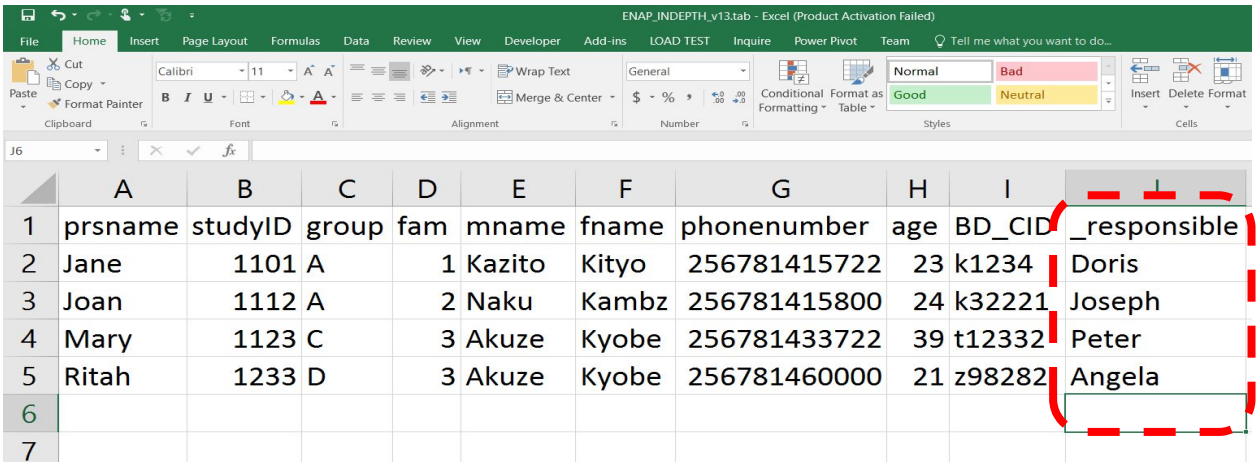
NOTE

OPTION ONE: At this point you are able to select the person responsible from the batch of the survey assignments. This is possible by selecting the username from the drop-down menu. For this to be possible, the HQ must upload one .TAB text file per person responsible (e.g., upload five cases at ago)

OPTION TWO: The good news is that you can upload one .TAB text for the survey by adding a column in the .TAB file with the heading "_responsible" and populating all the fields for the identified cases ready for interviews with the appropriate interviewer or supervisor usernames that we created in #1.

Below is a sample file of Listing variables that "may be" appropriate for identifying a woman required for interviewing into the survey.

Remember when you create a spreadsheet of the cases you want to interview in the next set of assignments, you should save it as a TAB delimited file as we have already seen under #1.

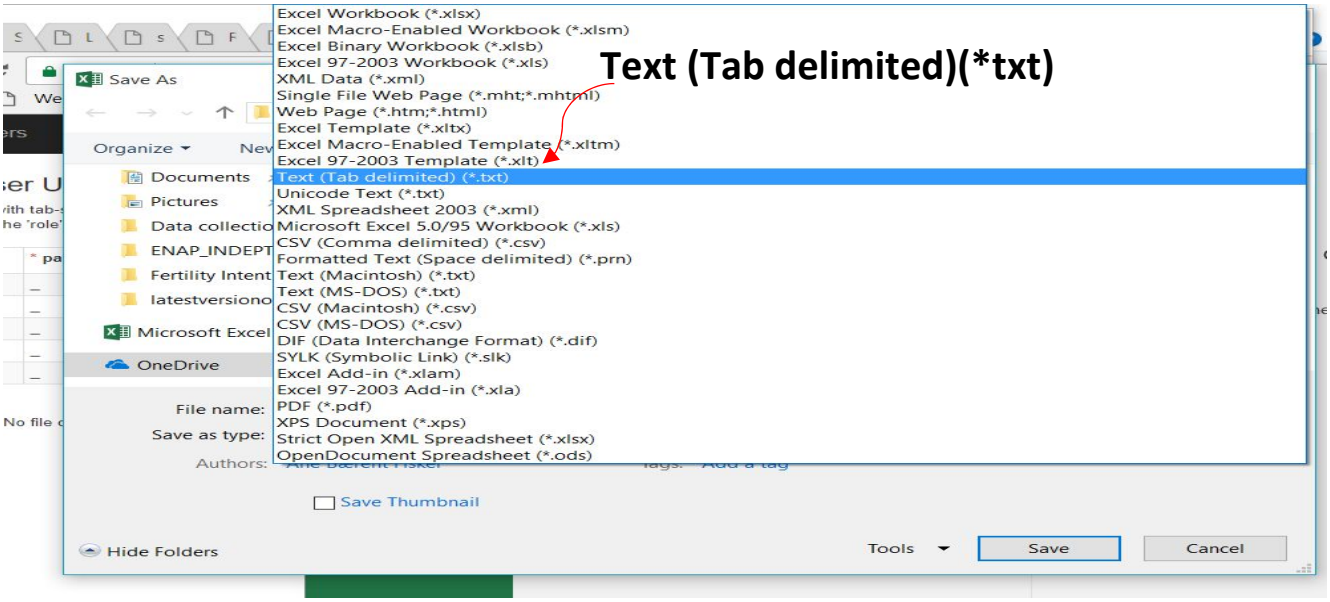


	A	B	C	D	E	F	G	H	I	J
1	prsname	studyID	group	fam	mname	fname	phonenumber	age	BD_CID	_responsible
2	Jane	1101	A	1	Kazito	Kityo	256781415722	23	k1234	Doris
3	Joan	1112	A	2	Naku	Kambz	256781415800	24	k32221	Joseph
4	Mary	1123	C	3	Akuze	Kyobe	256781433722	39	t12332	Peter
5	Ritah	1233	D	3	Akuze	Kyobe	256781460000	21	z98282	Angela
6										
7										

BEGINNING REPEATED SECTION

Once you have added the fields with the same exact wording, you should now save the workbook on your computer. Save as type "**Text (Tab-delimited) (*.txt)**" file.

Ensure that you save it in a location that you can find easily on your computer. For example in a folder on your Desktop. **Note: Only 'Supervisor' OR 'Interviewer' are allowed in the 'role' column.**



END OF A REPEATED SECTION

NOTE: Similarly with this approach you do not need to load a different file for each person responsible (e.g.. if you have 5 supervisors each supervising 5 interviewers and the interviewers are required to do 5 interviews a day, you should upload all the 125 cases (5 supervisors x 5 interviewers x 5 interviews = 125 cases).

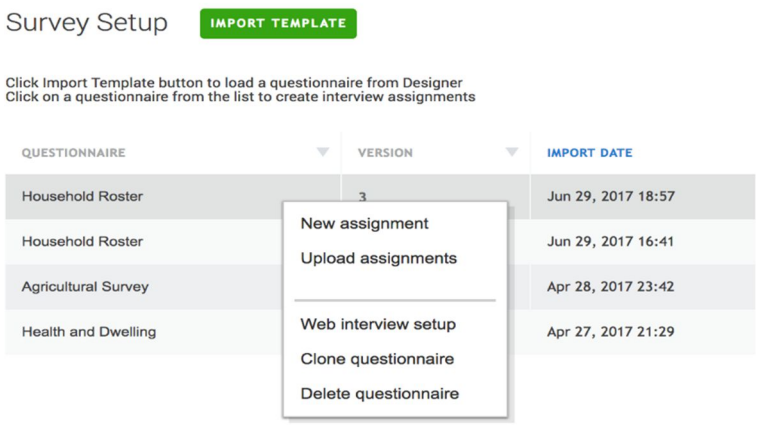
To use this option when you download the LISTING template from the survey solutions page for the questionnaire template, always remember to add an additional column to the end named as _responsible

When the upload is successful or unsuccessful, you will be notified by the system as it was shown for the first option above.

We recommend that you use option 2.

C3. Copy questionnaire templates (ADMINISTRATORS ONLY)

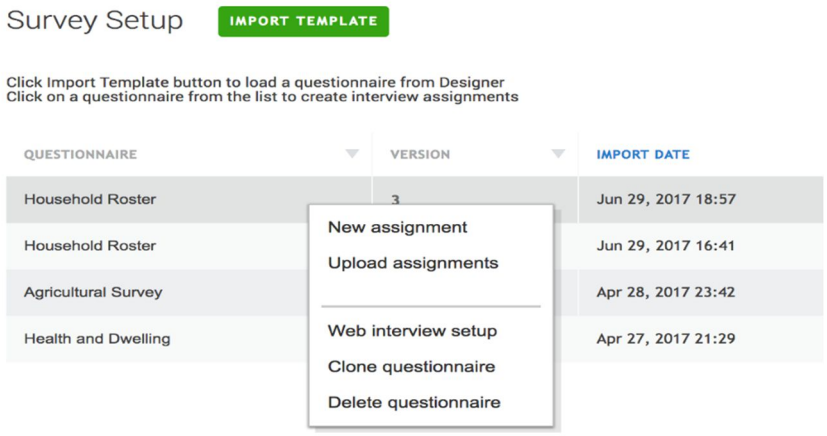
Cloning questionnaires on headquarters allows you to start a new survey using the same questionnaire that has been imported to the HQ previously. To copy a questionnaire in headquarters, click on the questionnaire you would like to copy and then select Clone questionnaire from the menu.



C4. Delete questionnaire template (ADMINISTRATORS ONLY)

IMPORTANT!!! This feature is only available to the administrator. It should only be used under special circumstances with extreme caution because it will delete all the data collected with that questionnaire from the server. Please beware - it is not possible to recover deleted data.

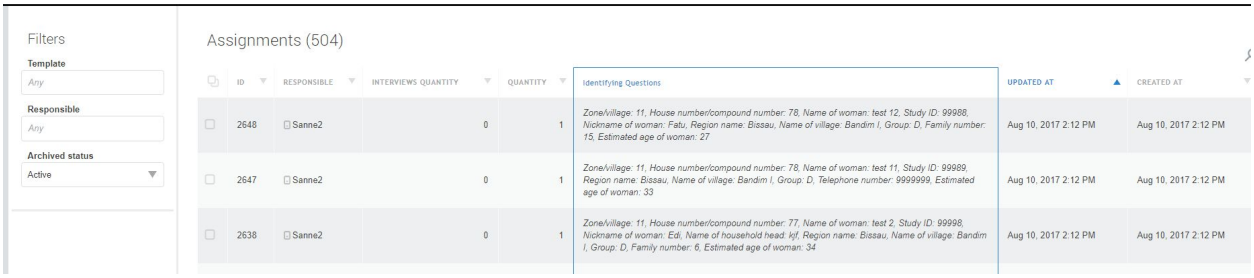
To delete a questionnaire, click on the questionnaire you would like to delete and select Delete questionnaire from the menu.



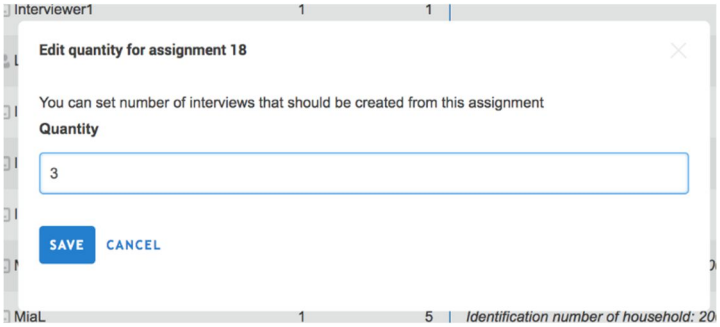
D. ASSIGNMENTS

At this point, you will be able to see the details of the assignments including the date created, modified, ID number and the person responsible for that assignment under the Assignments tab (HQ>>Survey Setup>>Assignments).

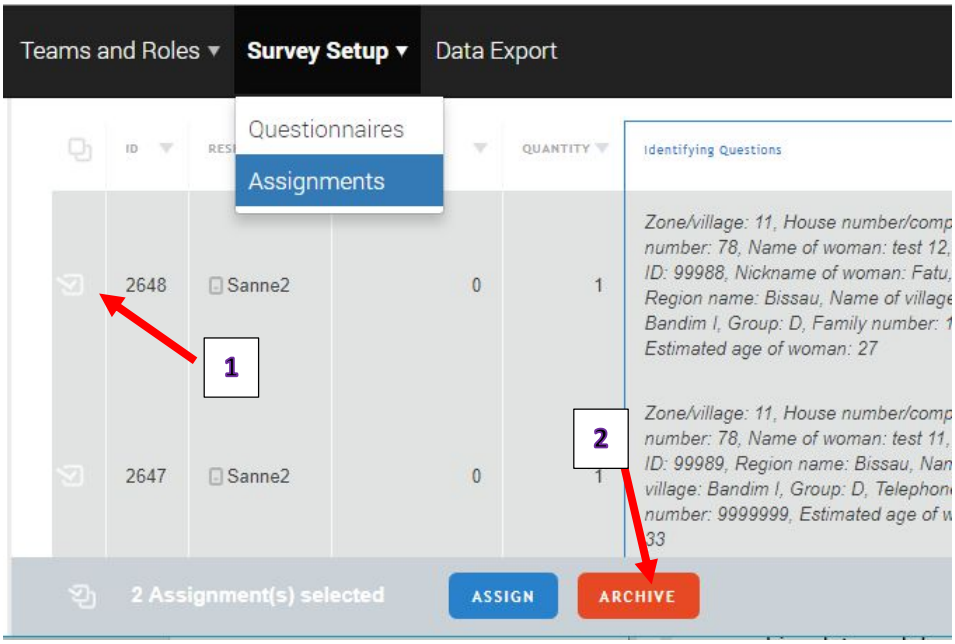
Under the Assignments, you will see the ID of the assignment, the person responsible (Supervisor/Interviewer), Quantity of interviews conducted under that assignment ID, the dates when the assignment was created and when it was updated E.g. by the interviewer. You can also see on the left hand side panel you are able to filter the assignments by template, person responsible and Archived status.



The quantity of interviews for each assignment can be changed on this page. This is especially useful for cases where Survey Solutions is being used for a listing exercise and you wish to set or later change a maximum number interviews for each assignments. To change the quantity, click on the number in the quantity column and edit the field before clicking on *Save*.



Under the Assignments, you can Assign and Archive the different assignments. This is possible by clicking on the check box in the column to the left of the ID column. See options 1 and 2 below.



NOTE:

FOR DETAILED NOTES ON HOW TO REVIEW INTERVIEWS ARE GIVEN IN THE HEADQUARTERS MANUAL UNDER TAB #3: INTERVIEWS

A6: Survey Solution's Headquarters Manual

INTRODUCTION

This manual contains a detailed guide on how to operate the HQ. It aims at guiding the survey manager/statistician/data manager at the HDSS sites during the survey implementation and data collection.

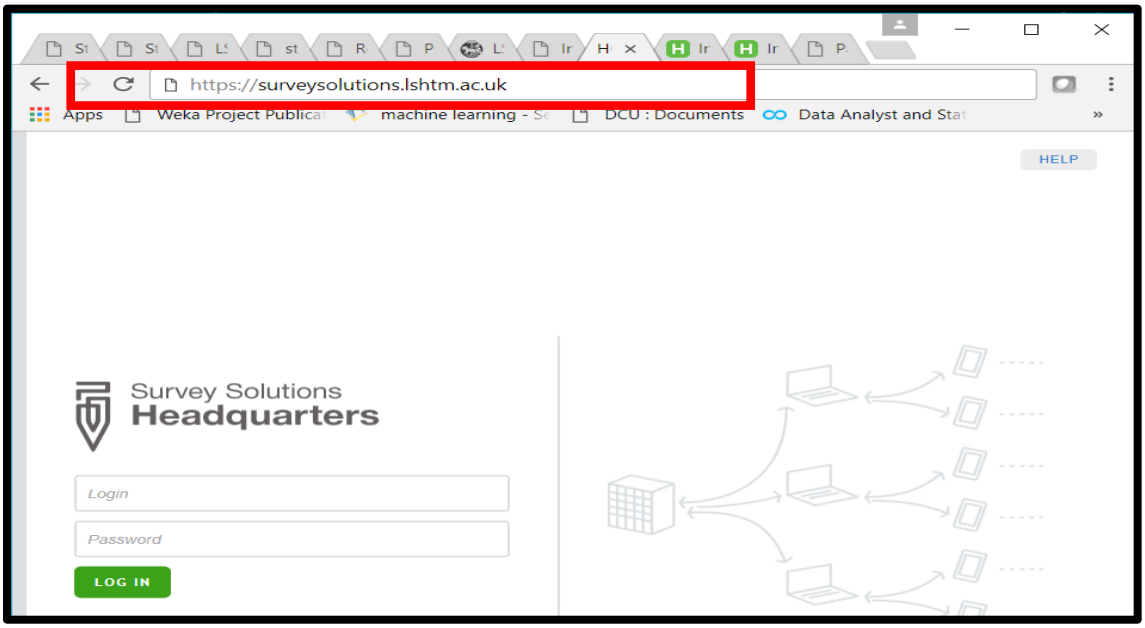
The HQ contains a number of connected tools for the administrator and headquarter users, which helps to track the overall progress of the survey (**Reports tab**), to review completed interviews (**Interviews tab**), to manage the human resources (**Teams and Roles tab**), to specify survey instruments, create survey assignments with those instruments (**Survey Setup**), to export the data collected from these assignments

This manual will start by showing you how to create an HQ account, then we will explain the functionality of each tab under the HQ account in the following order: **Teams and Roles**, **Survey Setup, Interviews, Data Export, Reports, Troubleshooting and Help.**

CREATING THE HQ ACCOUNT FOR YOUR TEAMS

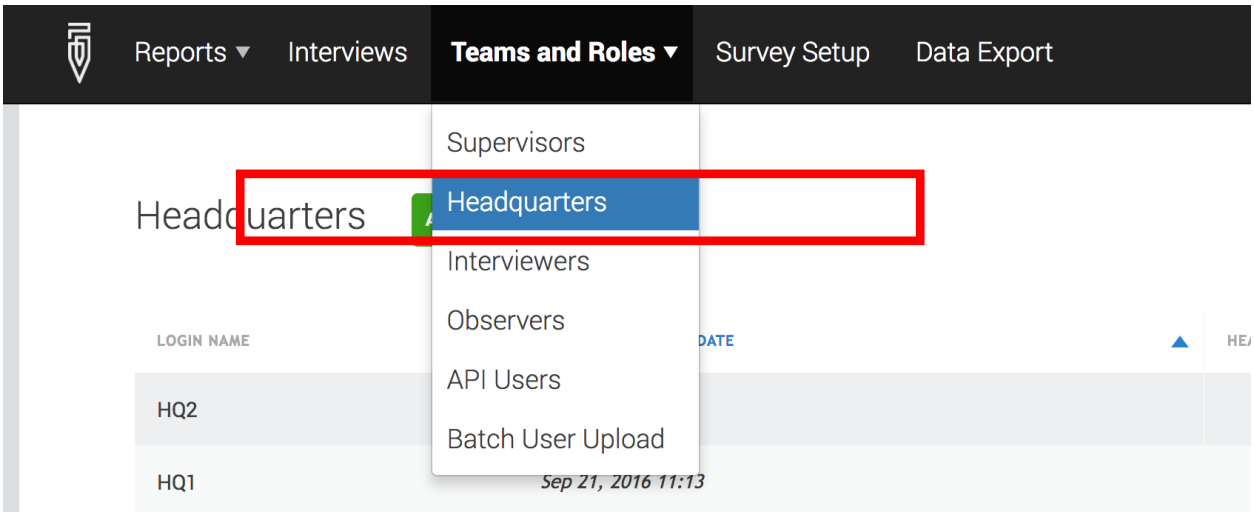
Option 1 – If you are using the LSHTM server

If you are using one of the dedicated virtual LSHTM servers, the country HQ will be set up for you. To log in into your HQ account, you will need to type the address of your dedicated virtual server into the address bar (see example below) and use the HQ username and password that will be provided to you (contact: vladimir.gordeev@lshtm.ac.uk):

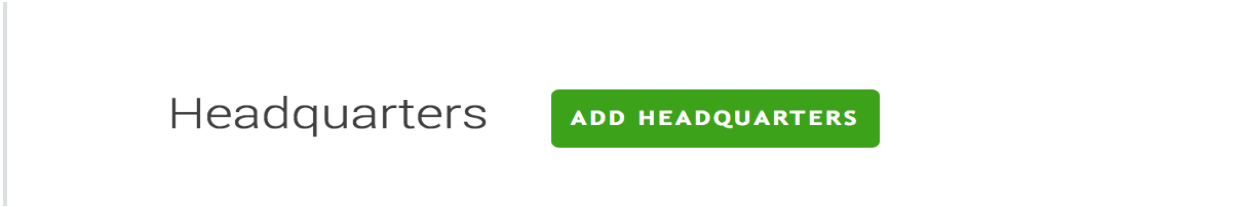


Option 2 – If you use your own server

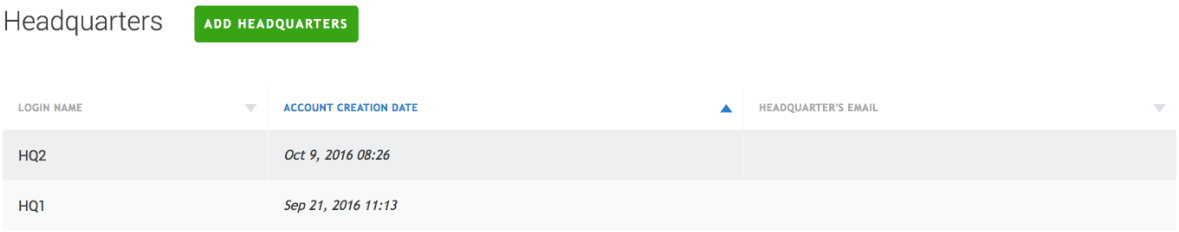
If you are using your own local server, your admin will need to create one or more HQ users. To manage HQ, the admin needs to log in into HQ, using the username and password created at the end of the installation of the Survey Solutions. Once admin logs in, to manage the HQ, click on ***Headquarters*** in the drop down menu under the ***Teams and Roles*** menu.



To add a headquarters user, click on the green ***Add headquarters*** button at the top of the page, and enter all the requisite information.

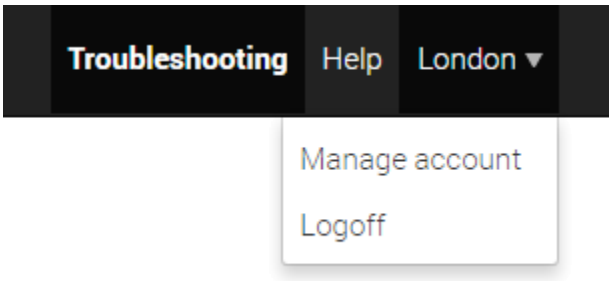


To modify a headquarters' account—change the password or lock the account—click on the headquarters username.



Now, the HQ user (or admin) can proceed with adding supervisors and interviewers. It is highly recommended to set up all accounts prior to active survey implementation and related fieldwork.

If you have successfully logged into your Survey Solutions HQ, the user name of your HQ will appear in the right top corner of the main menu. Using the drop down menu, you are able to manage your HQ, switch users and log off.



When you click on manage account you will be able to update the HQ information including: Full name, Email address, Phone numbers and also change the HQ passwords.

Note: Please remember, an admin and a headquarter user accounts have different purposes and privileges. Unlike HQ user, the admin can create headquarters/observer user accounts, archive user accounts, and delete questionnaire templates from the headquarters server. While you will have several HQ users, it is not recommended to have more than one admin per HQ and country.

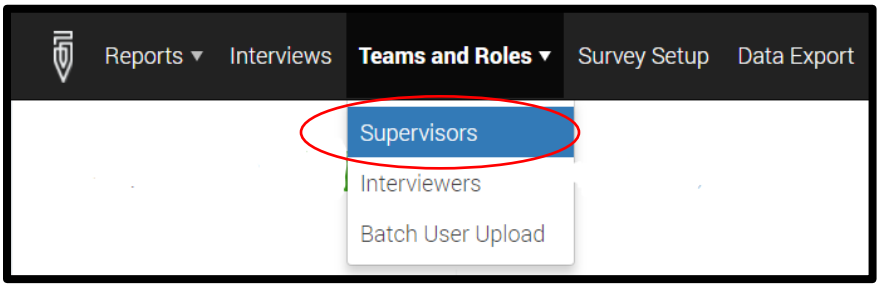
Once you have successfully logged into your HQ account by using either options outlined above, you can now setup accounts for other users, including the supervisors and interviewers using **Teams and Roles** tab.

TAB #1: TEAMS AND ROLES

Under the Teams and Roles tab, the survey data manager or statistician or survey manager is able to create the supervisors and interviewers accounts for the individuals who will participate in the survey implementation.

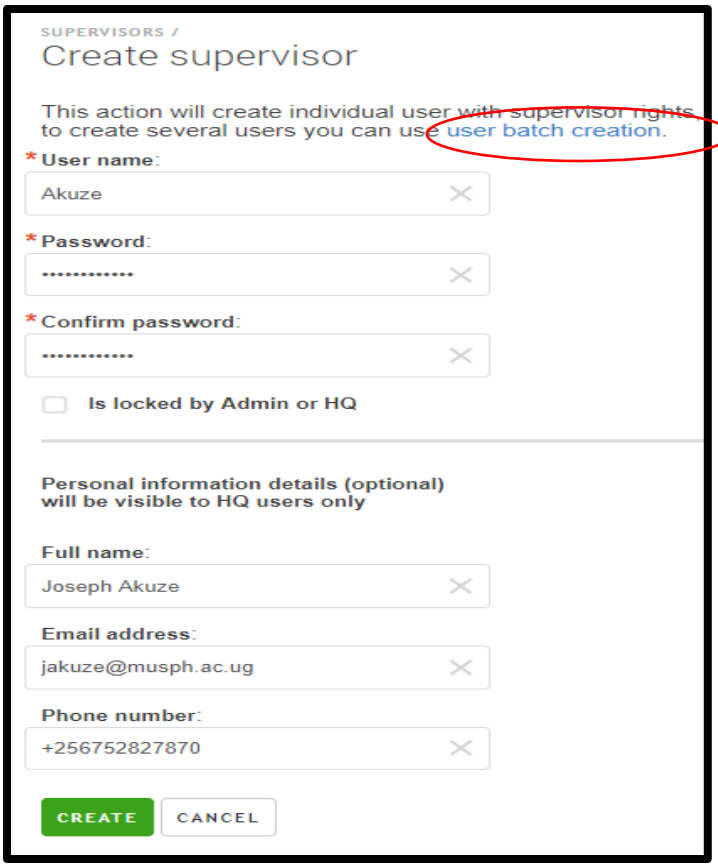
Creating accounts for supervisors

Ensure that you add supervisors for your HDSS. The supervisors responsible for assigning tasks to the interviewers and reviewing the data collectors completed questionnaires. To add a supervisor, click on the *Teams and Roles*

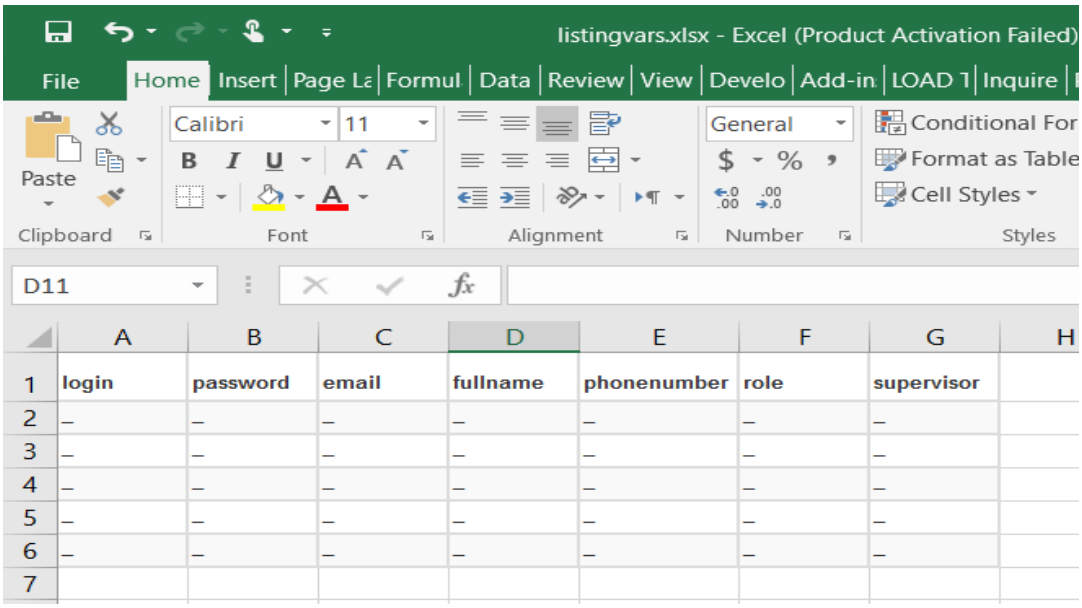


Then, complete the form displayed and click on ***CREATE*** to add the supervisor in the system.

Note: This option only allows you to create one supervisor account at a time. It is possible to create several supervisors using the ***USER BATCH CREATION***



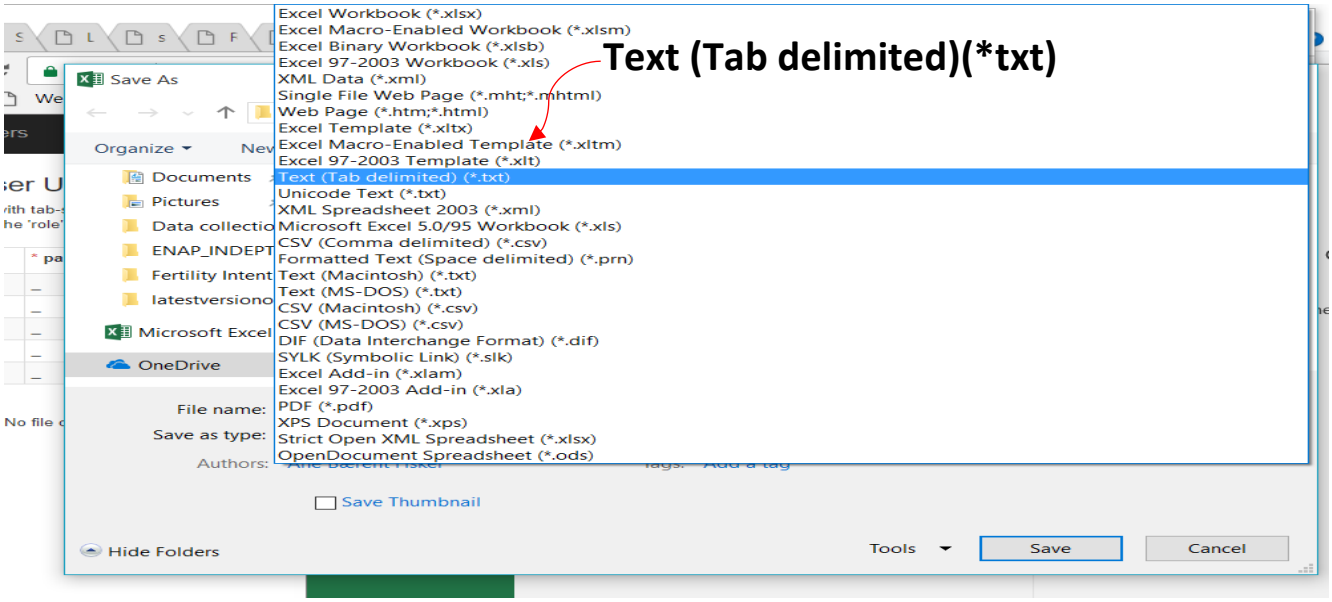
In order to upload more than one supervisor using the user batch upload. You will need to open a new MS excel workbook, add the following fields; (***login, password, email, fullname, phonenumber, role, supervisor***) as is illustrated below.



Once you have added the fields with the same exact wording, you should now save the workbook on your computer save a “tab-delimited” file. Or “.txt”

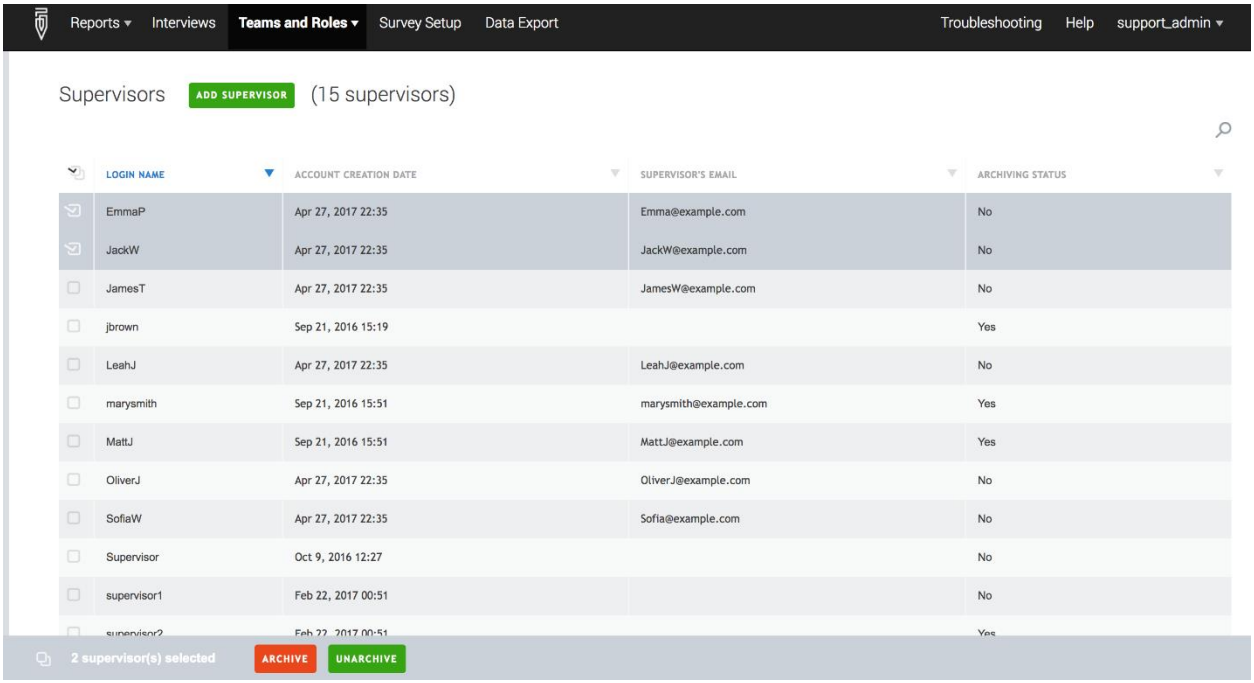
Ensure that you save it in a location that you can find easily on your computer.

Note: Only 'Supervisor' OR 'Interviewer' are allowed in the 'role' column.



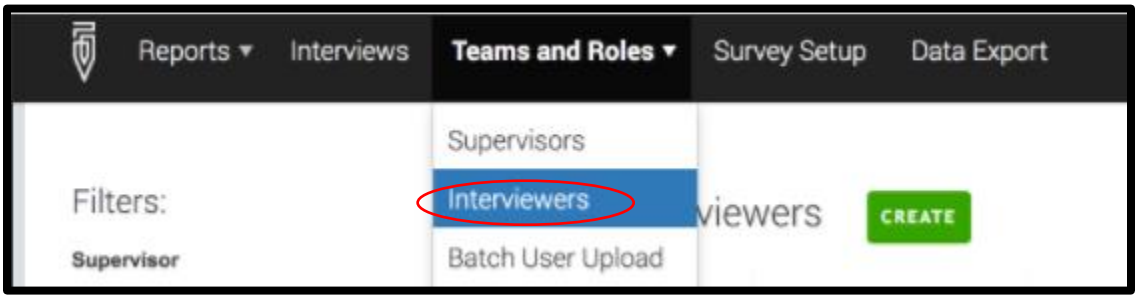
Archiving supervisor accounts

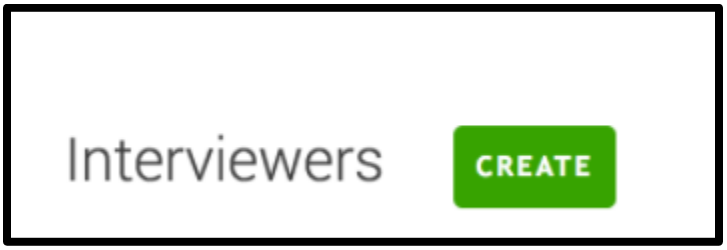
Administrators can archive supervisor accounts. By archiving a supervisor, you are also archiving all the interviewers on that supervisors' team. Supervisors and their data collection team will no longer be able to sync with headquarters. To archive a supervisor, select the check boxes for the supervisor(s) you wish to archive and click on the red archive button. Archived interviewer accounts can be restored. On the supervisors page, find the archived users using the Archive Status filter on the right most column. Then, select the check



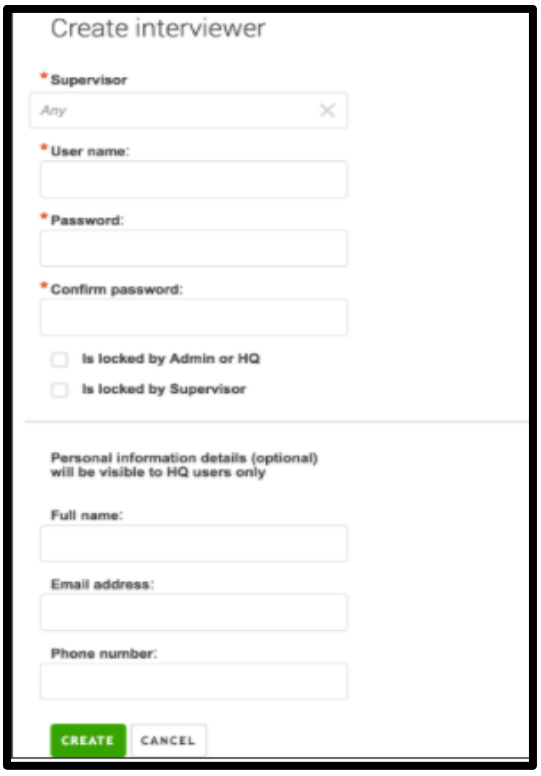
Adding interviewers

This process is similar to adding supervisors. You need to go to the **Teams and Roles** tab and then select Interviewers, then create Interviewers, then complete the form displayed.

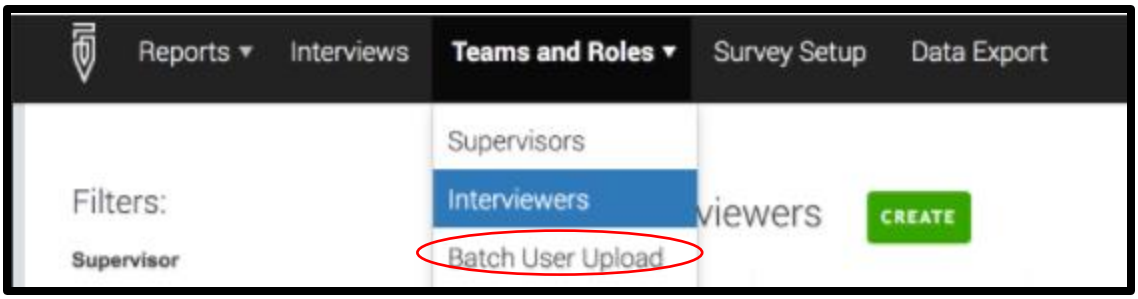




This approach allows you to upload only one Interviewer at a time.



To upload many Interviewers, you need to use the “Batch User Upload” option under the Teams and Roles tab. When you click on the Batch User Upload option under the Teams and Roles tab, you will be required to upload a text-delimited files.



The batch upload uses the same fields used when creating the supervisors' batch upload. Therefore, this means that you are able to great both the supervisors and the interviewers accounts in one click. This option saves you a lot of time.

NOTES:

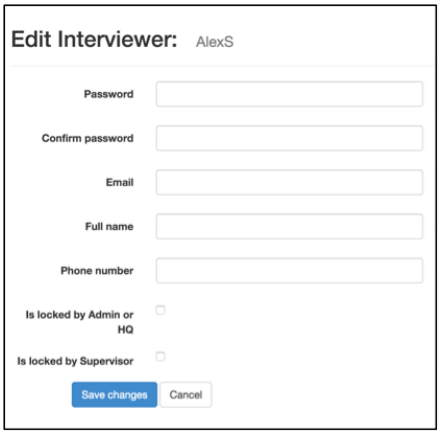
The steps of creating the text-delimited file that have been described under the supervisor batch upload option above are the same steps you need. Up to 10,000 user accounts can be created from a single tab-delimited file, which must include the login, password, and role of the user, and may include additional attributes, like email or phone.

Batch creation of user accounts comes handy when: the same accounts need to be replicated on a different server; migrating from a different CAPI system with established accounts; the survey is so huge that manual creation of accounts via a web interface would take a significant amount of time.

Creating user accounts in batch mode is available to both the administrator and the HQ users. For new supervisor accounts the role must be specified as Supervisor (case-sensitive). For interviewer accounts the role must be specified as Interviewer (case-sensitive), and in addition the login of the supervisor must be specified, which determines the team where the interviewer account will be added. Interviewers may be added to existing supervisors, or to new supervisors mentioned anywhere in the users list. When creating new users with a batch upload feature, select the Batch user upload menu item of the Teams and Roles menu.

Modifying data related to interviewers and monitoring interviewers

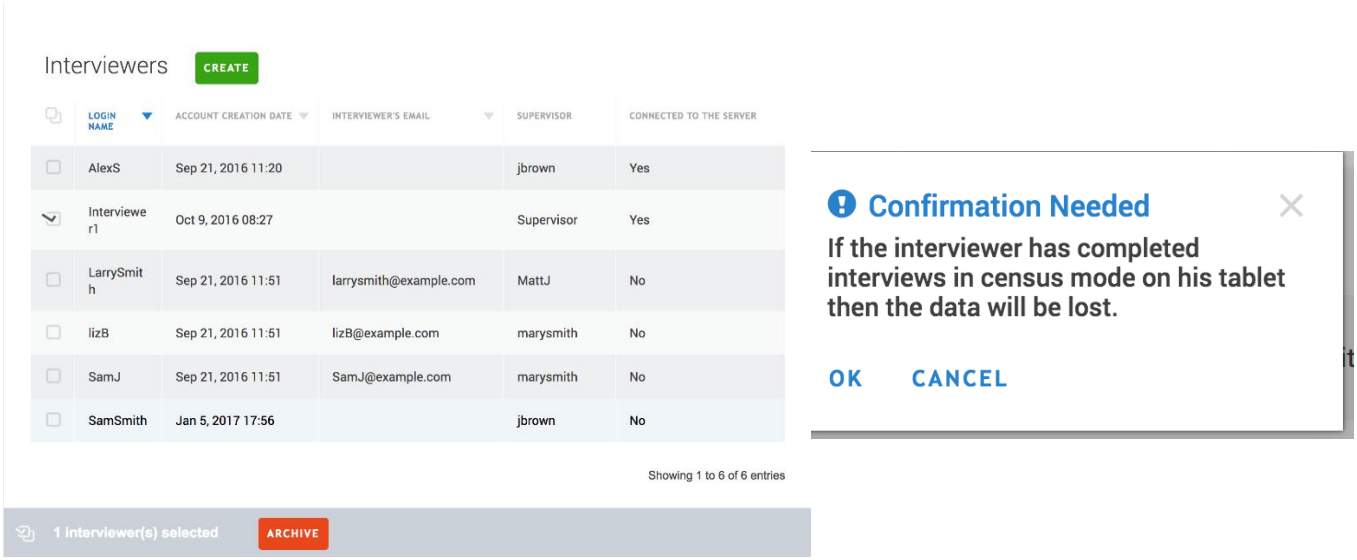
To modify an interviewer's account—change the password or lock the account—click on the interviewer's name. To monitor an interviewer, check the activation status in the Connected to device column and the synchronization log. The Connected to device column displays the status of the tablet's activation or in other words it will show if an interviewer has successfully activated a tablet using their login credential



LOGIN NAME	ACCOUNT CREATION DATE	INTERVIEWER'S EMAIL	SUPERVISOR	CONNECTED TO THE SERVER
AlexS	Sep 21, 2016 11:20		jbrown	Yes
lizB	Sep 21, 2016 11:51	lizB@example.com	marysmith	No
SamJ	Sep 21, 2016 11:51	SamJ@example.com	marysmith	No
LarrySmith	Sep 21, 2016 11:51	larrysmith@example.com	MattJ	No
Interviewer 1	Oct 9, 2016 08:27		Supervisor	Yes
SamSmith	Jan 5, 2017 17:56		jbrown	No

Archiving interviewer accounts

Administrators can also archive individual interviewer accounts. An interviewer will no longer be able to sync with headquarters. To archive an interviewer, select the check box associated with the interviewer(s), then click on the red archive button. Lastly, confirm the action.



Archived interviewer accounts can be restored. On the interviewers page, find the archived users by selecting *Archived users* on the menu on the left hand side. Then, select the check box associated with the interviewer(s), and click on the unarchive button at the bottom of the page. Lastly, confirm the action.

TAB #2: SURVEY SETUP: import, copy, and delete questionnaire(s)

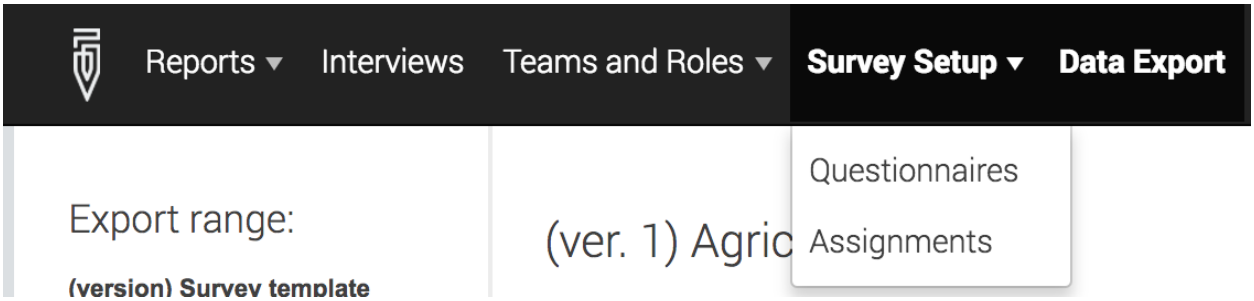
Under **Survey Setup** tab, the survey manager will be able to run the survey through creating different assignments for the different team members created in Tab #1. The survey setup tab lets you do several things:

Questionnaires:

- Import template. Import questionnaires from the Designer site to the headquarters software.
- Create assignments: Headquarters provides two ways of creating assignments for a given questionnaire: one at a time, or in a many at a time (batch).
- Clone questionnaire template: Copy the questionnaire templates previously imported to the server. This feature is only available to the administrator.
- Delete questionnaire. This feature is only available to the administrator.

Assignments:

- Monitor assignments. Displays a tabular view of all assignment and enables you to edit the quantity of interviews for any assignment.



Uploading questionnaire template

Before the questionnaire template is imported, it is important to test the questionnaire using the TESTER app prior, as well as during the training of interviewers. All translations and country specific modifications should be incorporated prior to the importing of the template and the beginning of the data collection.

The final survey questionnaire will be uploaded to the HQ by the overall survey administrator (vladimir.gordeev@lshtm.ac.uk), upon request made by your country admin.

The major cumulative updates to the questionnaire would be applied once in 4-6 weeks following the feedback from all countries, unless errors are very critical. Please note that any corrections, whether minor or big, would require re-importing of the questionnaire into HQ which will also create a new separate database, hence, frequent re-uploading is not recommended if it can be avoided.

Creating assignments

Once the questionnaire template has been uploaded into the survey headquarters questionnaire setup system. It is now possible to start distributing assignments to the different interviewers.

There are two options available for creating assignments. One at a time and many assignments at a time.

OPTION ONE: Creating one assignment at a time

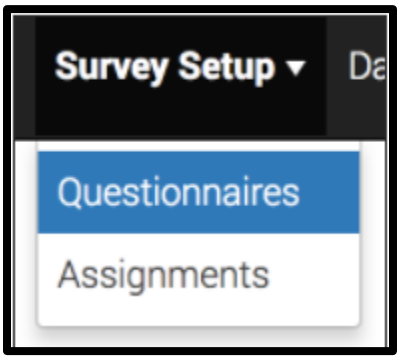
With this option, you can only upload only one assignment at a time. This is not the best option for issuing assignments as you will have to add assignments one at a time.

STEP 1

The HDSS data manager should log into the server using the headquarters username and password as already shown earlier.

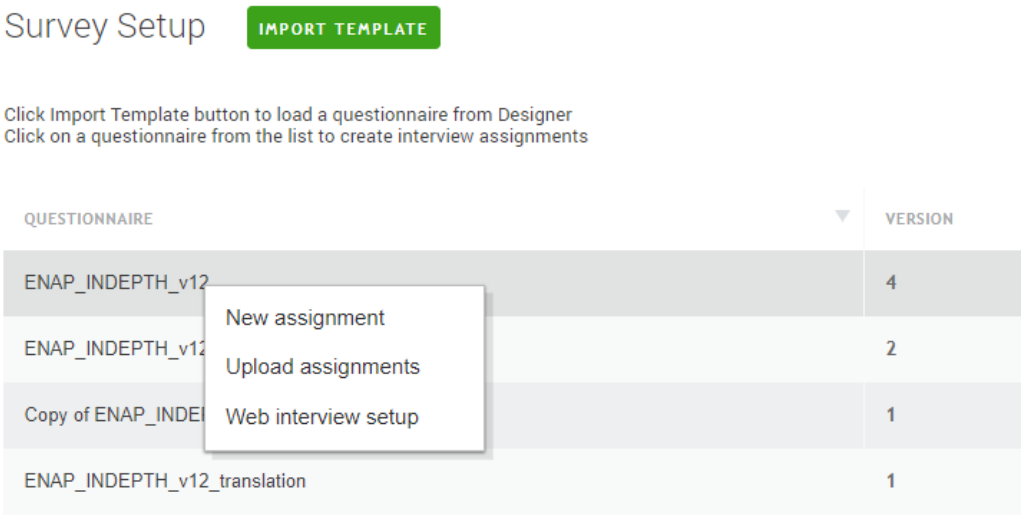
STEP 2

You should select the ***Survey Setup*** tab on the main menu, and click on **Questionnaires.**



In order to create a new assignment, we need to click on the questionnaire template that you want to create an assignment for and then click on the *New Assignment* from the pop-up menu that you want to send out.

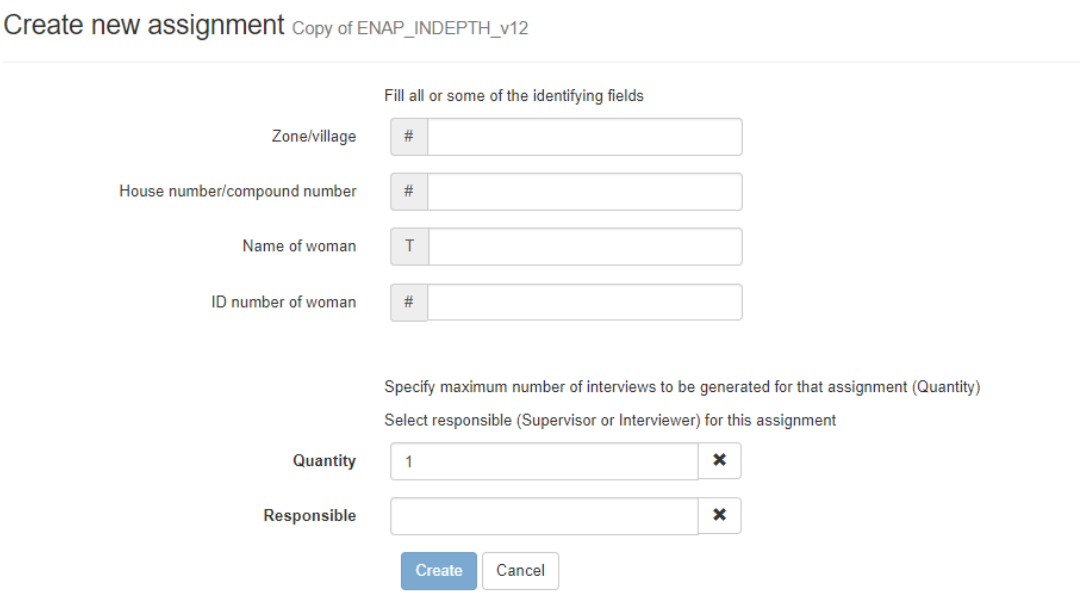
NOTE: *In the diagram below, the option for importing a questionnaire template is highlighted in green. However, as already mentioned under the uploading questionnaire template above, only the survey administrator should do this task as he has the sole access to the final questionnaire template from the designer account. In case any changes are made to the template, the INDEPTH-ENAP technical secretariat will communicate to the HDSS teams and will also guide them on how to update or upload the new questionnaire.*



STEP 3

After clicking on the new assignment option, a new window will open up. This is where you should fill in the identifier variables for the interviewer to be able to identify the respondent correctly.

After specifying the identifier information, the quantity and the Interviewer responsible, you can now complete the assignment creation by clicking on the create button.



The only downside of this option is that you to repeat this process over and over again in case you have many new assignments to the interviewers.

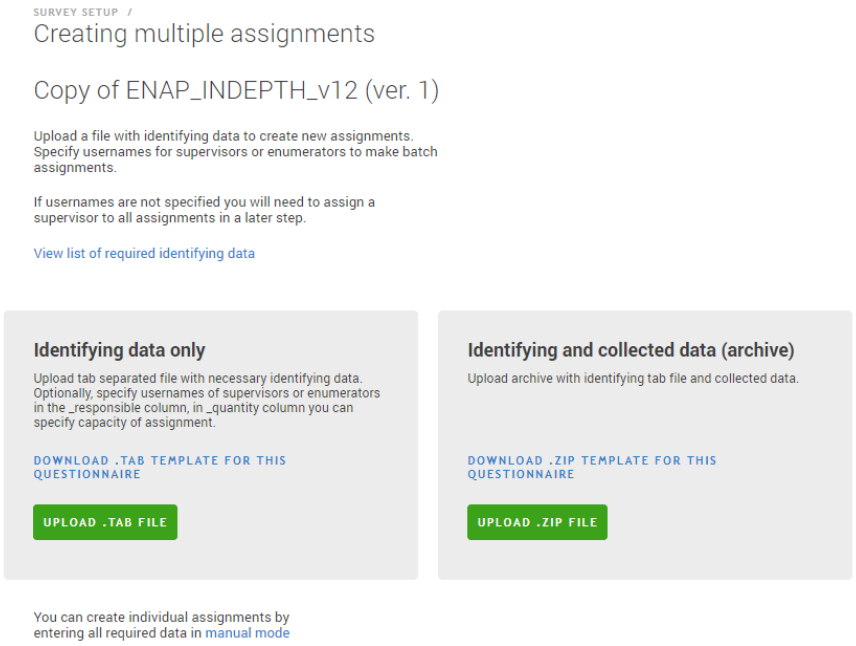
OPTION TWO: creating several assignments at a time

Before choosing this option you need to first create a tab-delimited text file with the unique identifiers for the woman who has had a pregnancy in the last five years within the HDSS setting. Examples of these

identifiers may include, study ID, woman's ID, name, age, residence etc. These set of variables need to be compiled by the HDSS statistician or data manager.

Repeat STEP 1 and 2 for option one

When using this option, of creating assignments, you will need to repeat the same procedures for steps one and two above for creating one assignment at time. But this time you should select upload multiple assignments.

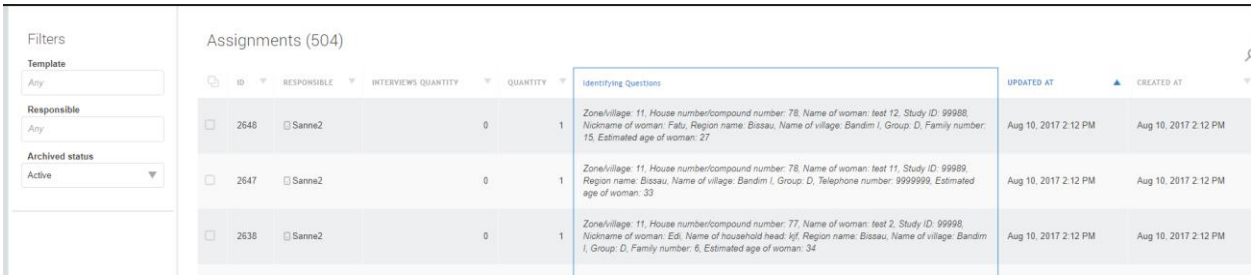


STEP 4

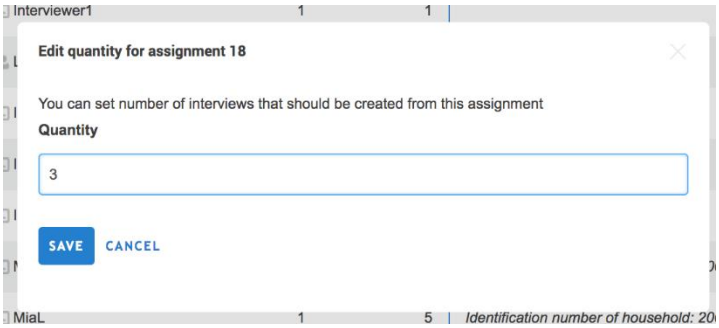
Upload the prepared listing or identifier data in text-delimited version. Select identifying data only if you have the prepared file on your PC. This is the only option where you can specify your interviewers and supervisors a priori. If your upload is successful, the system will tell you that it is successful.

At this point, you will be able to see the details of the assignments including the date created, modified, ID number and the person responsible for that assignment under the Assignments tab (HQ>>Survey Setup>>Assignments).

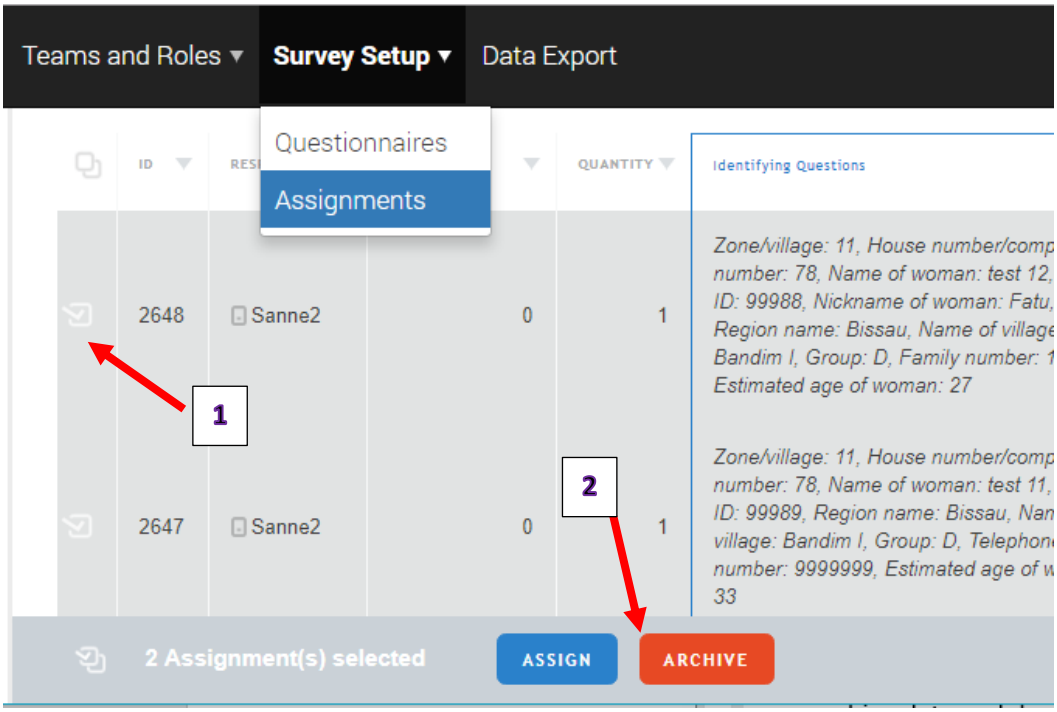
Under the Assignments, you will see the ID of the assignment, the person responsible (Supervisor/Interviewer), Quantity of interviews conducted under that assignment ID, the dates when the assignment was created and when it was updated E.g. by the interviewer. You can also see on the left hand side panel you are able to filter the assignments by template, person responsible and Archived status.



The quantity of interviews for each assignment can be changed on this page. This is especially useful for cases where Survey Solutions is being used for a listing exercise and you wish to set or later change a maximum number interviews for each assignments. To change the quantity, click on the number in the quantity column and edit the field before clicking on *Save*.



Under the Assignments, you can Assign and Archive the different assignments. This is possible by clicking on the check box in the column to the left of the ID column. See options 1 and 2 below.

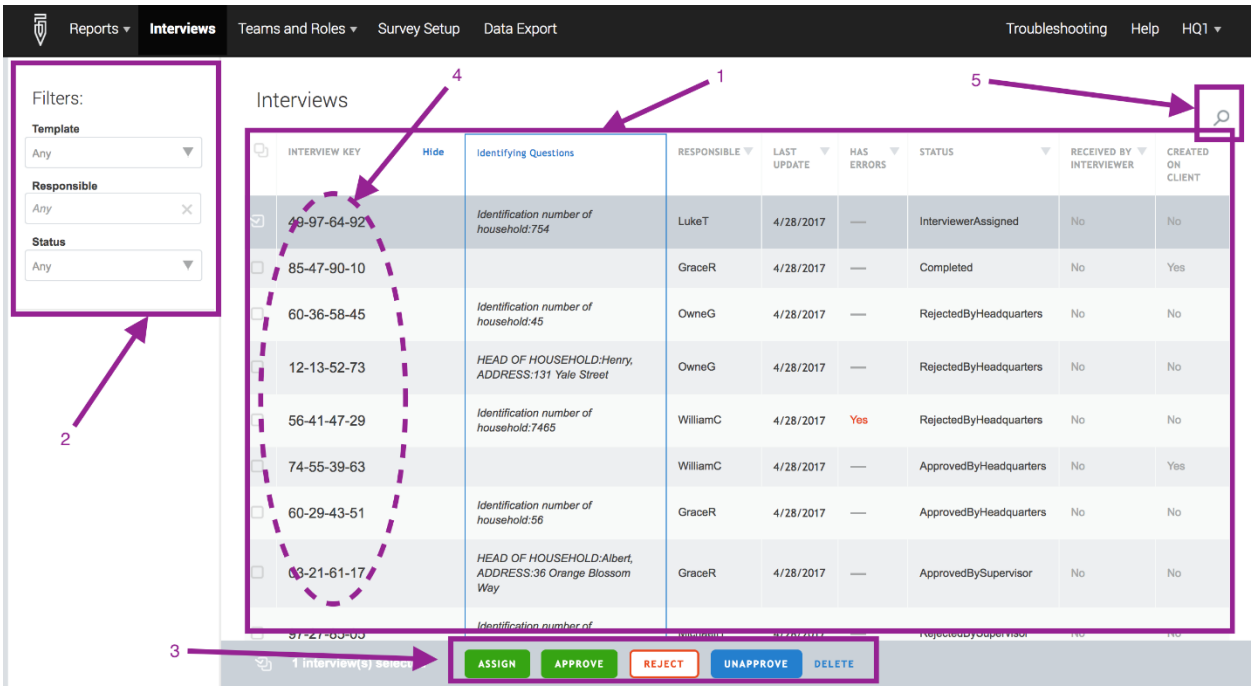


TAB #3: INTERVIEWS

The Interview tab is designed for taking action. Under the Interview tab, headquarters may search for survey cases, reallocate them to different teams, review them for quality assurance, and delete them to fix problems.

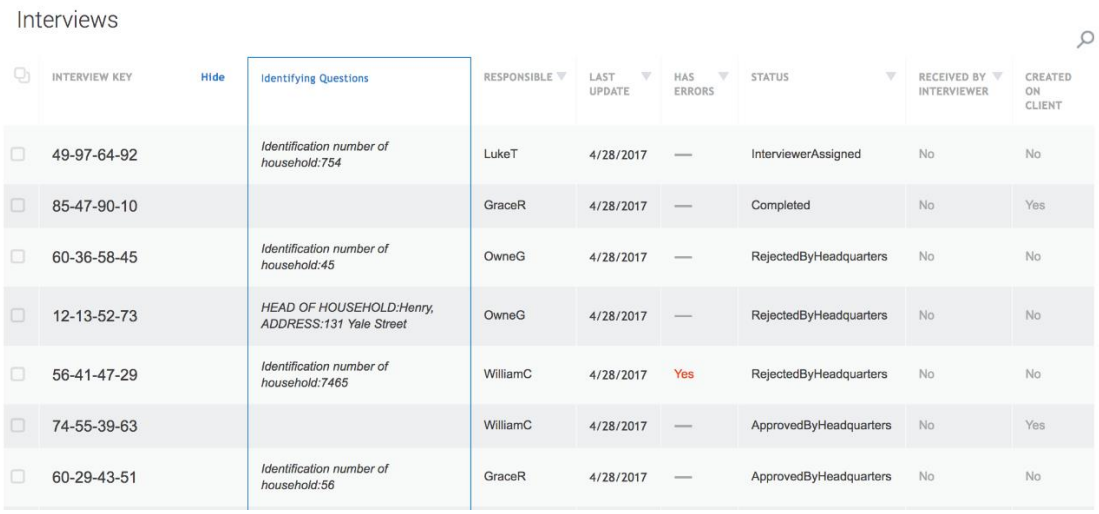
The interview tab consists of the following tools:

1. **Database of survey cases.** The rows are survey case entries. The columns capture details about survey cases, such as the respondent's address, interview status, and an indicator that the interviewer has or has not received the survey case.
2. **Filters for querying the database.** The database can be filtered to show only the survey cases of interest—for example, those with a particular status or those done by a particular supervisor.
3. **Toolbar with a Delete, Approve, and Reject button**. Survey cases can be deleted, approved, or rejected individually or as a group by selecting the check box next to each case. The toolbar will appear if one or more cases have been selected. If you would like to select all the survey cases displayed on that page, click on the first check box at the top of the list.
4. **Interview Key.** This is a system generated random number for each survey assignment. Completed survey cases can be reviewed in detail by clicking on the interview key for that case. This will open another interface that will be discussed below.
5. **Search bar** to navigate through the survey cases quickly.



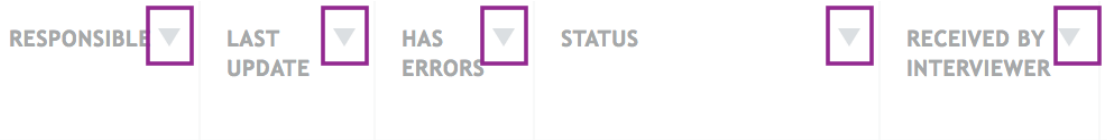
Searching for survey cases

The Interview tab contains a detailed list of survey assignments—those that headquarters has assigned to a supervisor, those that a supervisor has assigned to enumerators, and those that supervisors have approved and sent for headquarters' review. The details can be used to search for particular survey cases. The interview panel in the lower right-hand quadrant of the screen presents a list of interviews and information about them—where the respondent is located (Identifying Questions), which supervisor is responsible (Responsible), when the assignment was last updated (Last Updated), whether the interview contains errors (Has Errors), what status the interview has (Status), whether the interviewer has received the assignment (Received by Interviewer), and in the case of census mode surveys whether the case has been created on the device by the interviewer (Created on Client).



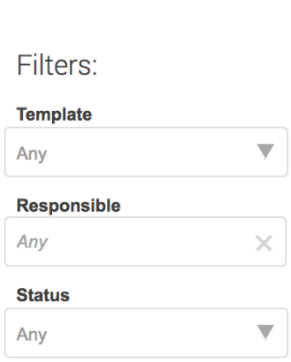
There are three mechanisms for finding a survey:

The **first** is by clicking on the arrow associated with any column of the information in the lower right-hand quadrant. This will sort the list of surveys in ascending (^) or descending (v) order by that chosen column.



The **second** mechanism is through the search field located in the toolbar. Use keywords to narrow down the list of survey cases.

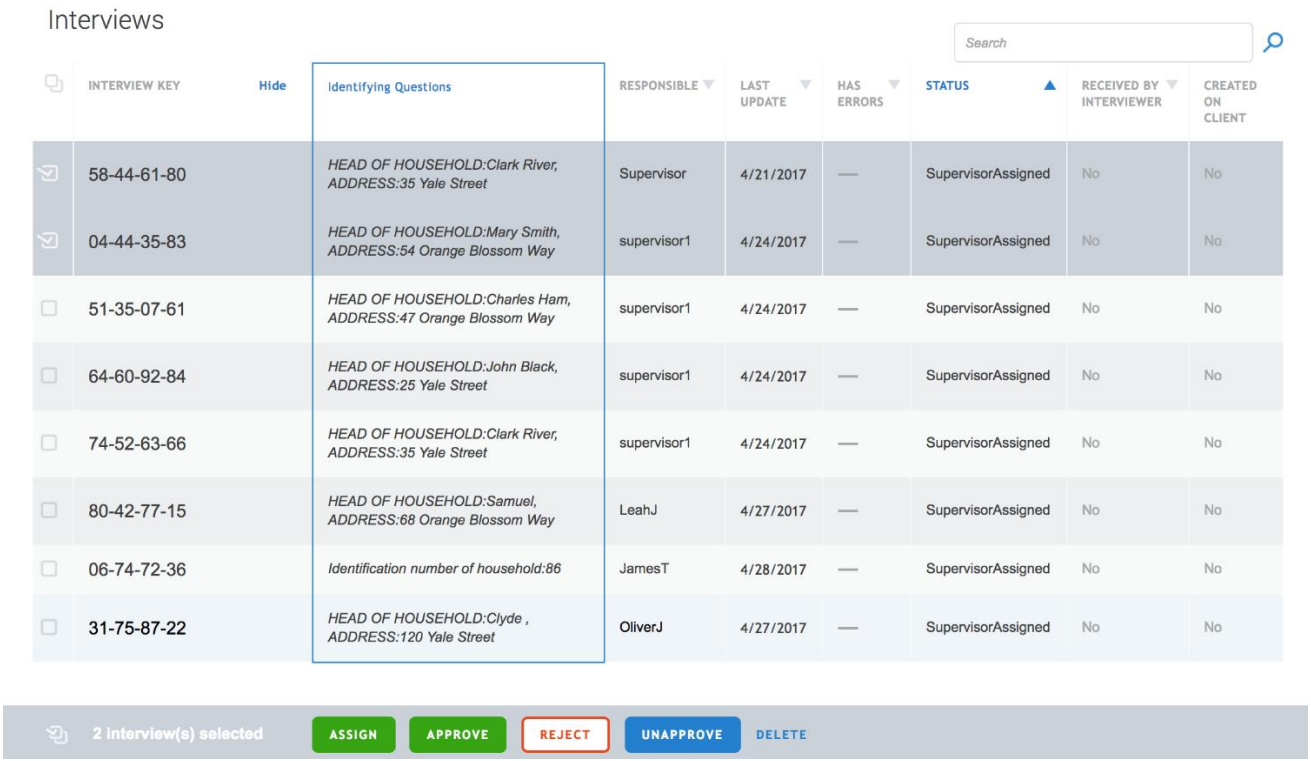
The **third** mechanism is through the filters located on the left-hand side of the screen. To narrow down the list of survey cases, headquarters may filter by the questionnaire (Template), team member responsible (Responsible), and/or the status of the interview (Status).



Reallocate survey cases between teams

Survey cases will need to be reallocated when a new team is created and needs to take over some of the work from existing teams. Cases can be reassigned by a Headquarter user from one team to another.

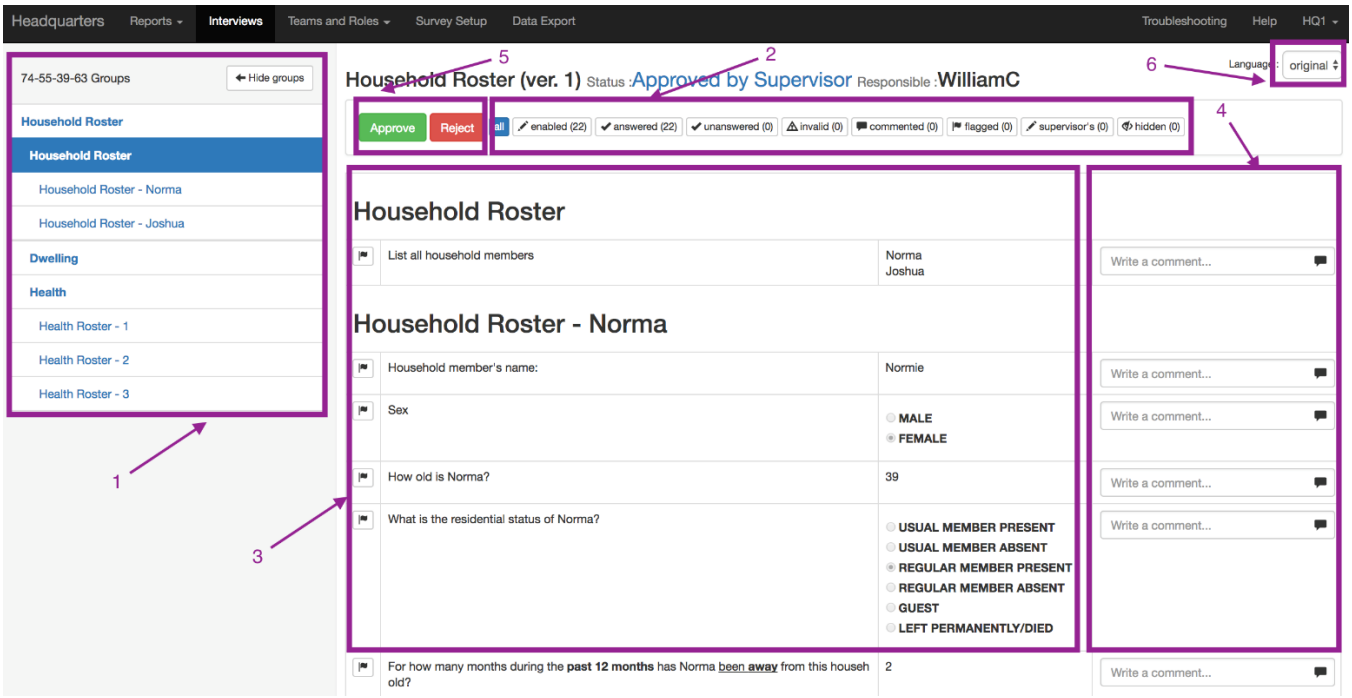
Reassignment is possible when the interview is in any status except the following two: *"ApprovedByHeadquarter"* and *"ApprovedBySupervisor"*. To reassign cases, select the check boxes for the cases you wish to reassign, and click on the green *assign* button.



Review survey cases

The questionnaire review interface consists of five components:

1. **Navigation pane.** Displays a table of contents for the questionnaire, where you can use each entry in this pane to see a particular part of the questionnaire. Clicking on the top entry, the name of the questionnaire, brings the full questionnaire back into focus.
2. **Filters.** These buttons can be used to see questions that have certain attributes—for example, comments or flags.
3. **Questionnaire review pane.** The main part of the screen displays the questions asked in the questionnaire, in the first column, and the answers given, in the second column.
4. **Comment pane.** This pane appears when a question is clicked. This shows the conversation associated with the selected item.
5. **Approve/reject buttons.** These buttons determine whether a questionnaire returns to the supervisor or is integrated into the final data set.
6. **Language dropdown menu.** If the questionnaire is a multilingual questionnaire, the user can toggle between languages using this menu.



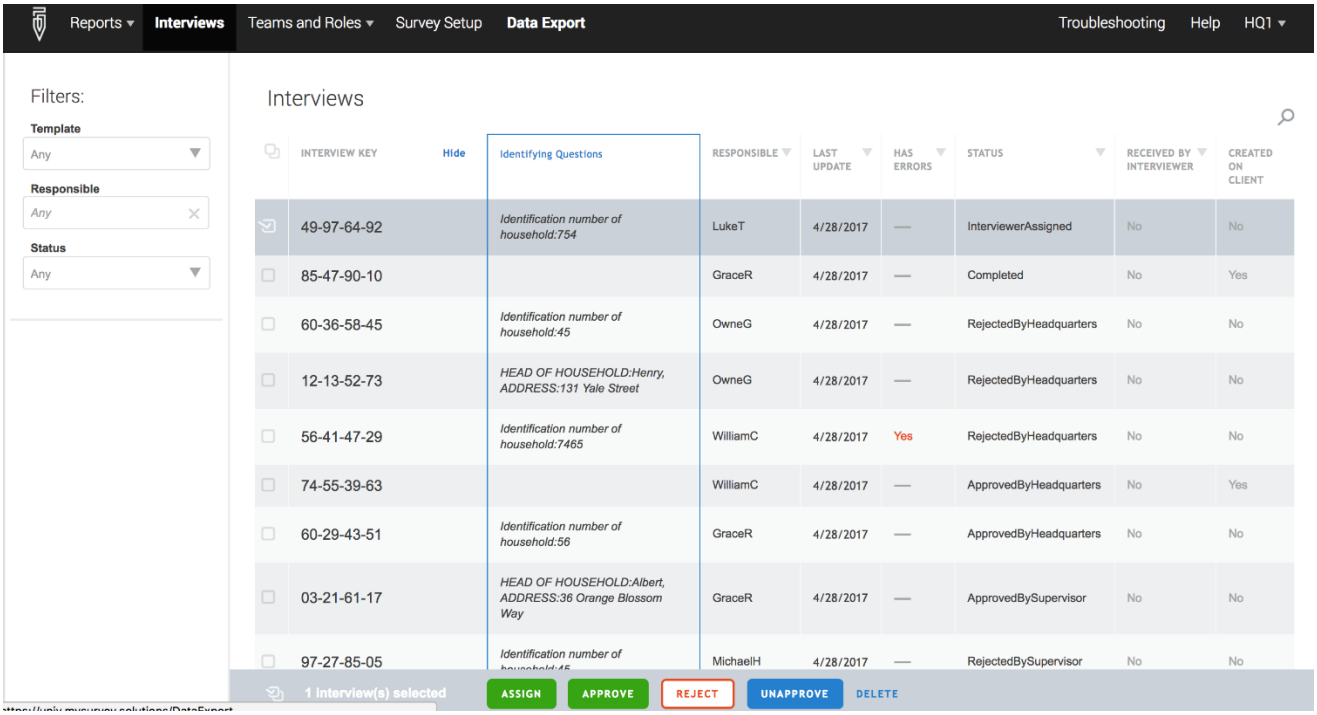
To review the questionnaire for completeness, consistency, and plausibility of answers one should be using the questionnaire review interface. The questionnaire pane, located in the middle of the screen, contains the questions and their answers.

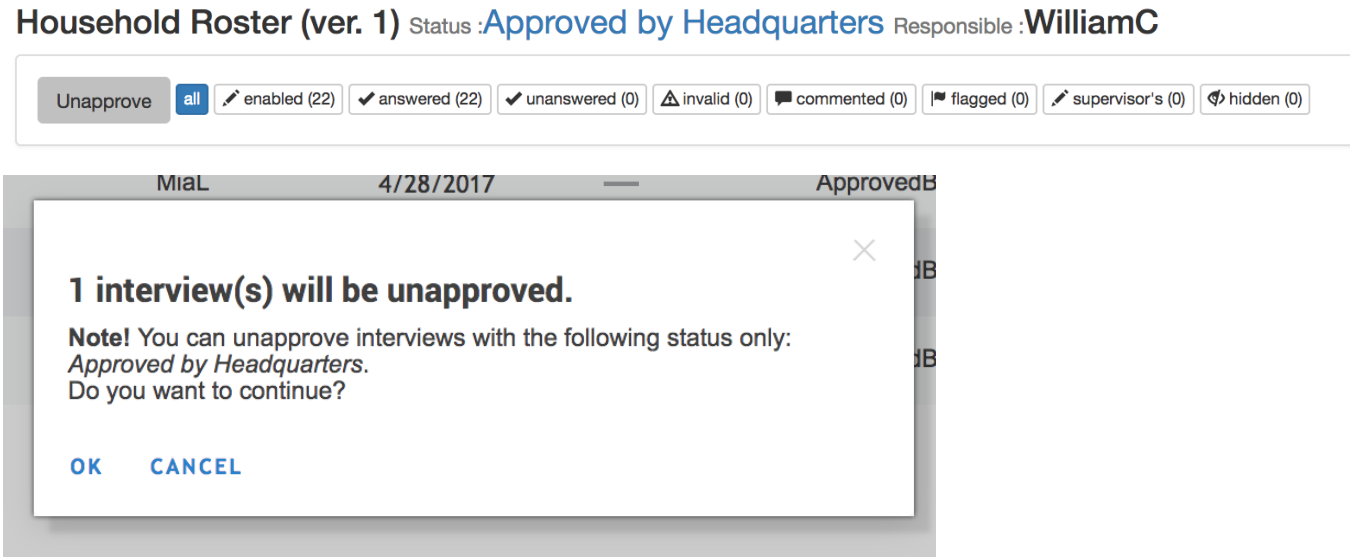
Groups, on the left-hand part of the screen, allow headquarters to see different parts of the questionnaire when any element of the pane is clicked. The top heading corresponds to the complete questionnaire. Bolded headings lead to modules (chapters) within the questionnaire. Indented headings are linked to either groups of questions or rows of rosters.

Filters, arrayed atop the questionnaire pane, facilitate the display of questions that have particular attributes. The "All" filter presents all questions. The "commented" and "flag" filters display questions with comments or flags, respectively. The "answered" questions, as the name suggests, are all those that have an answer. The "invalid" filter displays questions that failed a validation check (e.g., range, consistency with other responses, etc.). The "supervisor's" filter yields special questions that only the supervisor sees and can answer. The "enabled" filter returns questions that are available to be answered based on the skip logic of the questionnaire and answers provided to key questions.

Next, headquarters either approves or rejects the questionnaire by clicking the appropriate button. The headquarters' choice determines whether the assignment is included in the survey database (Approve) or returned to the supervisor for correction (Reject).

To unapprove any interview in the *Approved by Headquarters* status select this interview(s) by checking the checkbox next to it and click the *Unapprove* button. Alternatively, open the interview case and click on the *Unapprove* button on the toolbar. The interview will be reverted to the *Approved by Supervisor* status, and the headquarters user will have a possibility to reject it back to the supervisor if necessary. You can apply this operation to multiple interviews if necessary.





Additionally, the headquarter user, has the authority to approve completed interviews directly, without waiting for the supervisor's decision. This results in faster turnaround of the data circulating in the system in case the supervisor is swamped with work, in which case the headquarters can reduce the workload by pulling the good-quality interviews from the supervisor's stack. There is no change in the user interface, but the button will now permit headquarters approvals of the interviews in status "Completed". Note that you cannot reject an interview to an interviewer directly.

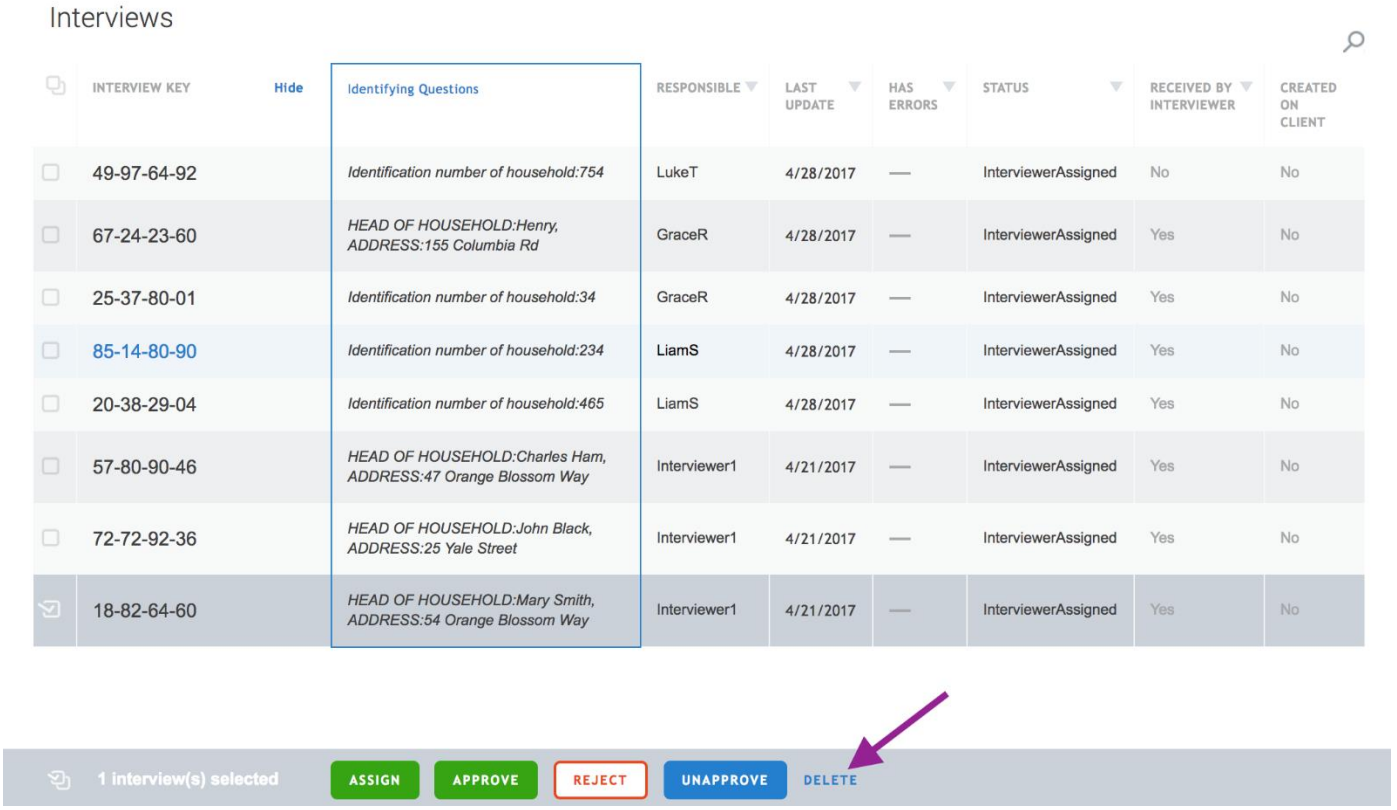
Delete survey cases

Deletion is a dangerous tool reserved for rare occasions—for example, when errors arise in survey sampling or in creating survey cases.

Because deletion is dangerous, only the **administrator/headquarters users** have the authority to delete selected interviews. It is only allowed for interviews with the status Supervisor Assigned or Interviewer Assigned. Those with the former status have been assigned to a supervisor, but not yet assigned by the supervisor to an interviewer. Those with the latter status have been assigned to an interviewer, but have not been completed. In both cases, headquarters should exercise extreme caution. Once deleted, an interview cannot be restored. Hence, administrator should only delete interviews that have been carefully confirmed as errors to be deleted.

To delete an interview, first find an eligible case. The most efficient strategy for doing so may be to filter the questionnaires based on status so that only those with status Supervisor Assigned or Interviewer Assigned appear. (See the section **Searching for survey cases** above for more details.)

Next, select the case(s) to delete by clicking in the check box to the left of the case. Make sure the correct cases have been selected before proceeding. Then, click on the Delete interview button located immediately above the interview panel. This will permanently delete the selected survey cases.


Interviews
INTERVIEW KEY
Hide
Identifying Questions
RESPONSIBLE
LAST UPDATE
HAS ERRORS
STATUS
RECEIVED BY INTERVIEWER
CREATED ON CLIENT
49-97-64-92
Identification number of household:754
LukeT
4/28/2017
InterviewerAssigned
No
No
67-24-23-60
HEAD OF HOUSEHOLD:Henry, ADDRESS:155 Columbia Rd
GraceR
4/28/2017
InterviewerAssigned
Yes
No
25-37-80-01
Identification number of household:34
GraceR
4/28/2017
InterviewerAssigned
Yes
No
85-14-80-90
Identification number of household:234
LiamS
4/28/2017
InterviewerAssigned
Yes
No
20-38-29-04
Identification number of household:465
LiamS
4/28/2017
InterviewerAssigned
Yes
No
57-80-90-46
HEAD OF HOUSEHOLD:Charles Ham, ADDRESS:47 Orange Blossom Way
Interviewer1
4/21/2017
InterviewerAssigned
Yes
No
72-72-92-36
HEAD OF HOUSEHOLD:John Black, ADDRESS:25 Yale Street
Interviewer1
4/21/2017
InterviewerAssigned
Yes
No
18-82-64-60
HEAD OF HOUSEHOLD:Mary Smith, ADDRESS:54 Orange Blossom Way
Interviewer1
4/21/2017
InterviewerAssigned
Yes
No
1 interview(s) selected
ASSIGN
APPROVE
REJECT
UNAPPROVE
DELETE

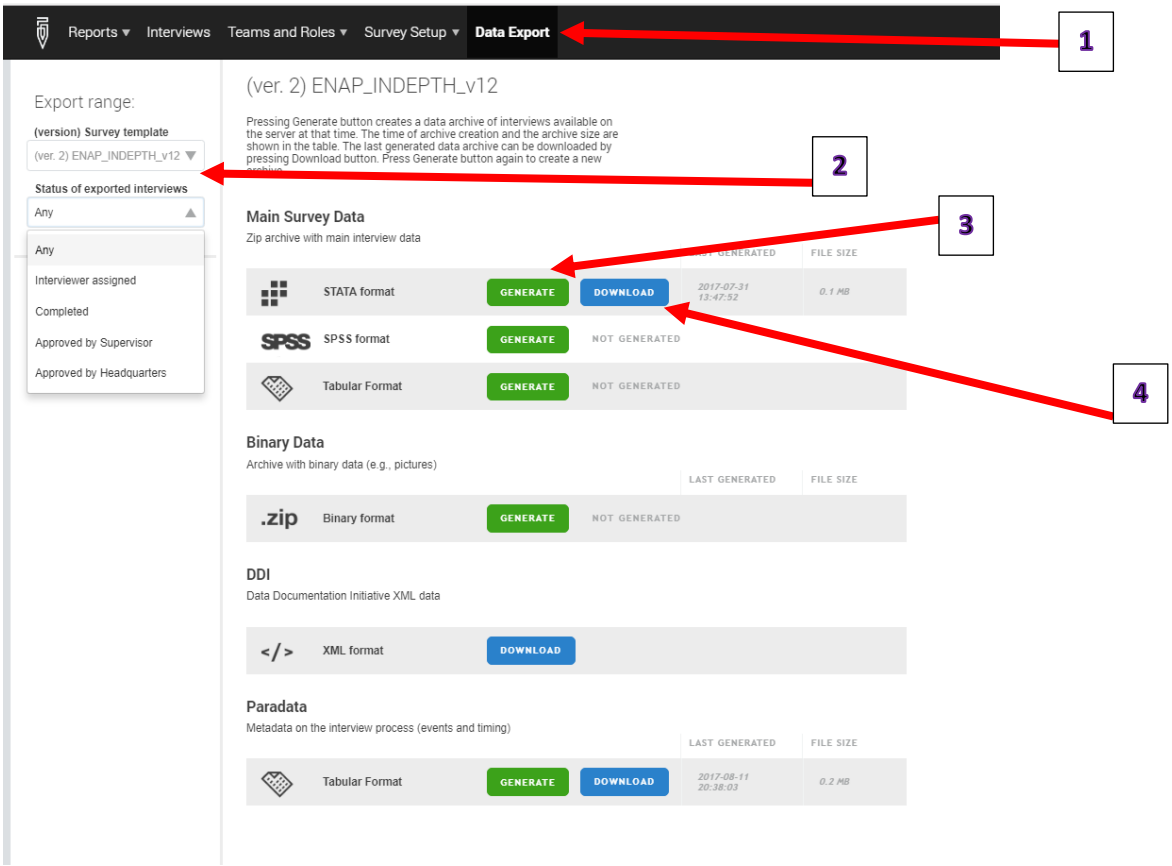
TAB #4: DATA EXPORT

Under the Data Export Tab, you will be able to export the survey datasets (Main survey data, Binary Data, Data Documentation Initiative (DDI) XML data and Paradata).

Before you export the data, you are required to select the version (survey template). You can also select the status of the interviews but this is optional. The different statuses include; Any, Interviewer assigned, Completed, Approved by Supervisor and Approved by Headquarters. These filters are found on the left hand panel.

After specifying the data range using the filter options, you need to generate the dataset to export by clicking on the green “GENERATE” tab for Main survey data, Binary data and Paradata.

After generating the datasets, you should now click on the blue “DOWNLOAD” tab to download



Datasets

Main Survey Data: Survey Solutions allows you to download the main survey data in three formats including; STATA version 14+, SPSS, and Tabular format

Binary data: This dataset contains the archives of the binary data including pictures, sounds, maps etc

DDI: This is the XML data format

Paradata: This is the metadata on the interview process and includes the events and timing data.

Data Export Files

The exported data will be in a .zip file containing one or more files. The number of such files is a function of the structure of the questionnaire, meaning that exporting produces one file per level of observation in the questionnaire, or in other words in addition to one file at the questionnaire level, each roster in the questionnaire will generate another export file. For example, one file for household-level variables, one file for household member-level variables, and one file for asset-level or consumption item-level variables. Additionally, the .zip file will contain interview comments and interview actions files.

Note that the exported data differentiates between missing values caused by the logic of the questionnaire (skipped questions) and missing values due to absence of an answer. Only the values of the logically skipped questions will be have missing values (blanks in tab-delimited files), whereas -999999999 or ##N/A## (for string variables) will be exported whenever an answer to a question was expected, but not recorded for string variables).

The actual files exported will differ slightly based on the file format you have chosen to use: Tab, Stata, SPSS and DDI.

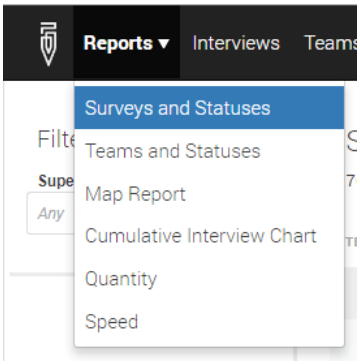
TAB exports tab-delimited data in .tab files. Each file is accompanied by a Stata .do file that reads the .tab file into memory, and applies variable and value labels. **STATA** exports DTA files directly. If .do file for applying variable and value labels desired, they can be found accompanying the tab-delimited data. Version 14, which supports Unicode standard for string data. **SPSS** exports SAV file directly. **DDI**: exports the list of data files, variables, their types, labels, question texts, interviewer instructions, etc.

Questionnaire and Data Export Structure

Data collected using the questionnaire with a roster (same set of questions asked about different events) is exported in the following number of files: one file at the questionnaire/household level of observation, another file at the household members/parent level, and two system generated service files on interview actions and comments. The detailed anatomy of the export file for questionnaire data can be found at: http://support.mysurvey.solutions/customer/portal/articles/2579806-questionnaire-data-export-file-anatomy. For system generated export file anatomy please refer to: http://support.mysurvey.solutions/customer/portal/articles/2579821-system-generated-export-file-anatomy.

TAB #5 REPORTS

This Tab appears on the left of the main menu, it is where all reports and summaries about the survey implementation and status are created. The reports include: Surveys and Statuses, Teams and Statuses, Map report, Cumulative Interview Chart, Quantity and Speed.



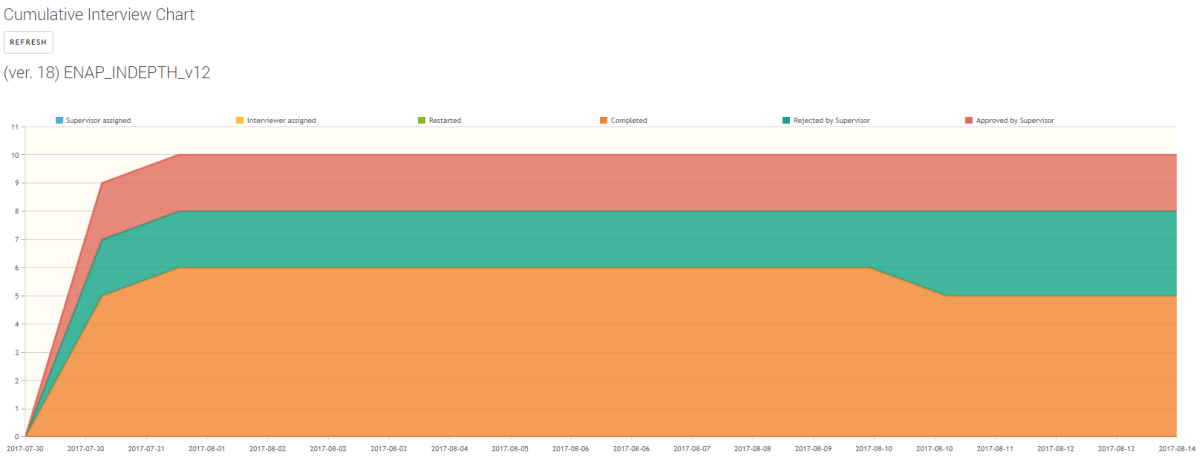
See below a brief description about the different reports (detailed description can be found at: http://support.mysurvey.solutions/customer/portal/articles/2481544-reports?b_id=12728).

Surveys and Statuses: This shows the overall questionnaire template summary. It includes the version number of the template, the questionnaire template name, the number of supervisors assigned, number of interviewers assigned, number of completed interviews per template, number of rejected interviews, number of approved and interviews.

Teams and Statuses: This reports details the summary of interviews completed, rejected, approved and total interviews by supervisors in the survey.

Map Report: This report shows the map reports for the different questionnaire templates using the Google maps and GPS overlay.

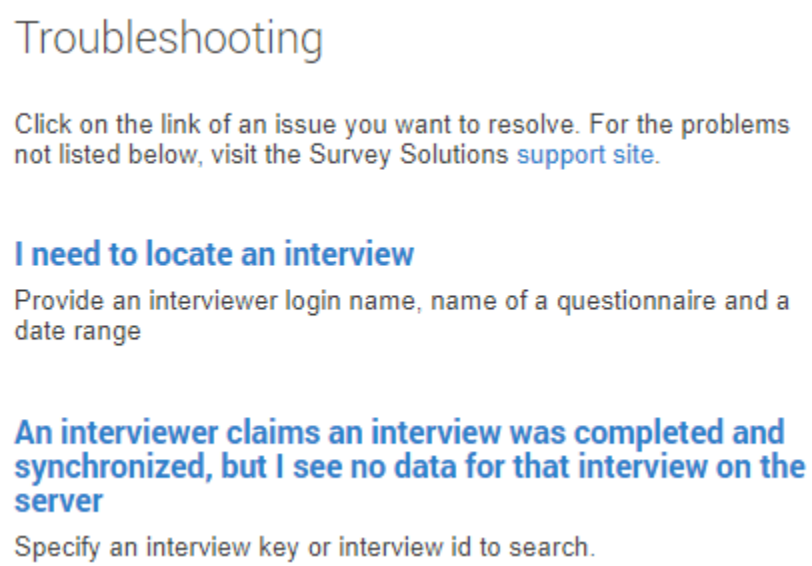
Cumulative Interview Chart: This report shows the cumulative interview status (Supervisor assigned, Interview assigned, Restarted, Completed, rejected by supervisor and Approved by Supervisor) by dates.



Quantity: This details the number of completed interviews, interview transactions by HQ, approved by HQ and interview transactions by supervisor

TAB #6: TROUBLESHOOTING

This tab is used for troubleshooting for problems or issues that the survey manager may want to resolve. Some preempted issues are shown under the troubleshooting tab namely; locate the interview and lost interviews that are claimed to have been submitted. For problems that are not anticipated, the survey manager is required to visit the Survey Solutions support site.



TAB #7: HELP

When this tab is click on, it will open the survey solutions help files in another tab on your internet browser.

A7: EN-INDEPTH Survey Implementation Guidelines

INTRODUCTION

This document is an instruction guide for setting up the INDEPTH-ENAP Survey Solutions system to support the data collection for the surveys at the participating Health and Demographic Surveillance Sites (HDSS).

OVERVIEW OF THE SURVEY SOLUTIONS APP

In order to have a fully functional system, you are advised to follow this guide step by step. Before we go into the details of how to set up a functional Survey Solutions system, you will need to ensure that you have some minimum software and hardware requirements place.

SYSTEM REQUIREMENTS

A. SERVER

For this study, three options are available for the server setup and support.

1. The World Bank cloud server (This is freely available on request).
2. The London School of Hygiene & Tropical Medicine will provide a server to support the HDSS during the data collection and management. This option is available if the HDSS does not have a server or if it has one that does not meet the minimum requirements for the survey solutions software to run.
3. The HDSS can also use its own existing server if it meets the minimum requirements

See below the minimum requirements for own the server and setup.

1. PRE-INSTALLATION AND READINESS CHECK:

1.1. Please check that your server meets these physical prerequisites before opting to use your own server:

- **CPU and RAM**: 4 cores CPU, 8GB RAM and 50GB free space as minimum requirement. To achieve maximum performance, ensure you have better hardware. The World Bank technical support team recommends quadrupling the basic minimum requirements specified above (for RAM and drive), or having at least Intel Xeon E5-2670 v2 (Ivy Bridge) Processors, 8GB RAM, 50-100 GB SSD Drive.
- You should ensure that your server is dedicated to the project and not competing with other projects for server resources.
- You should ensure that you have in place a stable Internet access for 24 hours and 7 days a week that is fast and reliable for immediate synchronization and backup.
- You should have a Server back-up is in place. This can be either a physical external or virtual drive. You should also have a schedule for the data back-up.

1.2. Please check that SERVER ENVIRONMENT meets the Server Environment prerequisites:

- Properly cooled
- Physically and digitally secured

- Reliably powered 24/7
- Technical on-site support available

1.3 Please check that the SERVER SOFTWARE meet the Software prerequisites (installed and configured prior to Survey Solution installation):

- Windows Server 2012R2 or above
- Microsoft IIS (included in OS)
- Microsoft .Net Framework 4.6.1 or above
- PostgreSQL version 9.4 minimal, or above

1.4 Antivirus/firewall/integrated security package software installed. Firewalls, anti-virus, and integrated security packages need to be configured to permit Survey Solutions server component to communicate with other devices exchanging data with it. Below are a couple of suggestions:

- Instructions for Windows Server 2012 Firewall
- Avast Internet Security
- Comodo Internet Security
- Kaspersky Internet Security
- McAfee Internet Security
- Norton Internet Security
- ZoneAlarm Pro Firewall

2 INSTALLATION:

2.1 Data base installation (**step 1**)

Before running 5.21 **SurveySolutions.exe** installation setup, please make sure that you have installed and running an instance of PostgreSQL, and other software specified above (1.3 and 1.4).

The PostgreSQL can be downloaded here: https://www.bigsql.org/postgresql/installers.jsp/ . Both versions 9.6.3 and 9.5.7 are compatible with SS. Please refrain from installing v.10, as it is in the beta phase.

During the process, you will set the port on which server will be running as well as a password for built-in user (postgres by default).

2.2 Application installation

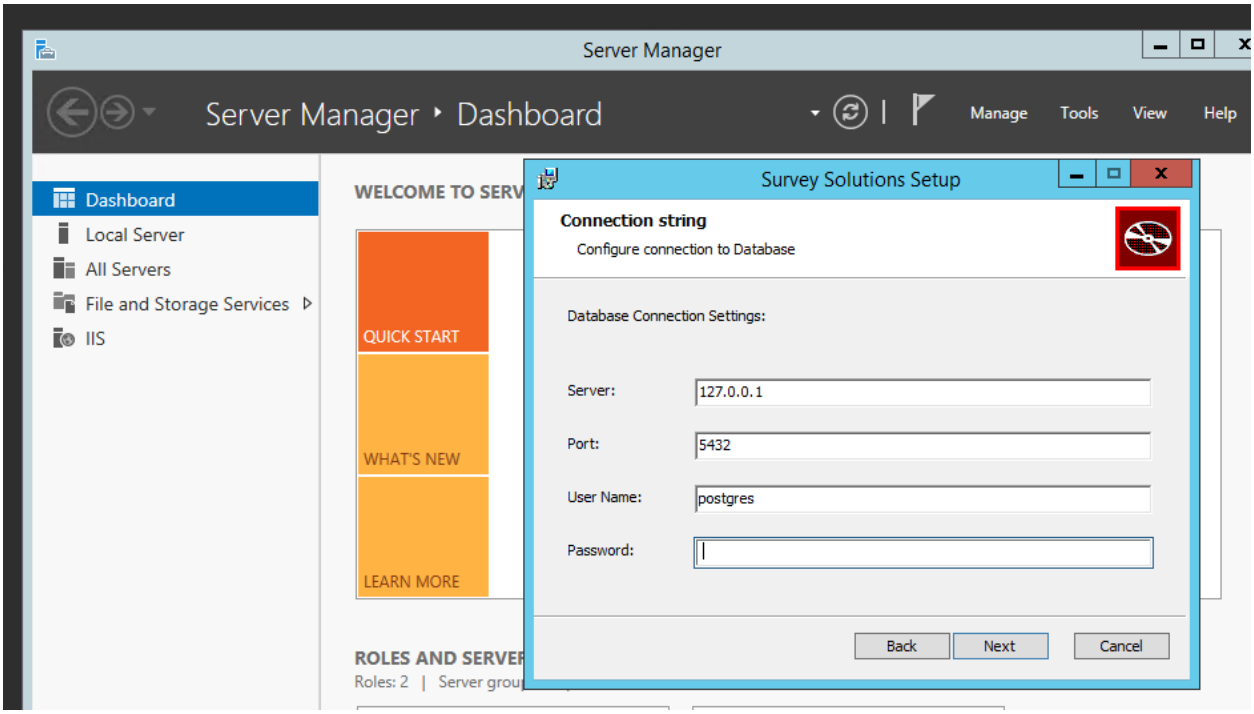
There are two options: manual local server installation (step 2a) or assisted installation by World Bank Survey Solution technical team (step 2b). Both options are described below:

Step 2a – manual local server installation

Survey Solution installation file can be found at: https://download.mysurvey.solutions/

Please make sure that you have downloaded (or upgraded) to the latest released version of the Survey Solution, which is **5.21.0**. You may experience some technical issues with the older version **5.20.0**, so please update to the most recent one (**5.21.0**).

During installation it will ask you to specify two parameters: **Installation location** – choose any or keep the default (C:\Program Files) and **PostgreSQL connection parameters** (default values will be already there) – here specify the required port and password from **step 1** (Data base installation).



Note: FOR ILLUSTRATIVE PURPOSES.

After you complete the Survey Solution installation on your server, the finish installation page will automatically open in your browser where you will be able to create administrator user for Survey Solutions Headquarters application. The link to the site will be automatically added to your start menu. If this does not happen or you want to access it directly by typing the address in the browser's address bar, you can do so by typing http://localhost:9700/ in the browser whilst in the server environment.

Once the installation is complete, please forward the following details regarding your server to Vladimir (vladimir.gordeev@lshtm.ac.uk): the username login for Survey Solutions Headquarter that is used by the Headquarter admin that you just created, associated e-mail, URL and real IP address of your server installation. For example, “indepthadm”, user@site.com, 255.22.54.233, andhttp://surveysolutions.lshtm.ac.uk:9700/. This information will be then forwarded to the World Bank for your server to be whitelisted for security purposes.

OR

Step 2b – assisted installation by World Bank Survey Solution technical team

Given that you can provide remote access (RDP) to your local/virtual server and admin rights to the World Bank Survey Solutions technical team, you can send a request to install Survey Solution software remotely by sending a request to support@musurvey.solutions

HARD WARE REQUIREMENTS

LAPTOPS/DESKTOP FOR SUPERVISORS

Each site should have the PCs (laptop(s) or Desktop(s) designated for the supervisor(s) whose role is to oversee the work of the interviewers. NOTE: It is also possible for the supervisor to use a Tablet, but one might find it easier to manage assignment tasks using the PC, given the size of the screen.

The Laptops or computers (or tablet) should be able to have an internet connection either using WLAN or WIFI or 3/4G dongle or modem. This is essential requirement to allow supervisors being able to review the submitted interviews by the interviewers. The PC should be running Windows 7/ Windows 8 Operating System (OS) or higher and have a minimum of 4GB RAM and 500GB HDD.

If the supervisor is working in the field, the connectivity in the field should be arranged using 3G or 4G connection (built-in, USB, external wireless modem, or dongle).

TABLETS

The Survey Solutions APP is designed to work on Android tablets. For maximum usability of the APP, the interface is optimized to work on tablets with screen size larger than 7-inches. The APP will also work with the smaller tablets (5-inches or smaller) but with limited visibility due to the small fonts.

Additional requirements

1. Tablets should ideally be operating **OS Android 5.0 or above.** Devices running OS Android 4.1-4.4 are incompatible with the TLS v. 1.2 protocol, a security protocol that is used on most of the servers. This incompatibility does not have a straightforward fix, and tablets running OS Android 4.1-4.4 are not able to synchronise the data with the server.

- 1GB of RAM and 8GB of flash memory storage are minimal. Better technical characteristics will improve responsiveness of the programme
- At minimum 1GB of available space must be available for the Survey Solutions' use.
- The tablets should be simcard enabled to allow for 3G or 4G connection for connectivity while in the field or should able to connect to a WIFI connection for connectivity at the office.
- The tablet should have cover for storage and protection from rain and sudden drops
- The tablet should have an extra external battery or recharger
- It is recommended to install a tablet lock APP to limit the use of the tablets for other purposes by the interviewers other than data collection, as it can reduce the battery performance of the tablet up to 50%.

CREATING ACCOUNTS FOR YOUR TEAMS

After completion of setting up the server, and checking the requirements for both the tablets and the laptop or desktops, you are now ready to set up the Designer, the Headquarters, Supervisors and Interviewer user accounts

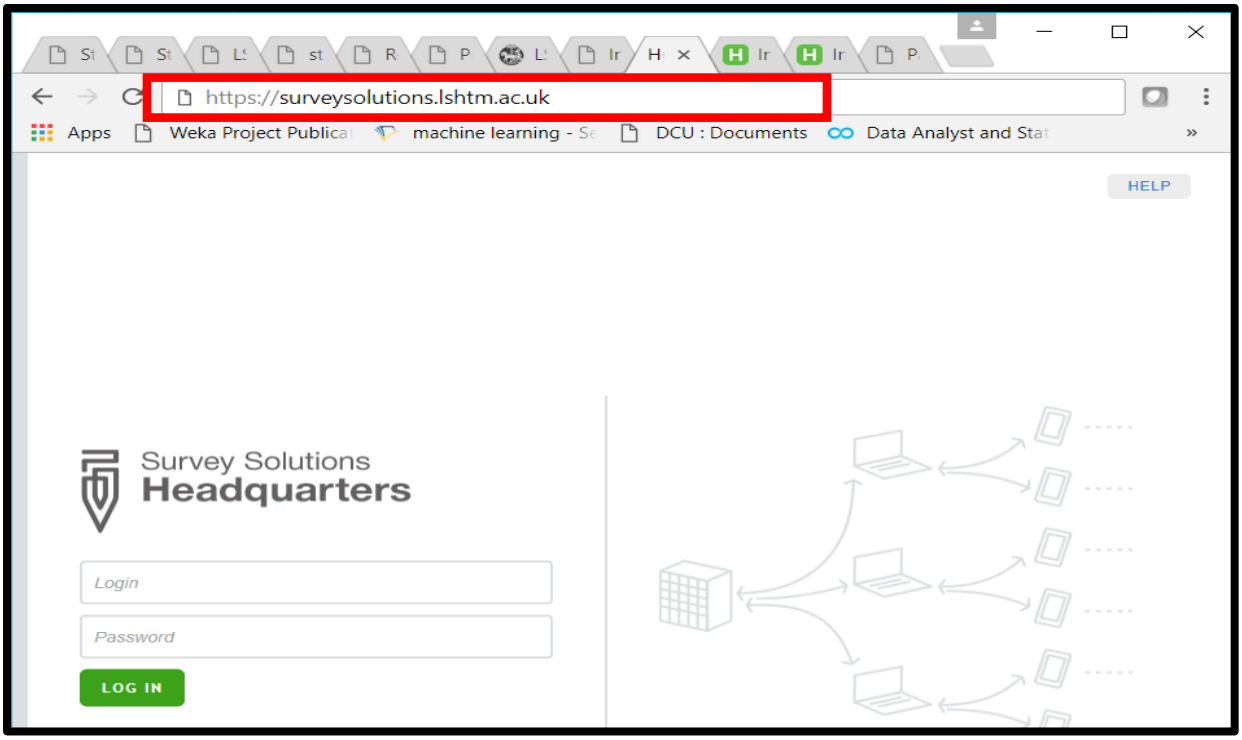
STEP 1

HEADQUARTERS

Option 1 – If you are using the LSHTM server

Vladimir (vladimir.gordeev@lshtm.ac.uk) who has access to the LSHTM server will create a Headquarters account for your HDSS site from his administrators account.

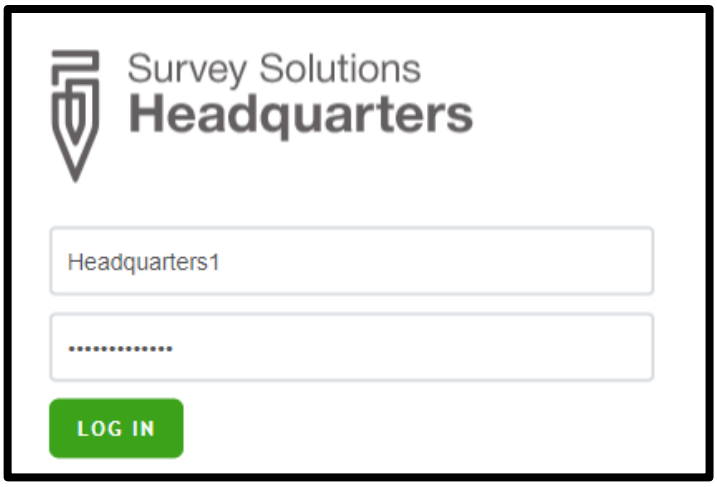
To log into your headquarters, you will need to type the virtual server address (for example, https://surveysolutions.lshtm.ac.uk), and log in using the headquarters username and password that will be provided to you as it is shown below:



Note: THIS OPTION IS ONLY AVAILABLE TO YOU IF YOU ARE USING THE LSHTM SERVER. THE ADDRESS IS FOR ILLUSTRATIVE PURPOSES.

Option 2 – If you have set up your own server

Log into your server using the headquarters username and password you created at the end of the installation. Each HDSS will utilize its own headquarters and will have its own username and password different from the one shown below.

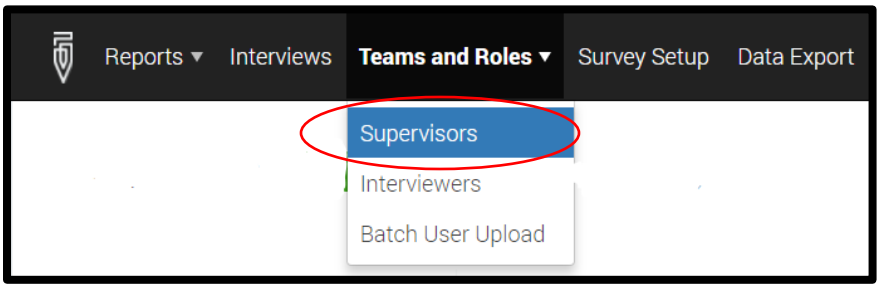


Once you have successfully logged into your Headquarters account from using either options provided above, you are now ready to setup accounts for other users, including the supervisors and interviewers.

STEP 2

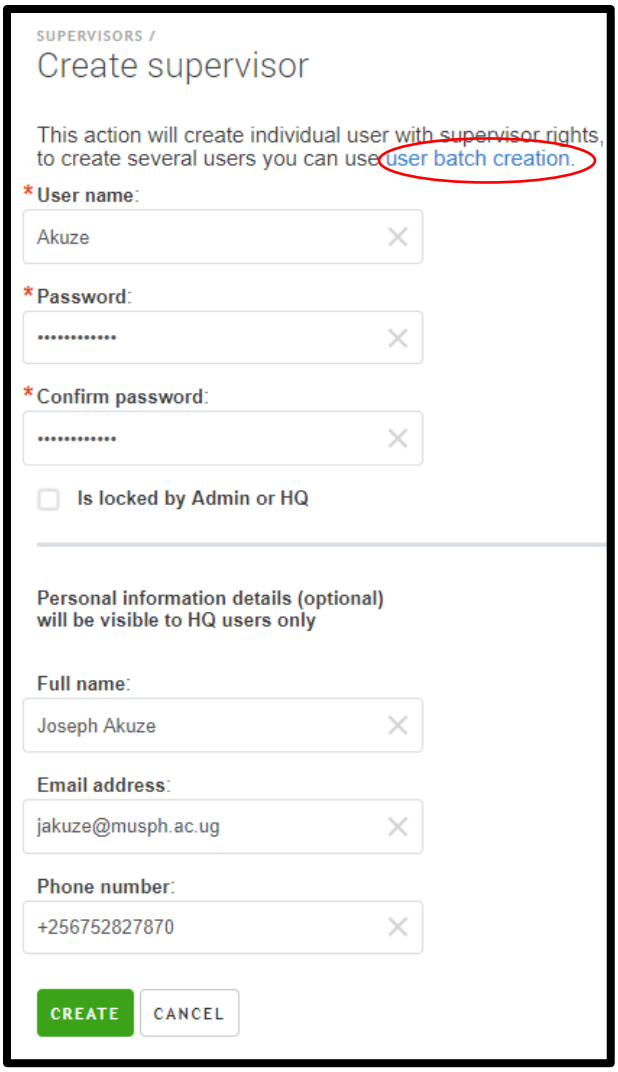
Creating accounts for supervisors

Ensure that you add supervisors for your HDSS. The supervisors responsible for assigning tasks to the interviewers and reviewing the data collectors completed questionnaires. To add a supervisor, click on the *Teams and Roles*

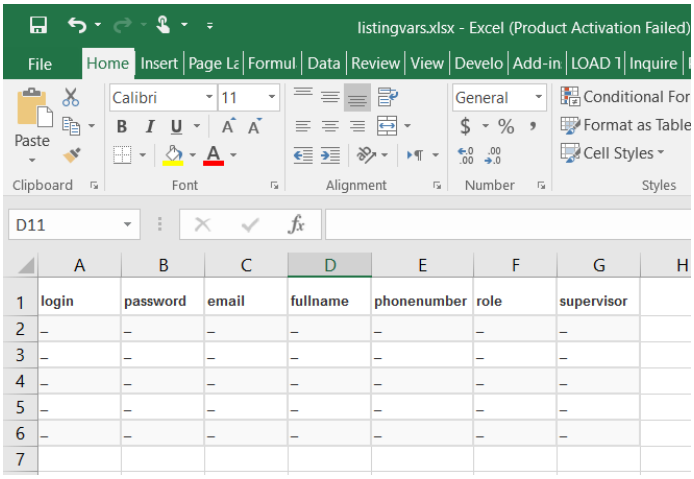


Then, complete the form displayed and click on ***CREATE*** to add the supervisor in the system.

Note: This option only allows you to create one supervisor account at a time. It is possible to create several supervisors using the ***USER BATCH CREATION***



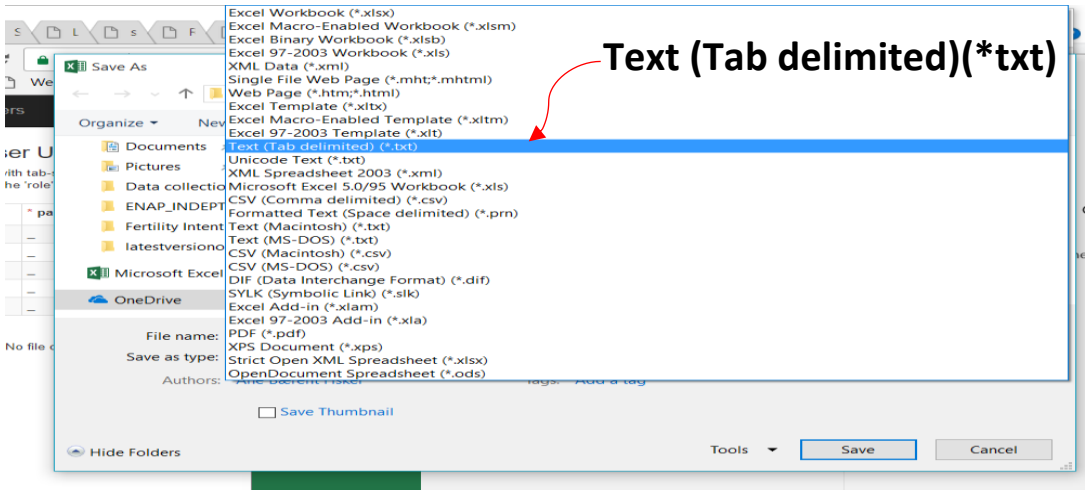
In order to upload more than one supervisor using the user batch upload. You will need to open a new MS excel workbook, add the following fields; (***login, password, email, fullname, phonenumber, role, supervisor***) as is illustrated below.



Once you have added the fields with the same exact wording, you should now save the workbook on your computer save a “tab-delimited” file. Or “.txt”

Ensure that you save it in a location that you can find easily on your computer.

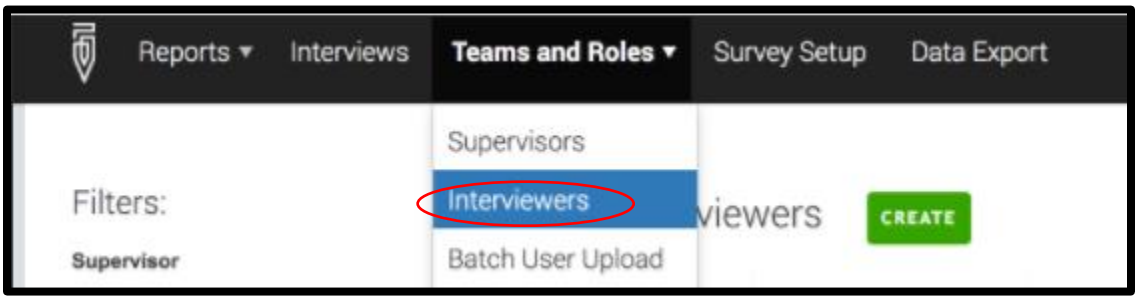
Note: Only ‘Supervisor’ OR ‘Interviewer’ are allowed in the ‘role’ column.

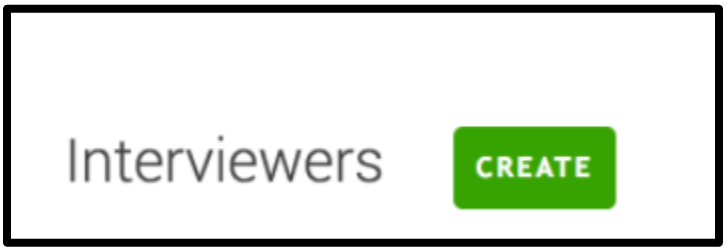


STEP 3

Adding interviewers

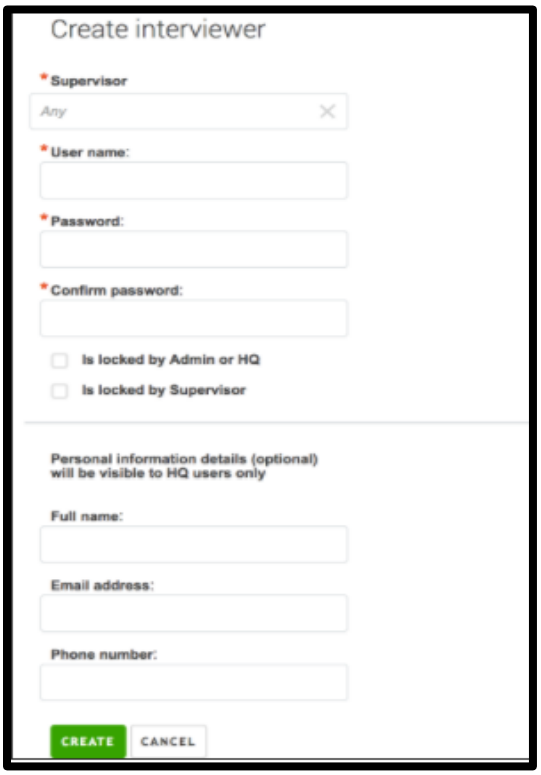
This process is similar to adding supervisors. You need to go to the Teams and Roles tab and then select Interviewers, then create interviewers, then complete the form displayed.





This approach allows you to upload only one Interviewer at a time. To upload many Interviewers, you need to use the Batch option which is similar to the one for Supervisors as is shown in STEP 2.

The batch upload saves you a lot of time.



NOTES

Up to 10,000 user accounts can be created from a single tab-delimited file, which must include the login, password, and role of the user, and may include additional attributes, like email or phone.

Batch creation of user accounts comes handy when:

- the same accounts need to be replicated on a different server;
- migrating from a different CAPI system with established accounts;
- the survey is so huge that manual creation of accounts via a web interface would take a significant amount of time.

Creating user accounts in batch mode is available to both the administrator and the headquarter users. For new supervisor accounts the role must be specified as Supervisor (case-sensitive). For interviewer accounts the role must be specified as Interviewer (case-sensitive), and in addition the login of the supervisor must be specified, which determines the team where the interviewer account will be added. Interviewers may be added to existing supervisors, or to new supervisors mentioned anywhere in the users list.

When creating new users with a batch upload feature, select the Batch user upload menu item of the Teams and Roles menu.

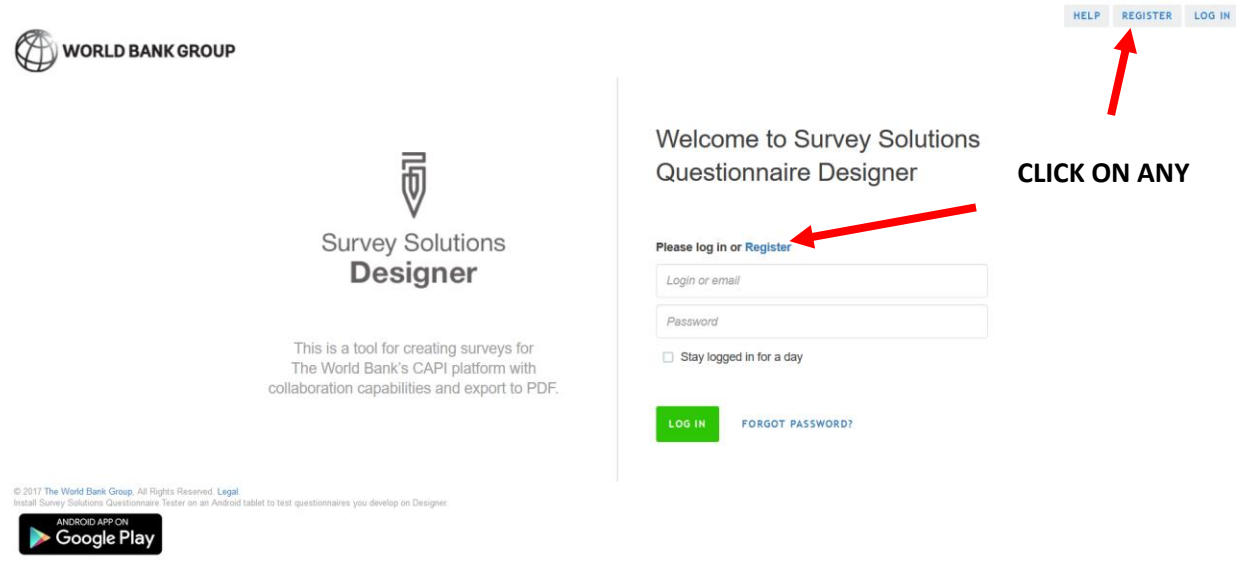
DESIGNER ACCOUNT

The HDSS data manager or the survey manager at the HDSS site will need to also create a designer account from which he/she will be able to access the latest version of questionnaire. The purpose of having a designer account is to be able to import the questionnaire template into the Headquarters so that is can be assigned to supervisors and interviewers during the data collection. The following steps illustrate how to create a designer account.

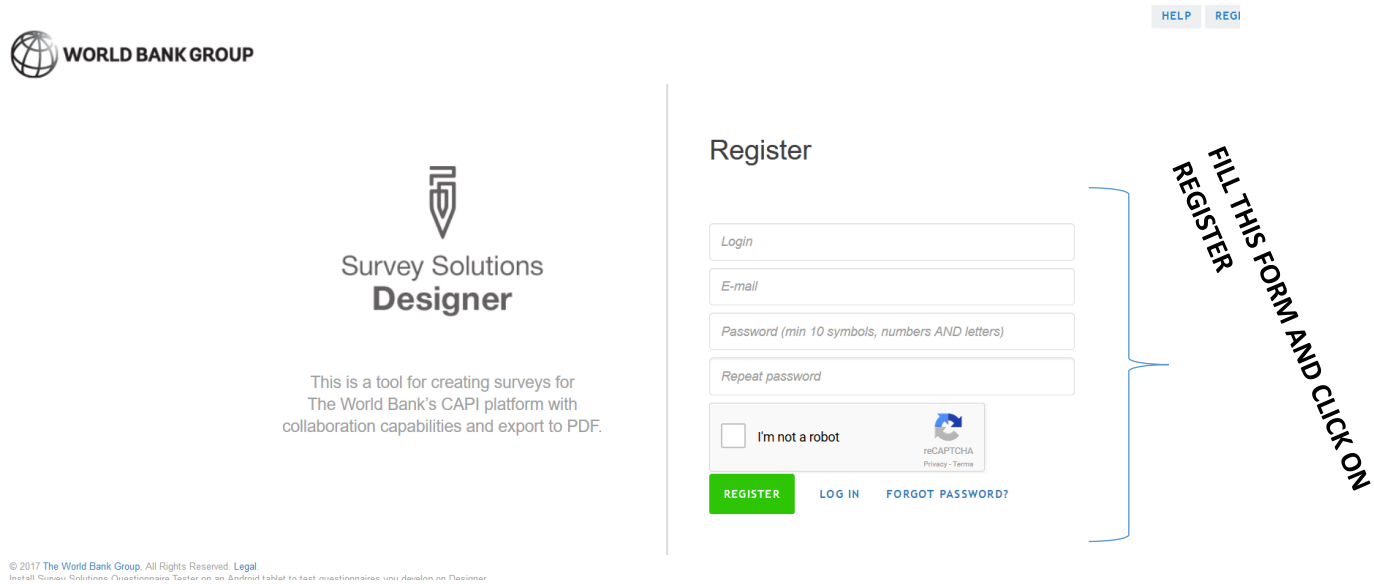
STEP 1

GO TO: https://solutions.worldbank.org/

STEP 2: Click on Register

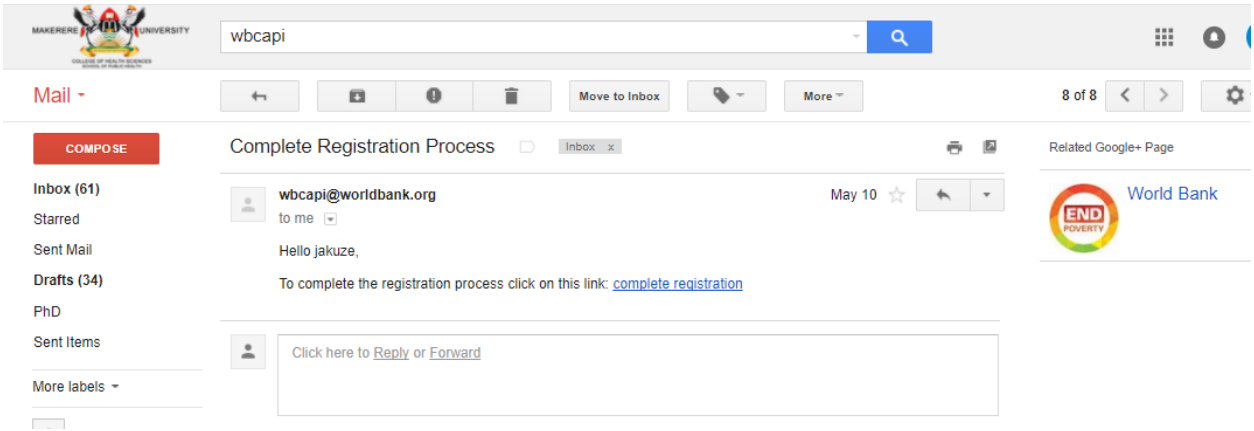


STEP 3: Fill out the form and click on register. Then wait for an email notification



STEP 4:

Complete the registration by clicking on the link "*Complete registration" in the email received.*



STEP 4:

When you click on the link with complete registration in step 3 above, a new browser will open so that you can log into the DESIGNER using the USERNAME and PASSWORD that you have created. With the designer account you are able to design your own questionnaire template.

SURVEY SETUP

Uploading questionnaire template

Now that the survey manager at the HDSS has created accounts including the headquarters, supervisors and interviewers. The final survey questionnaire will be uploaded by the overall survey administrator. It is only possible for original author of the questionnaire template to share it.

Creating assignments

Once the questionnaire template has been uploaded into the survey headquarters questionnaire setup system. It is now possible to start distributing assignments to the different interviewers.

There are two options available for creating assignments. One at a time and many assignments at a time.

OPTION ONE: Creating one assignment at a time

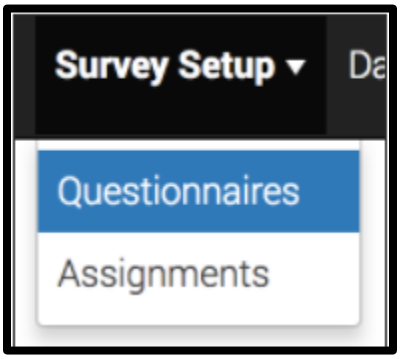
With this option, you can only upload only one assignment at a time. This is not the best option for issuing assignments as you will have to add assignments one at a time.

STEP 1

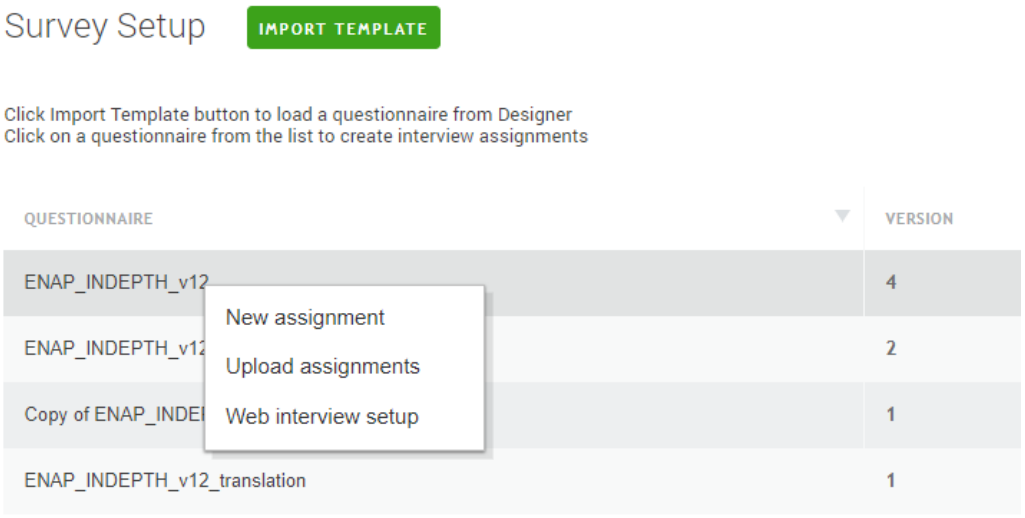
The HDSS data manager should log into the server using the headquarters username and password as was shown in STEP 1 of creating accounts for teams.

STEP 2

You should select the ***Survey Setup*** tab on the main menu, and click on **Questionnaires.**



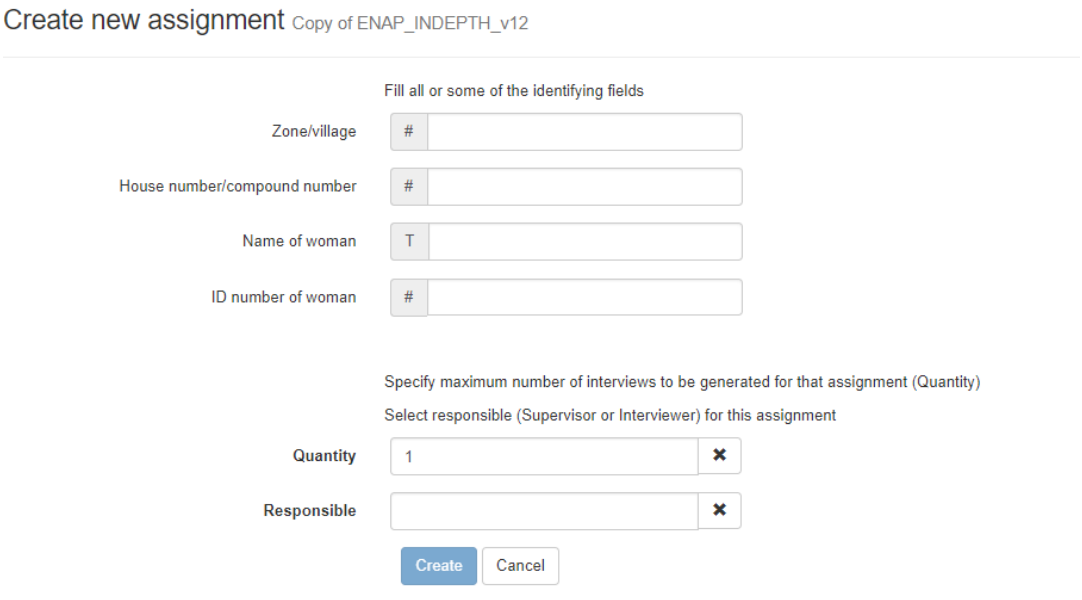
In order to create a new assignment, we need to click on the questionnaire template that you want to create an assignment for and then click on the *New Assignment* from the pop-up menu that you want to send out.



STEP 3

After clicking on the new assignment option, a new window will open up. This is where you should fill in the identifier variables for the interviewer to be able to identify the respondent correctly.

After specifying the identifier information, the quantity and the Interviewer responsible, you can now complete the assignment creation by clicking on the create button.



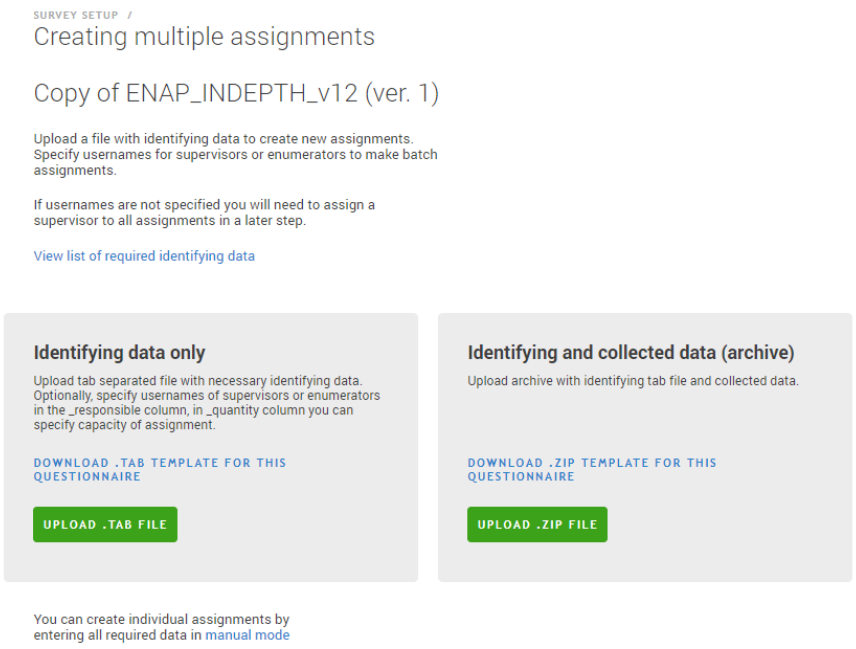
The only downside of this option is that you to repeat this process over and over again in case you have many new assignments to the interviewers.

OPTION TWO: creating many assignments at ago

Before choosing this option you need to first create a tab-delimited text file with the unique identifiers for the woman who has had a pregnancy in the last five years within the HDSS setting. Examples of these identifiers may include, study ID, woman's ID, name, age, residence etc. These set of variables need to be compiled by the HDSS statistician or data manager.

Repeat STEP 1 and 2 for option one

When using this option, of creating assignments, you will need to repeat the same procedures for steps one and two above for creating one assignment at time. But this time you should select upload multiple assignments.



STEP 3

Upload the prepared listing or identifier data in text-delimited version. Select identifying data only if you have the prepared file on your PC. This is the only option where you can specify your interviewers and supervisors a priori. If your upload is successful, the system will tell you that it is successful.

At this point, you will be able to see the details of the assignments including the date created, modified, ID number and the person responsible for that assignment.

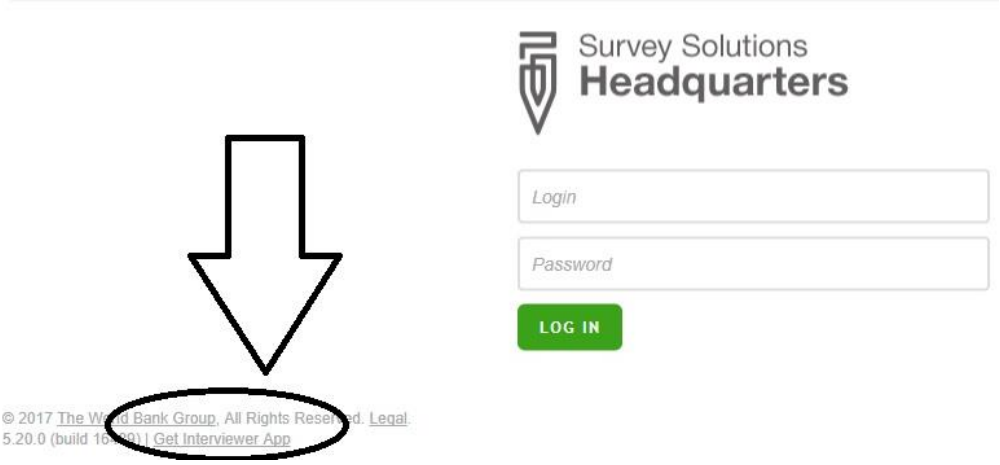
Within the headquarters you can review the completed interviews and can accept, reassign, assign, archive data and download data into usable formats like in STATA 14, SPSS, CSV among others.

Download and Install the Interviewer Application on the tablet:

IMPORTANT! You should not use the Tester app from Android PLAY store for data collection. Any data entered in tester will not be stores. You need to download an INTERVIEWER app.

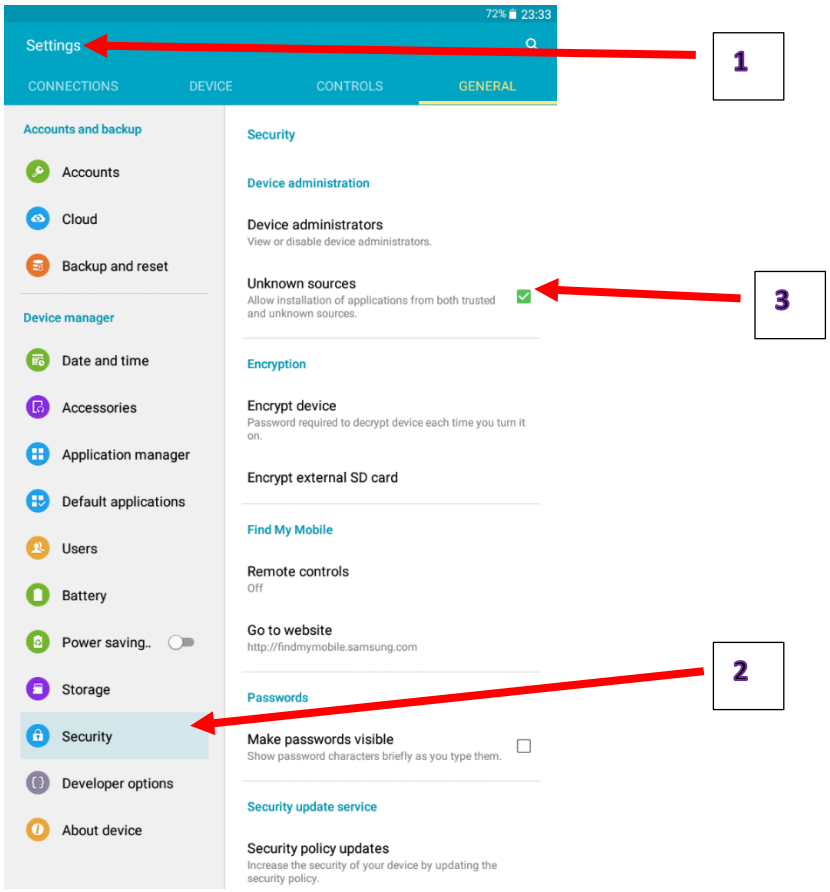
It is important that you download the Interviewer app from your own Survey Solutions server and do not download it from the generic page of the World Bank. To do so, you will need to go to the login page of your designated headquater. For example, if your virtual server is with LSHTM, it will be the same address as the virtual server - https://surveysolutions.lshtm.ac.uk. If you use your own server, you will need put in the address of your virtual server in the browser.

Once you are at your home headquarter log in page, at the bottom left corner of the page you can see the button which will allow you to download the INTERVIEWER app (shown below). The APP setup can be downloaded on your PC or directly to the tablet. To save up your internet traffic and data charges, we recommend that you download the setup .apk file once and copy it to all other devices. This mode of operation will save you a lot of time and money.

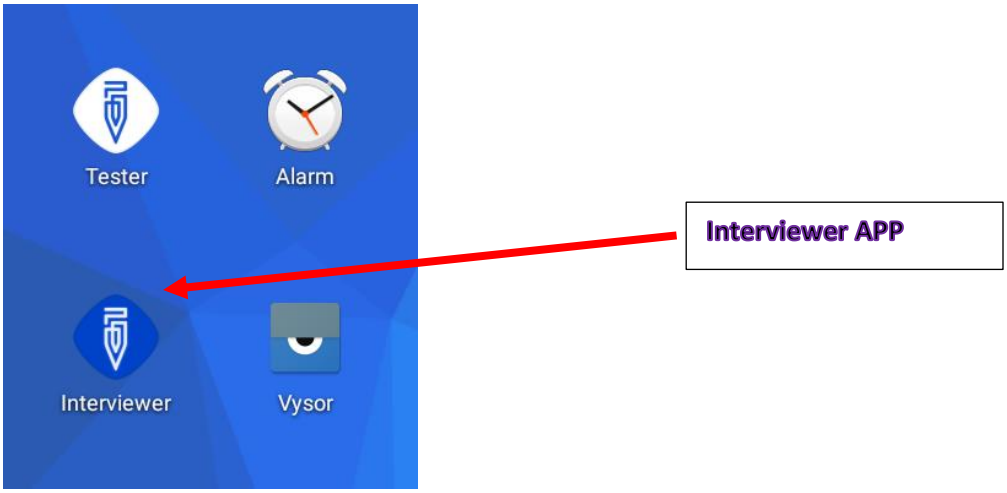


Further detailed instructions on further app installation:
http://support.mysurvey.solutions/customer/en/portal/articles/2475216-download-and-install-the-interviewer-application

Since the interviewer setup is not part of the Google APP store certified applications, you will need to accept installation of applications from unknown sources. This can be easily found on your device by following this path **SETTINGS >> SECURITY >> UNKNOWN SOURCES.** When inside the security settings, you need to find the unknown sources checkbox. This will allow you to run the installation of the Survey Solutions APP on your devices.



When the installation is complete, the interviewer APP which is blue in colour will appear on the home screen of your devices.



Logging into tablet for the first time.

When the installation is complete, the interviewer APP which is blue in colour will appear on the home screen of your devices. This should allow you to login into one of the Interviewer accounts that you created as is shown on pages 5-7 of this document. You will need to enter the link to the server which you selected to use – the address of your virtual server. For example if you are using the LSHTM server, then you can use the following link: https://surveysolutions.lshtm.ac.uk .

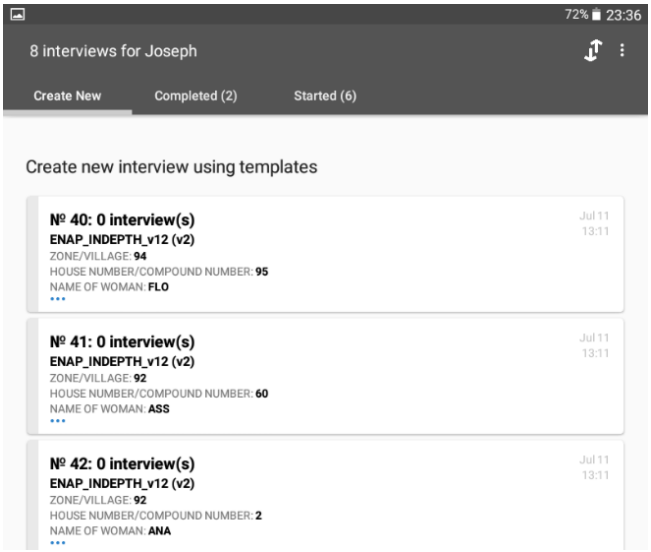
This should then be followed by the Interviewers' assigned user name and password that you created for them during the interviewer account setup.

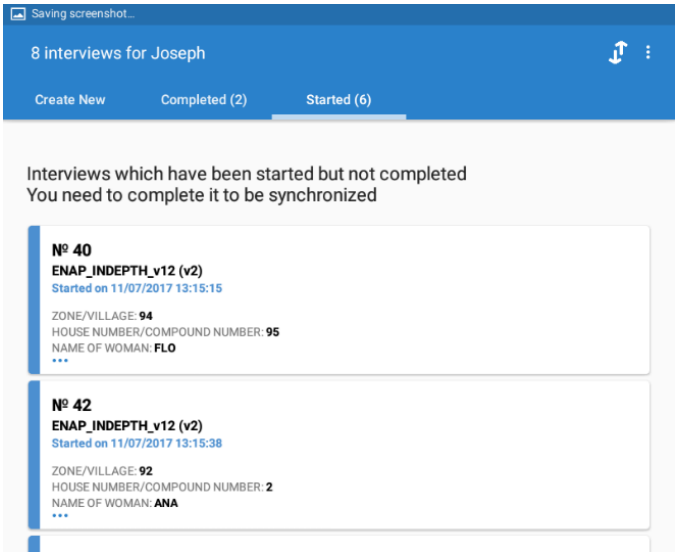
Remember that this information should match exactly as you set it up earlier otherwise you will receive an error message that the information provided is wrong, therefore, the tablet will not be able to link to the server.

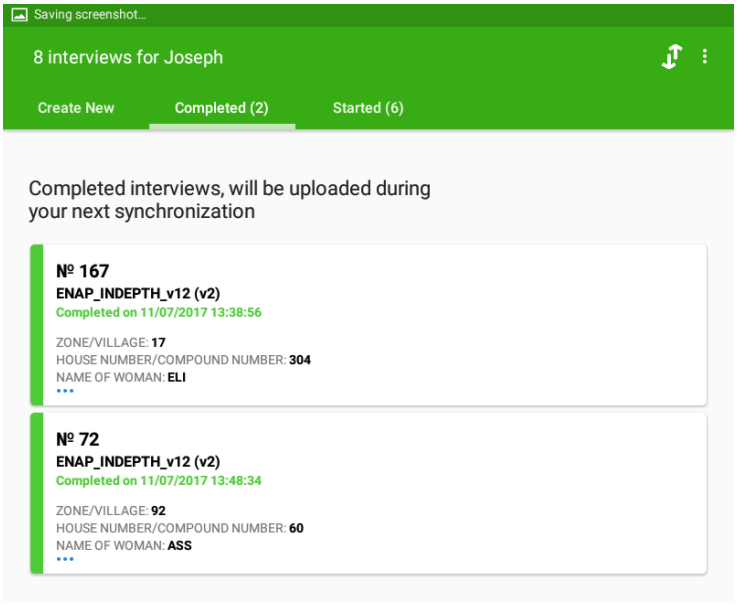
After logging into the interviewer APP, the device should be connected to the server by connecting it to the internet. This will allow for the synchronization to take place. After synchronization the interviewer will be able to see his assignments.

IMPORTANT! Due to security set-up and the way the app operates, once an interviewer logs into the tablet for the first time the tablet it registered as assigned to the interviewer credentials. This means, that it will be not possible for another interviewer to log in into the same device, as the IP and tablet details will already have been linked to the details of the first interviewer. Therefore, it is good practice to label your devices beforehand and assign them to a specific interviewer before they are deployed into the field to start the data collection.

The only way to reassign the device is to uninstall the Survey Solutions' interviewer APP and reinstalling it again but with this process you will lose all the data on the device that has not yet been synchronized and sent to the server to be reviewed by the supervisor. Therefore, it is important, that all devices are synchronized on a regular basis and should the need always sync the device before it is reassigned.







Top left: Interviewer APP interface for new assignments

Top right: Interviewer interface for started assignments

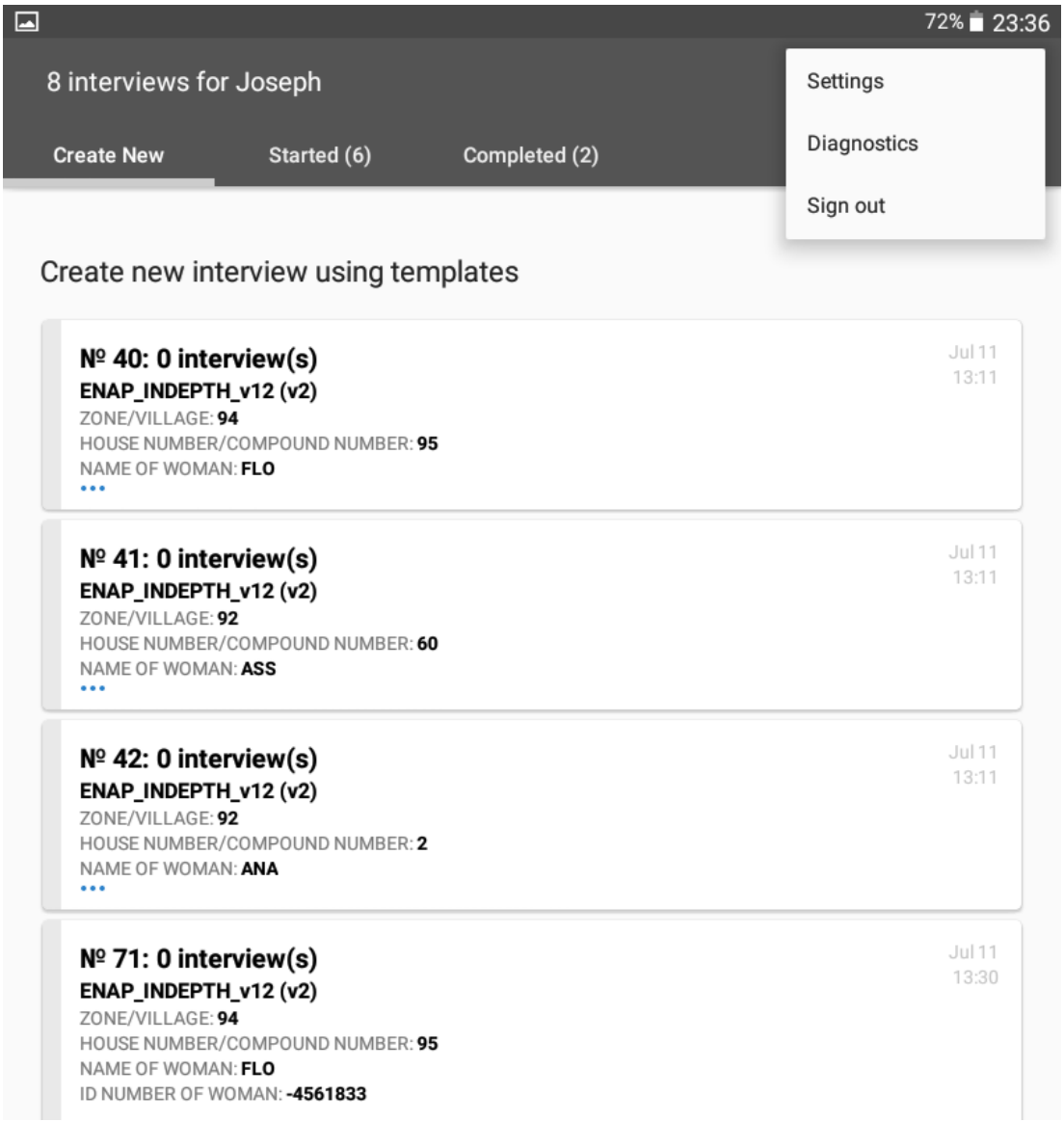
Bottom left: Interviewer APP interface for completed interviews.

The ICON on the interviewers APP top right corner with two arrows pointing in opposite directions is used for synchronization.

The interviewer dashboards in the pictures above give an overview if the interviewer's assignments at hand and the level or state of completion for each assignment. In case any of the submitted interviews are rejected, an additional dashboard will appear in between the started and completed dashboards.

Starting a new interview

The interviewer should make sure s/he is in the **Create New** tab to find her new assignments. For each assignment, the identifier information to find the woman of interest will be shown in a small rectangular cards. For example, if the supervisor or headquarters assigned five assignments to the interviewer, he or she will see five rectangular cards on her **Create New** tab.



When s/he taps on one of the rectangular boxes, the assignment will expand making the identifier details more visible and he or she can make an informed decision on whether to start that interview or not. If s/he decides to start the interview, s/he is meant to tap on the **START** button that appears for that specific interview in order to load it. Below is the tester app user guide in the Annex. Some of the functions are similar.

Annex

Guidance Note for using the Tester app and access the latest version of the ENAP questionnaire.

This document has been compiled to assist you with the step-by-step installation and use the tester APP to test the INDEPTH-ENAP questionnaire template.

You will be required to have access to an Android tablet or phone.

The Tester APP will allow you to access the latest version of the questionnaire.

Please follow the instructions below to download the mobile app and access the questionnaire.

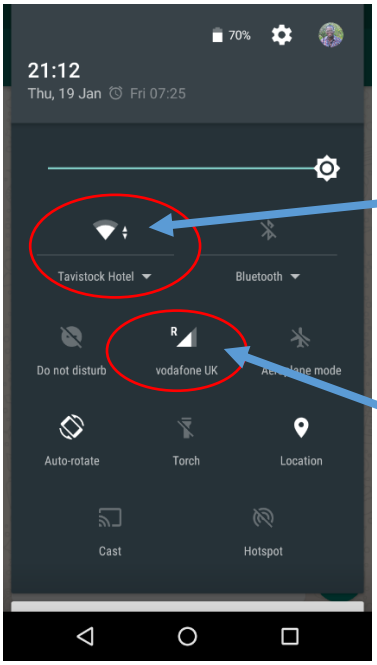
Please circulate this note to the researchers involved in the study in your site and compile and send us comments by **COB the 19th of July.**

These comments will be used to Improve the performance of the questionnaire template that will be implemented in the next few months.

NOTE: ***You need to ensure you have at least 204 Megabytes of storage space left on your device before installing the Survey Solutions Tester APP*.**

STEP 1: Connect to the WIFI from your Android device.

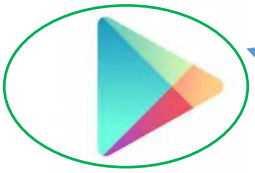
Connect your Android device to the internet using either 3G/4G or WIFI. This is made possible by taping the WIFI icon or the mobile data icons



WIFI connection

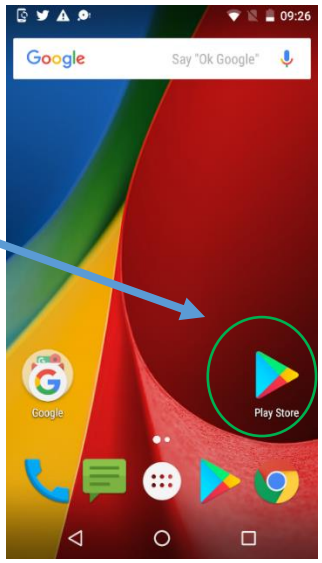
Mobile data

Once your device is connected to the internet you will now need to open the Google Play Store APP. If you cannot find this APP, you need to look for an APP that looks like the image below



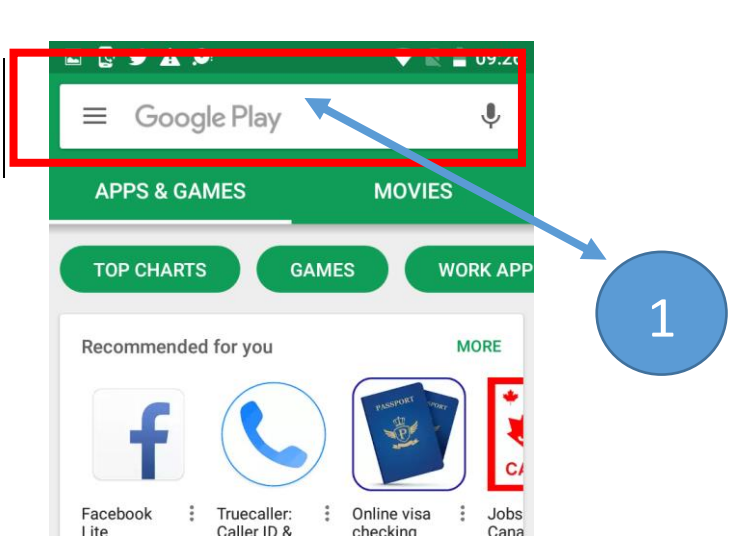
Google Play Store

The Play Store should appear like is shown here on your phone.

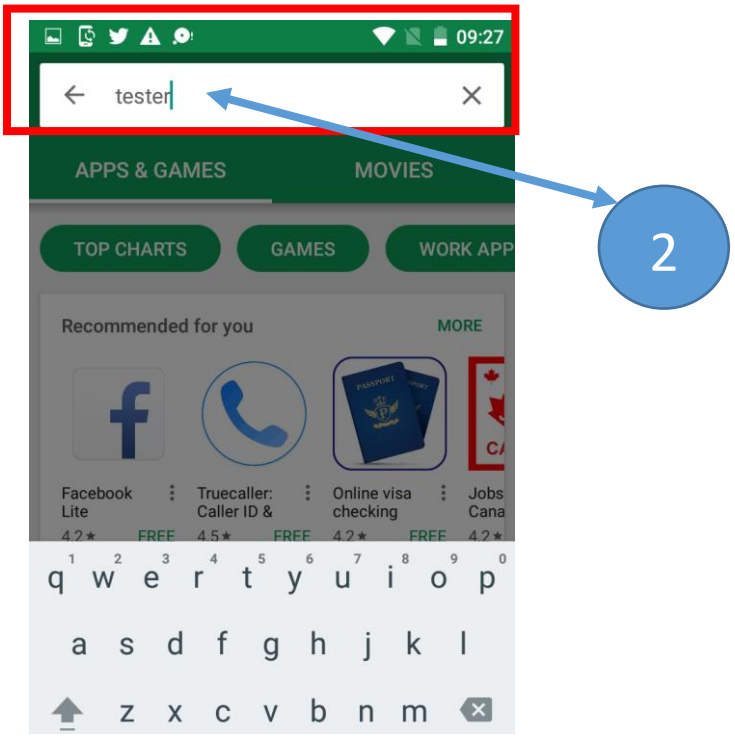


2. STEP 2 – Download Survey Solutions Tester from Google Play.

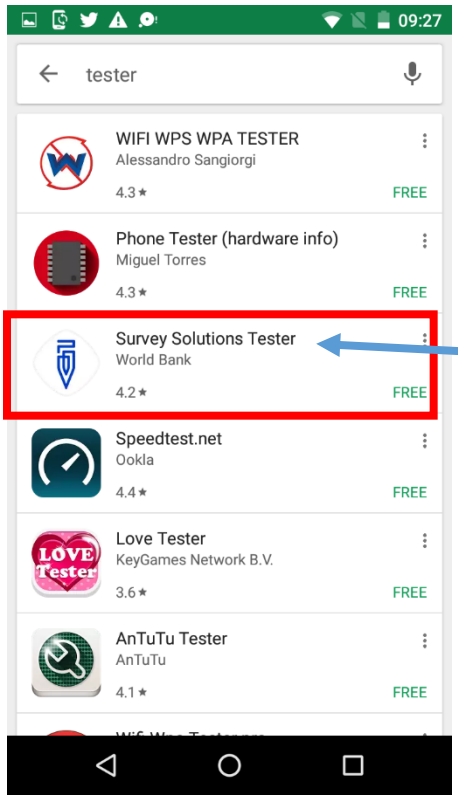
When your Play Store APP is open, you now need to search and download the Survey Solutions Tester APP using the steps shown below.



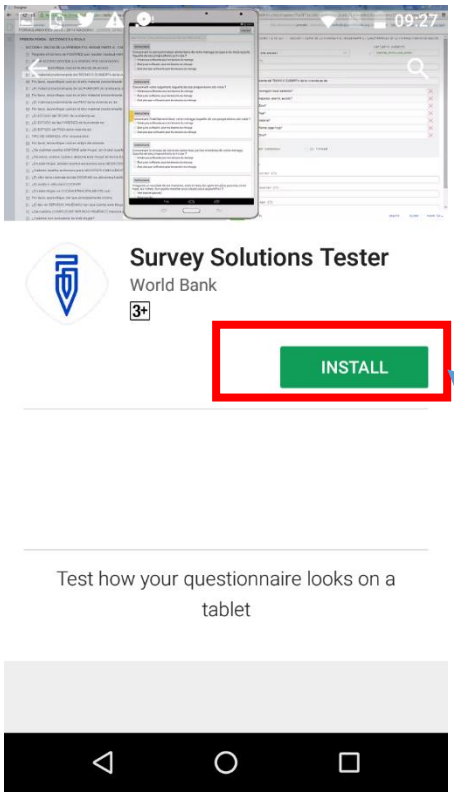
Tap on the highlighted part on your tablet/phone



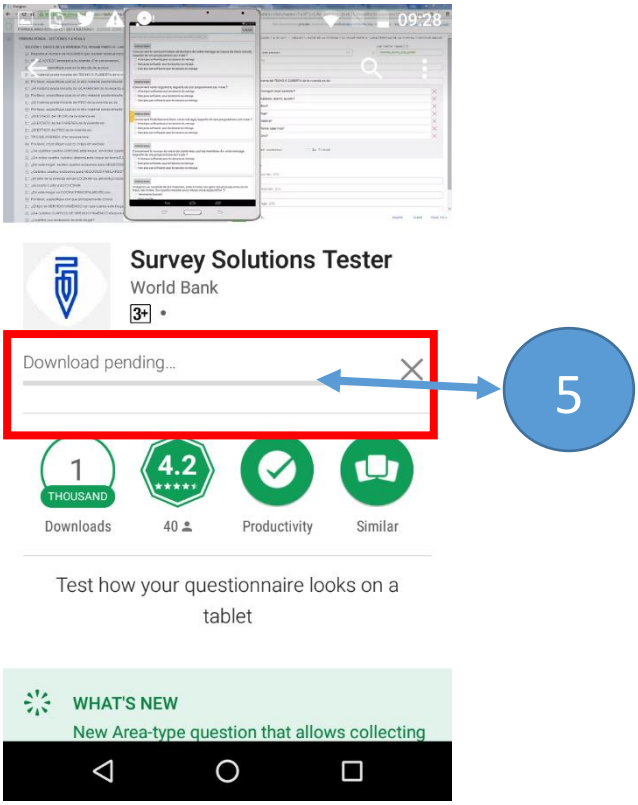
Type Tester in the highlighted part to search for the Tester APP



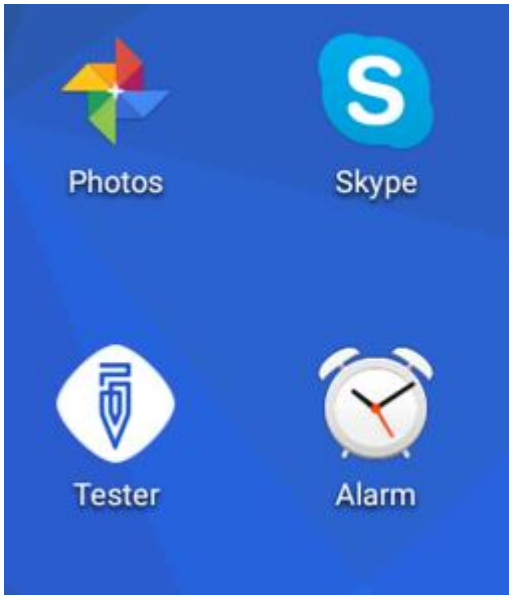
Tap on the Survey Solutions Icon to open for the installation



Tap on the installation Icon to install Survey Solutions Tester on your Tablet of Phone



Wait for the APP to completed download and installation



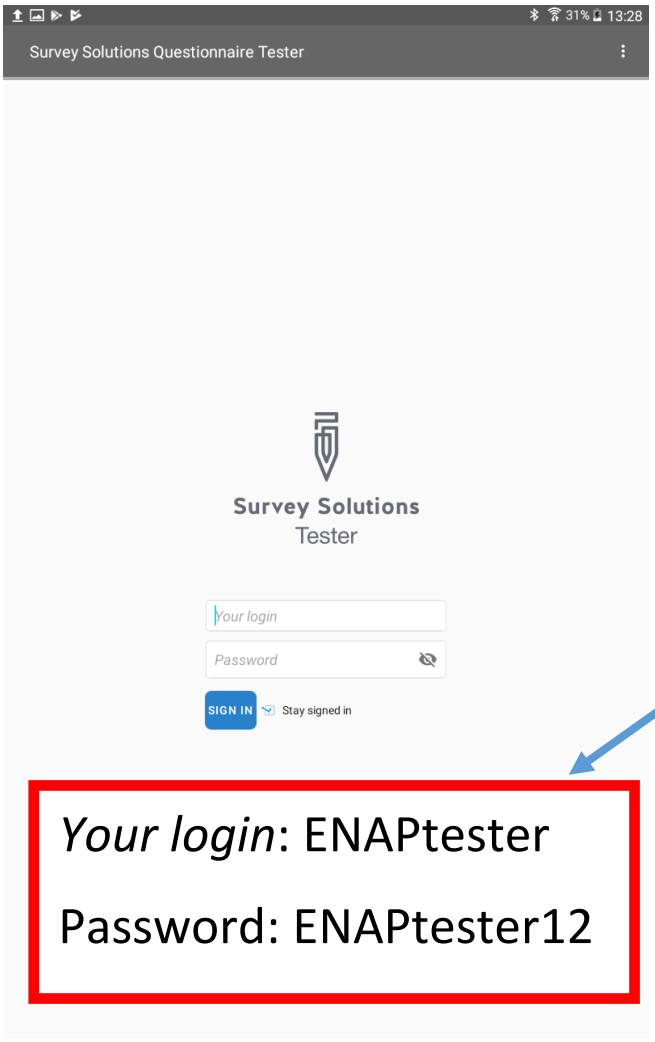
Once installation is complete, the Tester Icon should appear on your device.

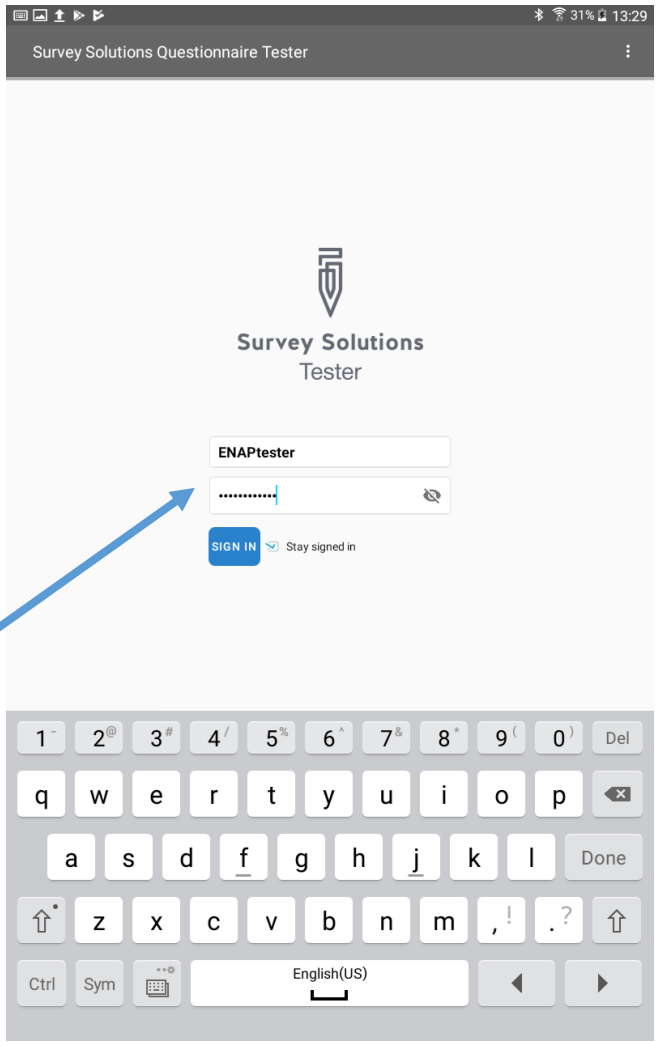
Note: ***The Tester APP has a white background. This should not be confused with the Interviewer APP.***

The Tester APP does not store any data. It is only used for training purposes only.

3. STEP 3 Log in in the Tester App.

When the download and installation of the Tester APP is complete, double click on the Tester APP to open it. The hope screen should look like the ones presented below. In order to access the INDEPTH-ENAP questionnaire template for testing, you should login using the **ENAPtester** username and **ENAPtester12** as the password.





Your login: ENAPtester

Password: ENAPtester12

When the login is successful, you should now be able to view the available questionnaire template for testing as is shown below in STEP 4. You should now click on the ENAP_INDEPTH_V## template link to activate the questionnaire for testing.

NOTE

BEFORE STARTING TO TEST THE QUESTIONNAIRE TEMPLATE, WE ADVISE YOU TO TAKE NOTES FOR SUGGESTIONS OR COMMENTS ON THE DIFFERENT SECTIONS AND QUESTIONS IN THE APP.

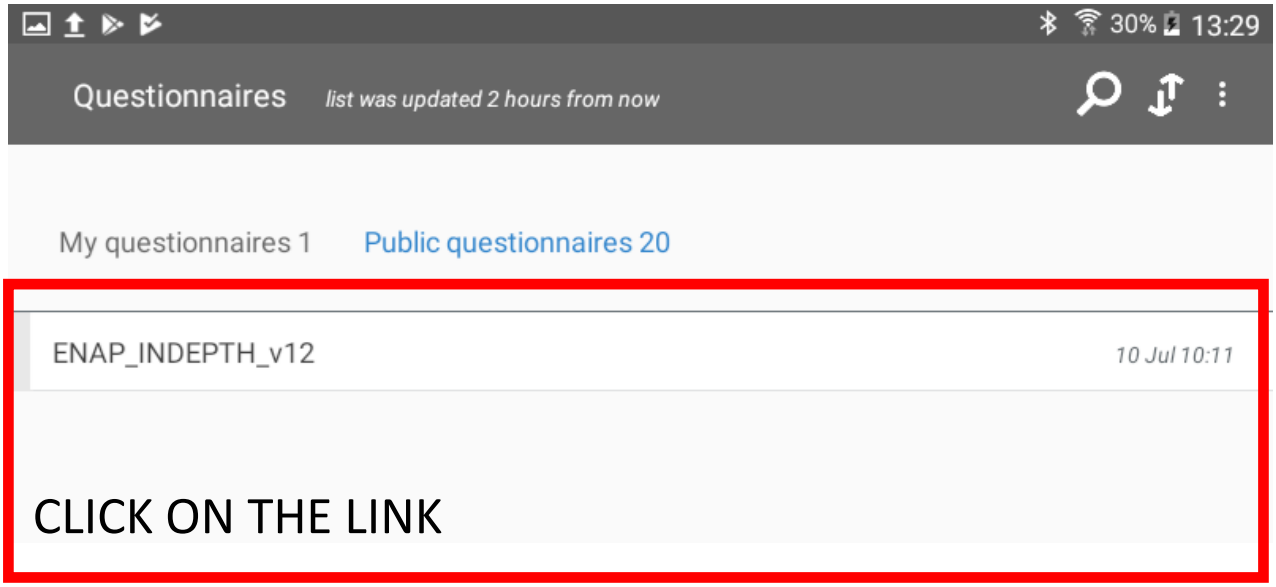
REMEMBER TO NOTE DOWN THE QUESTION NUMBER OR SECTION NUMBER.

REMEMBER THAT AT THIS STAGE WE EXPECT ONLY THE DIFFERENT HDSS CORE SCIENTISTS WHO ARE ALREADY FAMILIAR WITH THE ENAP STUDY TO CONDUCT QUESTIONNAIRE TEMPLATE TESTING. FEEDBACK ON ANY INCONSISTENCIES OR CHALLENGES WHEN TESTING THE APP SHOULD BE SUBMITTED TO THE INDEPTH-ENAP CORE TEAM FOR IMPROVEMENTS TO FOLLOWING EMAIL ADDRESS jakuze@musph.ac.ug BY 19TH JULY 2017.

LASTLY, PLEASE COMPILE ALL THE FEEDBACK FROM YOUR HDSS SCIENTISTS INTO ONE MS WORD DOCUMENT TO AVOID SUBMITTING DUPLICATED/MANY NOTES.

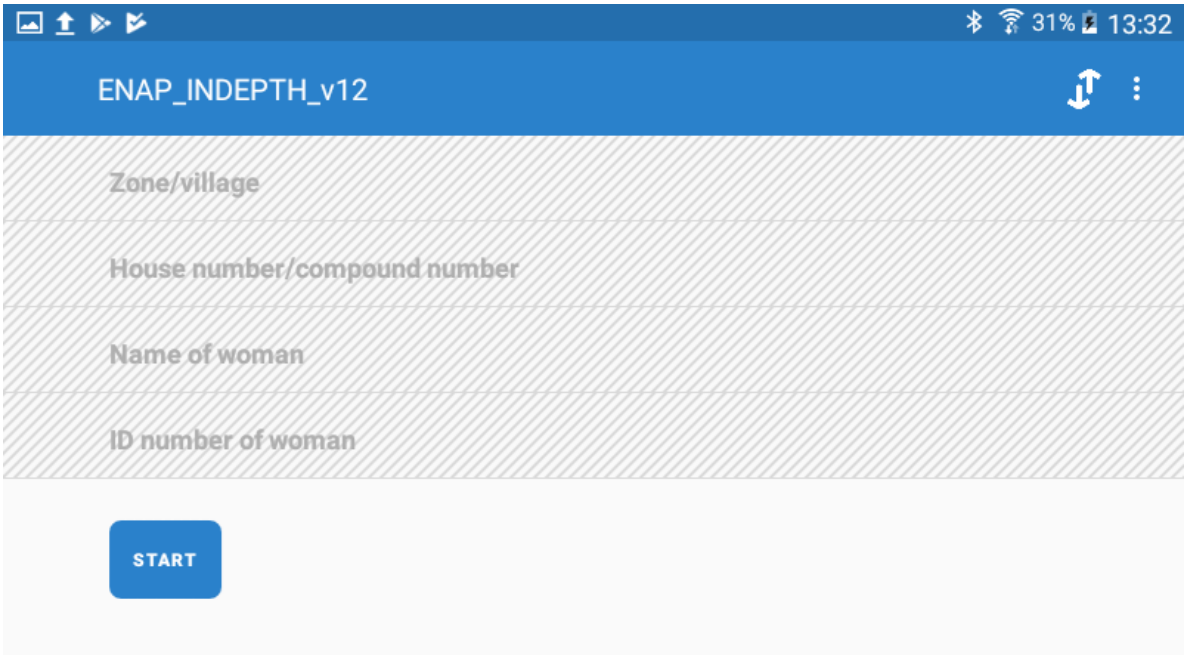
4. STEP 4: Access the latest version of the ENAP/INDEPTH Survey Questionnaire.

Click on the link to activate the questionnaire template. Click on the most recent or newest version of the INDEPTH_ENAP questionnaire template incase more than one version of the templates appears on your device.



5. STEP 5: Check the latest version of the survey questionnaire.

Tester should appear like in the picture below. The grayed out words are the identifier variables that your HDSS will have selected. You should tap on the **START** button to begin testing

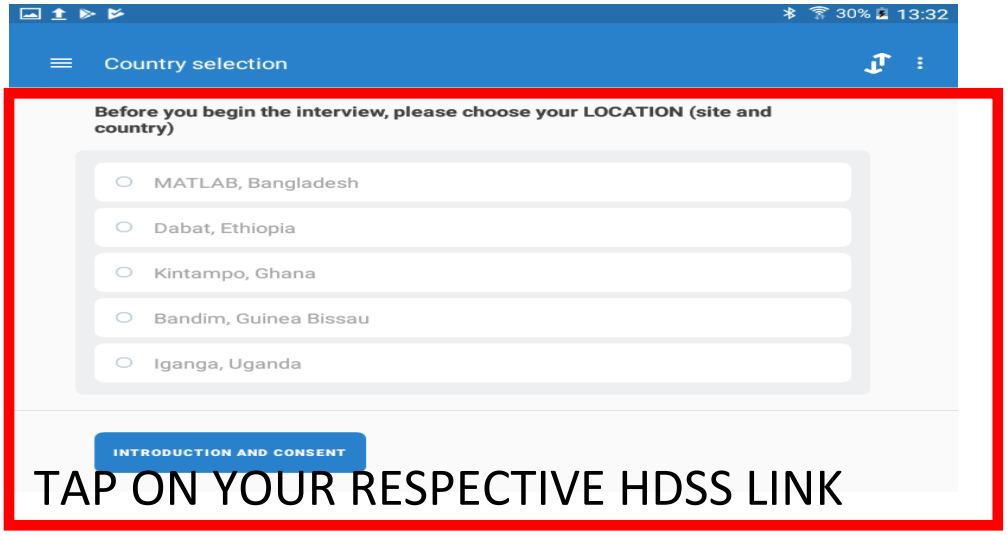


NOTE:

1. ***Some of the questions that are add-on to the INDEPTH-ENAP survey and will not appear for certain countries. An example vaccination questions that are highlighted YELLOW in the MS Excel version of the questionnaire template that have been selected by Bandim HDSS.***
2. ***All sections in the MS-Excel that appear in GREEN, the Survey Solutions APP will do this for you***

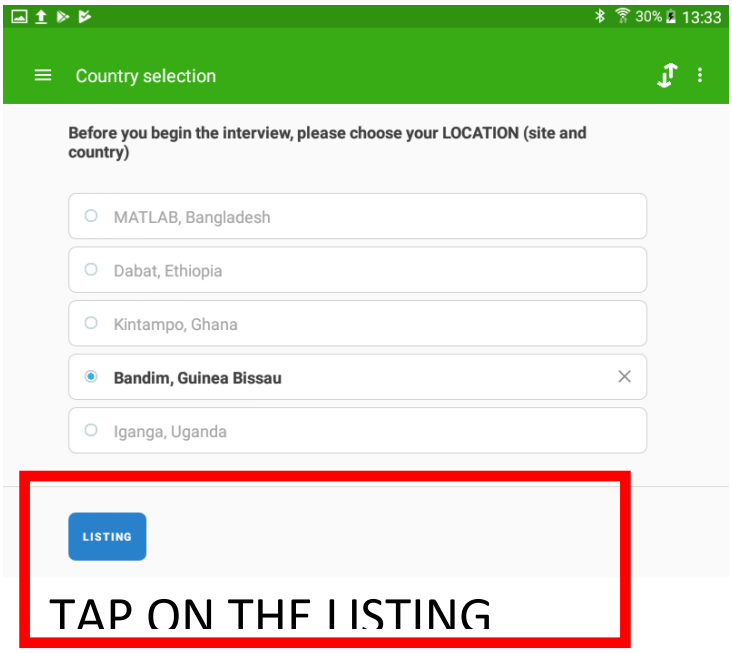
After tapping the **START** button, the APP should direct you to the next page where you are required to select the country in which your HDSS is located.

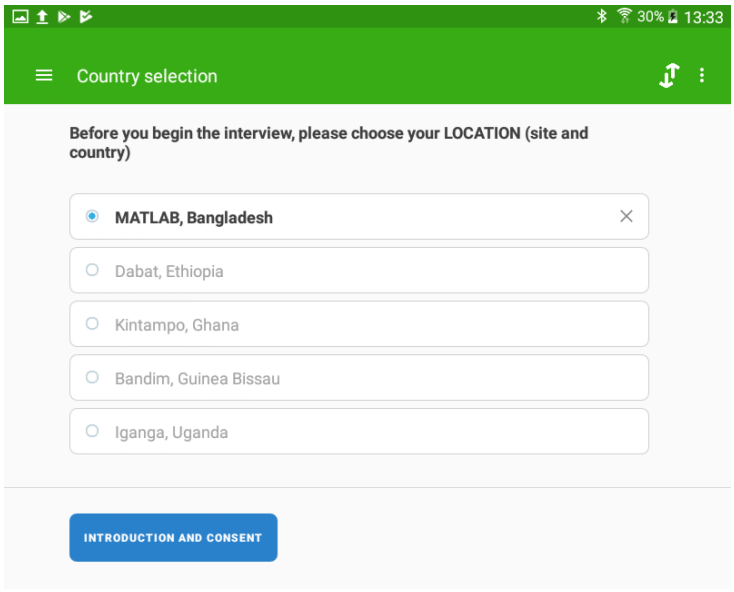
Country selection.



Listing variables

After selecting your respective HDSS, Since Dabat, Kintampo, Iganga-Mayuge HDSS and Matlab HDSS sites will not see the **LISTING** tab because a template for the woman listing has not yet been created. However, Bandim HDSS where the Listing has been created, the **LISTING** Tab will appear and this will be prepopulated with the woman's information in the final interviewers APP. These variables need to be identified by each HDSS.





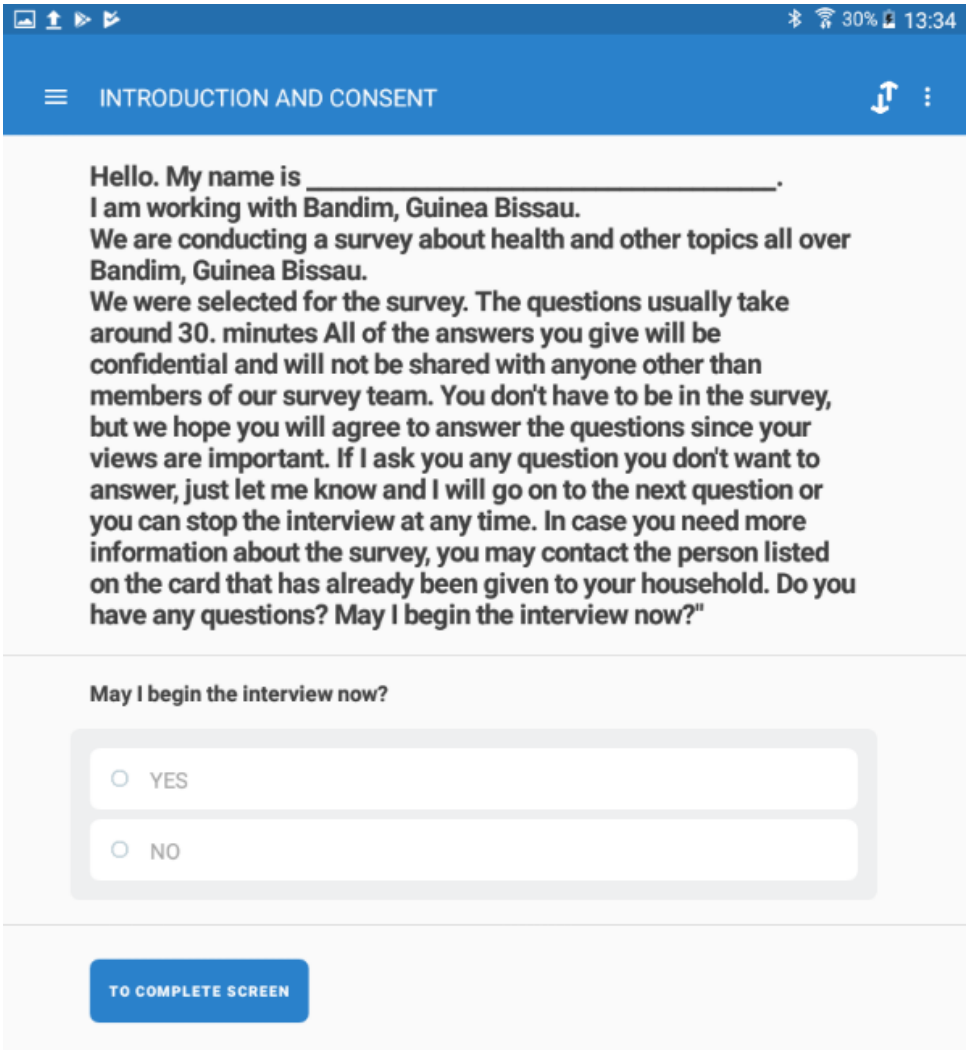
NOTE: YOU ARE USING THE TESTER APP THEREFORE, THE PREPOPULATED INFORMATION WILL NOT APPEAR.

Consent form and signature

After completing the Listing, the APP will direct you to the **INTRODUCTION AND CONSENT** section, where you can begin the testing the functionality of the questionnaire template.

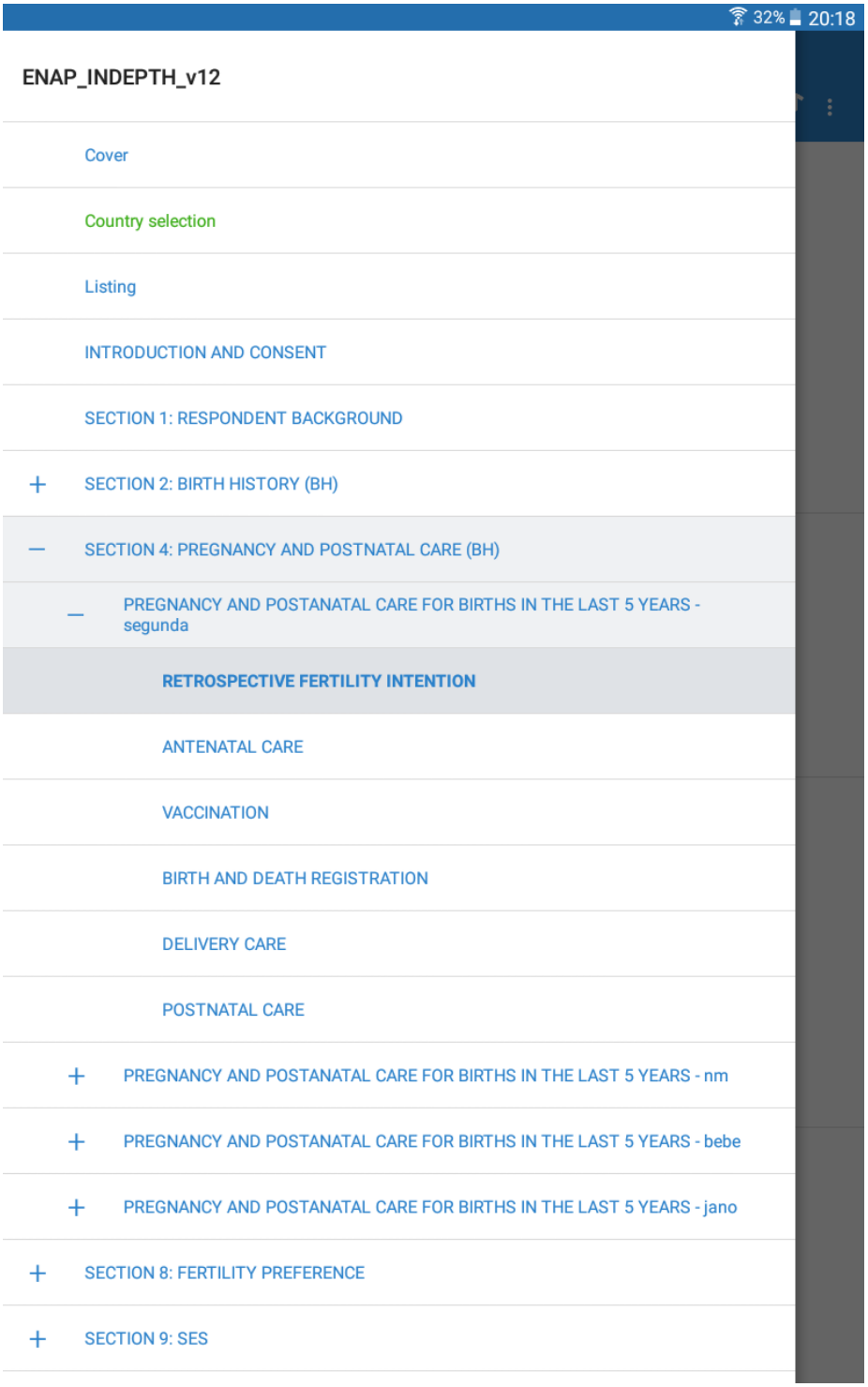
NOTE:

1. If you do not tap on the options for the question "***May I begin the interview?",*** You will not be able to proceed testing the questionnaire functionalities.



Sturecture of the questionnaire

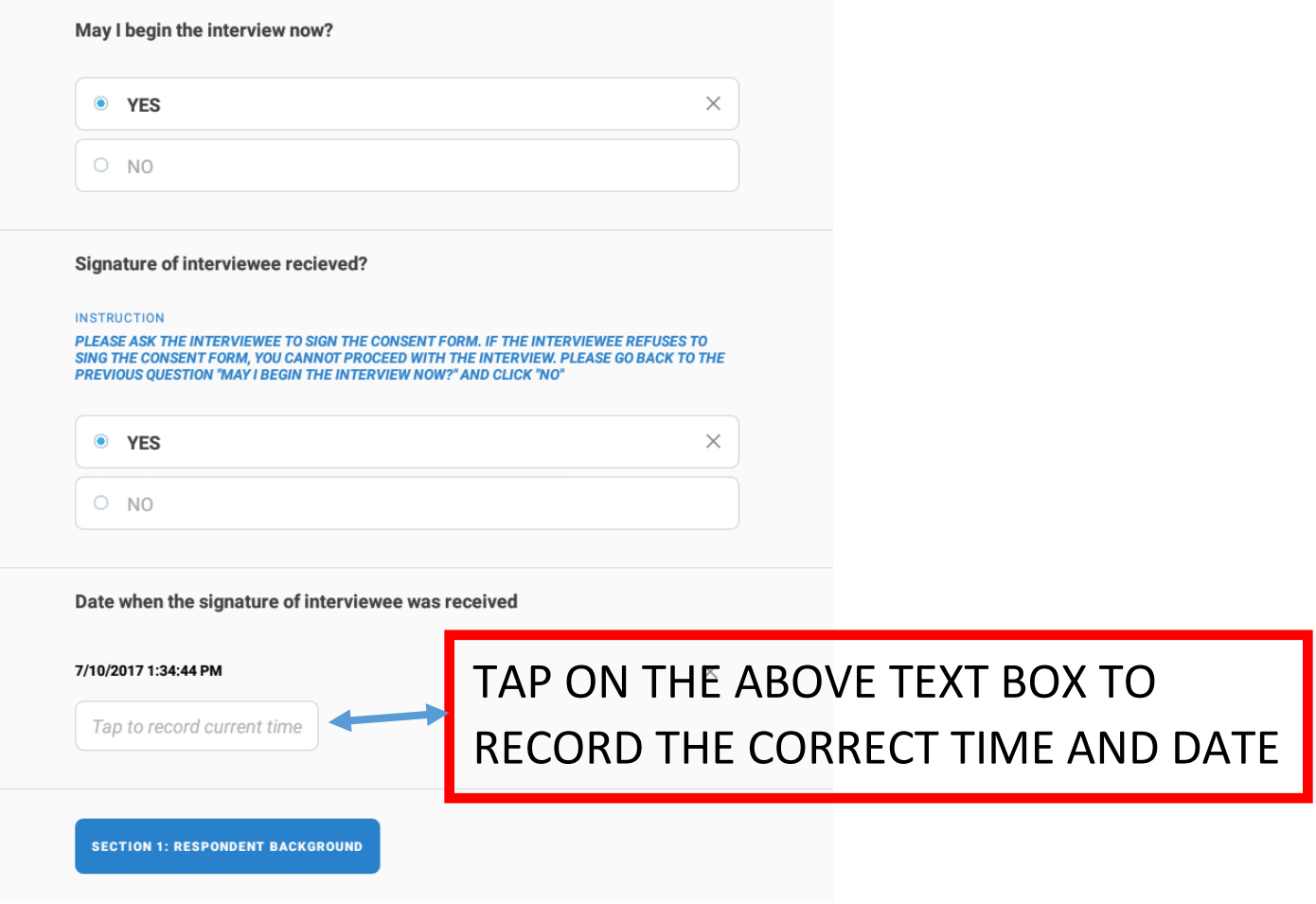
Below is the structure of the questionnaire template which can be accessed by tapping the expansion icon on the top left corner which appears like this



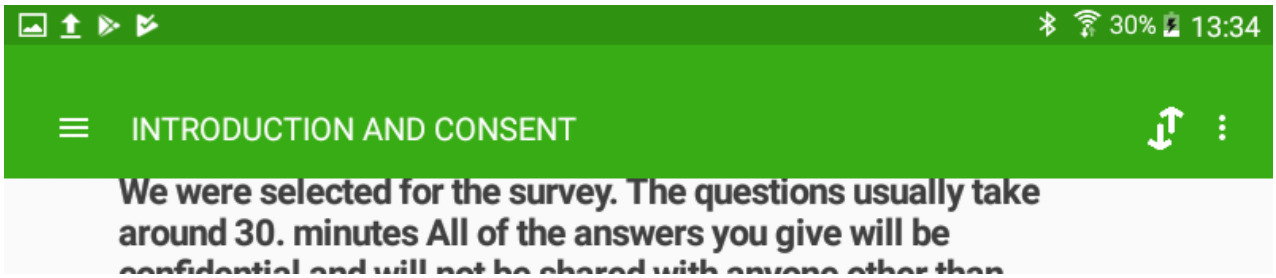
Interview Progress

At this point you are now ready to proceed testing the questionnaire template. Below are some of interface options that will notice while testing the APP and the questionnaire template.

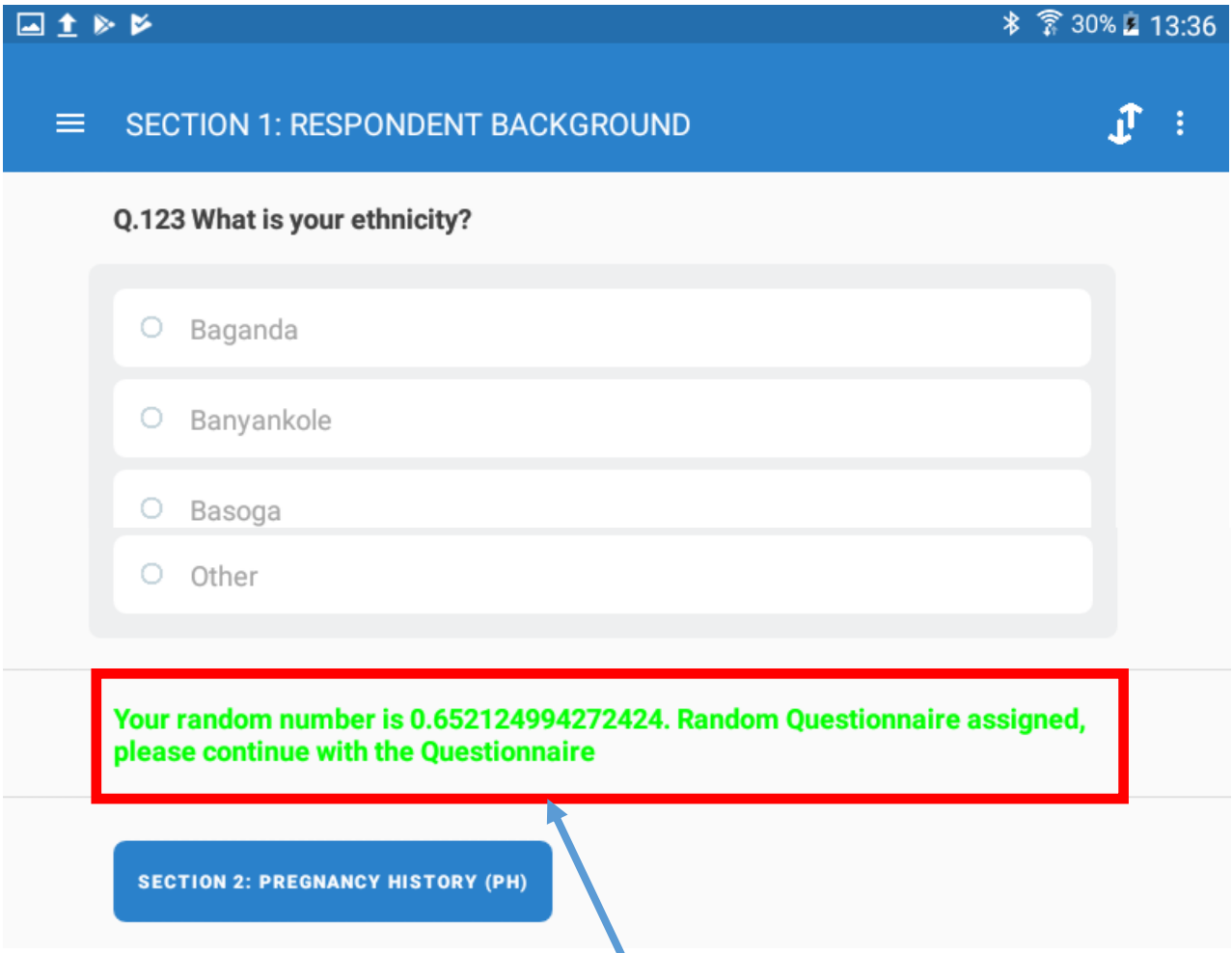
a. **When completing the INTRODUCTION AND CONSENT section, ensure that you remember to tap on the text box below the question "Date when the signature of the interviewee received" Otherwise the APP will not be able to compute correct values of the age of the individuals in question in the preceeding sections**



b. When you succefully complete a section, the section Header of that section will turn from BLUE to GREEN as is shown in the Pictures below



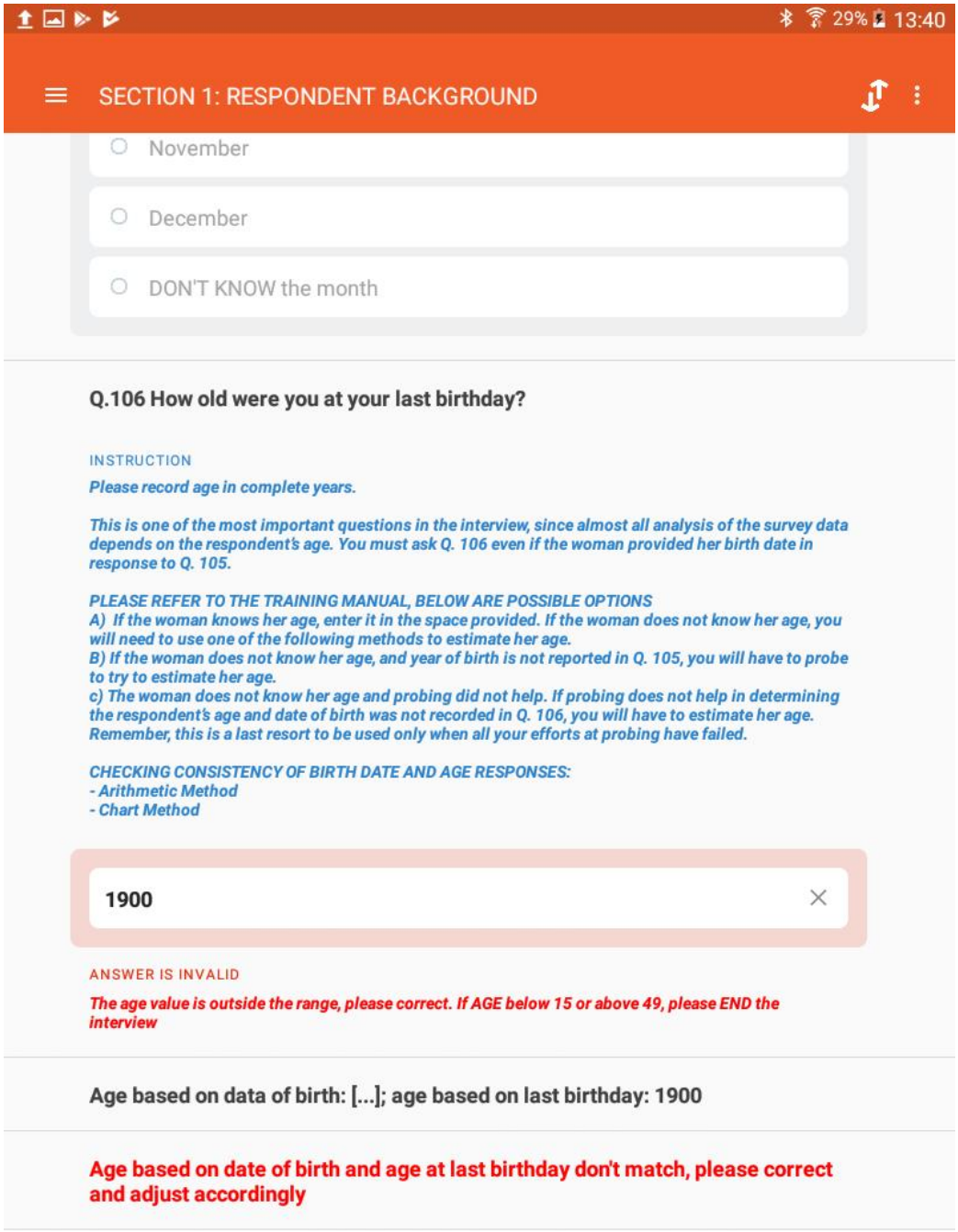
c. When you complete **SECTION 1: RESPONDENT BACKGROUND,** The APP will randomize you to either the **PREGNANCY HISTORY SECTION TWO** or the **BIRTH HISTORY SECTION TWO** as is expected and is clearly elaborated in the study protocols that you have all submitted to your Institutional Review Boards.



IF THE RANDOM NUMBER IS GREATER THAN 0.5 YOU WILL BE ASSIGNED THE PREGNANCY HISTORY SECTION 2, ELSE YOU WILL BE ASSIGNED THE LIVE BIRTH HISTORY MODULE FOR SECTION 2

Error messages

Erroneous entries that are entered or entries that are outside the expected range will flag a **RED** wording, and the **SECTION HEADER** will also flag a **RED** as is shown below.



A8: Data management plan

Data Management Plan for Research Students

Name	JOSEPH AKUZE WAISWA
Email	Joseph.waiswa@lshtm.ac.uk
Title	Research Degree Student
Date	16th May, 2018
Supervisors	Dr. Hannah Blencowe
	Prof. Simon Cousens
	Assoc. Prof Peter Waiswa

Support

Information on writing a Data Management Plan can be found at http://www.lshtm.ac.uk/research/researchdataman/plan/

One-to-one advice is available through the RDM Support Service researchdatamanagement@lshtm.ac.uk

DATA DESCRIPTION

What data will you collect or create?

Describe the data that you are collecting or creating in your project. Relevant information to provide includes:

The type of information that will be contained. E.g. MRI scans, interview transcripts, spatial data, etc.

Methods of capture. E.g. face-to-face interview, web survey, etc.

Amount of data. E.g. 100 patients will undergo an MRI scan, 500 people will be interviewed.

This project will utilise data from the Every Newborn-INDEPTH (EN-INDEPTH) household survey. The EN-INDEPTH household survey is a randomised survey which randomises the pregnancy history and birth history reproductive modules to women who are residents within five Health and Demographic Surveillance Sites (HDSS). The EN-INDEPTH study is part of Every Newborn Action Plan (ENAP) Measurement improvement Roadmap which is aimed to improve household capture of stillbirths and neonatal deaths in surveys. The survey was set to collect data on a sample of about 68,000 births in the five HDSS sites namely; Bandim, IgangaMayuge, Dabat, Kintampo and Matlab.

The survey collects data on background characteristics of the woman, reproduction career of the women using either birth history or pregnancy history including outcomes for the pregnancy or birth, fertility intentions, postnatal care and socioeconomic characteristics using the World Bank's Survey Solutions computer assisted personal interviewing (CAPI) application on android tablets through a household face-to-face interview.

The data collected on the tablets is synchronised with the Survey Solutions web based database hosted by either the London School of Hygiene and Tropical Medicine virtual server or dedicated HDSS local server. This data then undergoes consistency and quality checks.

Briefly describe the key activities that will be performed on your data, from its creation/capture to its eventual archiving or deletion.

Consider the lifecycle of your research data and the actions that will be performed during that time. For example, data may be captured using a web form, anonymised to remove personal information using software X, cleaned using Tool Y to enable it to be analysed, analysed using software Z, and so on. The lifecycle may be written as text or pictorial form (e.g. a gantt chart).

In addition, it's useful to consider the approximate time period when you will perform each action (e.g. data capture in month 2, data cleansing in month 4, etc.).

The data will be collected using the World Bank's survey solutions interviewer applications on android tablets. This will take about nine months from August 2017 up to July 2018. The interviewer will sync his tablet with the server to download a fresh assignment of questionnaires or submit a completed assignment. This process will happen daily when it is possible or weekly or bi-weekly for sites that may have limited internet connection. Before submitting an interview the interview will confirm that the questionnaire is completed to its

fullest. The interviewers will be guided by the colour coding of the interviewer app with blue indicating an incomplete questionnaire, green indicating a completed questionnaire and red indicating errors in the questionnaire. In addition to the colour coding the interviewer will be able to see a summary of the number of questions completed, unfinished or with errors.

Upon synchronisation the data received on the survey solutions server will be reviewed by a supervisor or data manager. This is the second line for the data quality check. The supervisor will review the submitted data for any errors and incomplete questions. He/she is also assisted by the questionnaire summary recorded on the survey solutions headquarters server. If any errors are identified, the supervisor will flag these questionnaire items and add specific comments to the interviewer and then will reject this interview. If no errors are found he/she will accept the interview. All rejected interviews will be returned to the interviewers' app on the next synchronisation.

The data will be downloaded by the data managers and backed up bi-weekly. A data monitoring template will be completed and shared bi-weekly to enable be keep track of the data collection processes and sample size. A descriptive summary of the data collection will be prepared by me.

When all data collection is completed in all HDSS sites around July 2017, all the sites will download a completed dataset from their servers in stata format, will anonymise it and will share it in stata 14/15 software. It will then be cleaned for data analysis using stata 14/15 software. Additionally paradata from the survey processes will be shared and prepared for response time analysis.

Data on questionnaire contexts will be generated with QUAID computerised software. The data cleaning process will start in July up to September 2018. The data analysis will start in October up to January when a manuscript for objective two will be submitted.

What data formats or standards will you use to store data produced by your project?

Outline the data formats, encoding standards, or software tools that you will use to create, analyse, or use data. E.g. data will be captured using a MySQL database and analysed using STATA and MS Access.

The data is stored on the survey solutions web databases will be exported into Stata format for analysis. The datasets from the five HDSS sites will be anonymised and merged for pooled analysis using Stata version 15.

In addition paradata on the timing of completion of the questionnaires in Unicode-text format will be exported.

What quality controls and thresholds will you establish to ensure that your data is fit for purpose

Quality controls may be applied prior, during and following data capture and processing. Possible QC practices include: testing instrumentation to ensure it is correctly calibrated, recording multiple measures, double-entry of information, checking validity of entered values

The survey dataset included several quality checks processes; including the pre-coded questionnaire on the survey solutions platform with value and consistency checks, skip patterns and error messages flagging out to the interviewers during data collection.

Once the data is synchronised and sent to the survey solutions server, it is reviewed at the survey solutions headquarters dashboard for consistency check. Any identified errors or inconsistencies or missing data in the submitted interview are flagged by the supervisors and the interview is then rejected. If the interview is completed successfully it is accepted and is stored in the survey solutions database. The rejected interviews are returned to the interviewers' tablet the next time he/she synchronises with the server. This give the interviewer to revisit the interview to make the necessary adjustments to the interview sections with errors or missing data after consulting the respondent.

What documentation or metadata is needed to understand your data?

Describe the documentation or metadata that you will create to enable the data to be understood and used by your future self and others. It is helpful to consider the following questions:

What information is needed to understand the content and context of its creation?

What documentation and metadata standards will be used?

How will potential users find out about your data?

Data dictionaries and metadata will be created be created for the dataset. In addition a document will be created containing all the variable names and value labels that will be included in this analysis. The final dataset will be anonymised.

Potential future users will be able to access the anonymised dataset upon request as per the data sharing agreement.

DATA STORAGE AND MANAGEMENT

Where will you store data during the project lifetime? (tick one or more)

School PC local drive (drive C: or D:)		Personal area on School network (drive H:)	ü	LSHTM Shared Network drive (e.g. I: drive)	ü	Dedicated server maintained at partner institution	ü
LSHTM-based project server	ü	School laptop or tablet		LSHTM Secure Data Server (for confidential data)	ü	LSHTM Novell Filr	

For-cost cloud service (e.g. Amazon S3)		Free cloud service (e.g. Dropbox, Google Docs)		Portable storage (e.g.USB disk or memory stick)	ü	Other. Please indicate	

Other

How will you organise and label your data?

Describe the approach you will take to structure and label your data. E.g. files and folders on a storage device, database tables and labels.

The datasets for this analysis will be extracted and from EN-INDEPTH core dataset and will be linked to the pregnancy history and birth history roster datasets using the unique interview identifiers. The roster datasets will be reshaped into the wide format.

The smaller dataset will be renamed accordingly, label values, variable labels and names will be generated and a list of these variables will be generated.

What security measures, if any, will you apply to protect data? (tick one or more)

Controlled access limited to authorized users only	ü	Physical security		Remove identifiable information (e.g. anonymisation)	
Data storage encryption	ü	Data transfer encryption	ü	Password protection	
Process on isolated machine in secure room		Secure deletion following analysis			
Avoid use of third party storage, such as Dropbox	ü	Other			

Other

DATA ARCHIVING AND SHARING

What data do you need to keep after your project ends and for how long?

The data from this project will be available and disseminated in the manuscripts and thesis, in addition the raw datasets will be available and kept at the London School of Hygiene and Tropical Medicine and other collaborating institutions as per the data standard of operation procedures and the data sharing agreements. The duration for storage before it is available for open access is ten years.

Where will data be kept after your project has finished (tick one or more)

Research data may be submitted to a data repository or data archive, which will handle the process of curation, preservation and sharing on your behalf.

I will keep the data myself		My supervisor will look after the data	ü	It will be looked after by the project team	ü
Held in the LSHTM Research Data Repository	ü	Held in a LSHTM-maintained project system	ü	Held in a 3rd party data repository. Please specify in Other field	ü

Other

Makerere University School of Public Health, INDEPTH Network, IgangaMayuge, Bandim, Dabat, Matlab and Kintampo HDSS data repositories

Can data be made available to anyone? If not state the reason it needs to be restricted and criteria for gaining access.

Can data be made freely available to anyone or do restrictions need to be applied? This question will help you to consider whether access controls need to be applied to limit data access. Potential reasons for restriction include the need to comply with consent agreements, which state:

- Data can only be used by specific users, e.g. researchers working in an academic environment, a specific skill set, etc.
- Data can be analysed only for specific purposes compatible with the consent agreement.

If data does need to be restricted, state the reason and the criteria that users would need to meet to gain access

The data from this project will be made available for sharing amongst the contributing partners according to the guidelines defined in the data sharing agreement. External parties will be welcomed to use the data upon requests made to the collaborating institutions. This will be based on the guidelines outlined in the data sharing agreement signed by all the collaborating institutions.

What actions will be performed to prepare your data for access?
(tick one or more)

Removal of personal information	ü	Add synthetic data (e.g. pseudonyms)		Copyright clearance	

Establish participant consent		Develop an access agreement	ü		

Other

RESOURCING

What do you consider to be the primary data management challenges in your project?

What problems or issues do you need to address in your project.

The EN-INDEPTH survey datasets usually have so many variables (both the survey data and paradata). Therefore I anticipate challenges in handling and manipulating them for analysis.

What resources would it be helpful for the School to provide to help deliver your plan?

How can the School help you to manage your data? E.g. training, specific IT Services, etc.

I will get a support from the school in terms of short course trainings in advanced analysis and management of both survey and paradata