Quality of care and quality of data for hospital births – tension or traction?

Dr Louise Tina Day

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy of the University of London.

Department of Infectious Disease Epidemiology and International Health
Faculty of Epidemiology and Population Health
London School of Hygiene & Tropical Medicine

January 2024

Work funded by
the Children’s Investment Fund Foundation,
the Bill & Melinda Gates Foundation,
the United States Agency for International Development-funded Data for Impact project,
the Chiesi Foundation

Supervisors: Associate Professor Cally Tann
Centre for Maternal, Adolescent, Reproductive and Child Health (MARCH)
London School of Hygiene & Tropical Medicine
Honorary Principal Scientist, MRC/UVRI & LSHTM Uganda Research Unit
Consultant Neonatologist & RCPCH College Tutor, University College Hospital

Professor Carine Ronsmans
Maternal and Newborn Health Group
Department of Infectious Disease Epidemiology
Faculty of Epidemiology and Population Health,
London School of Hygiene & Tropical Medicine.

Associate Professor Hannah Blencowe
Maternal and Newborn Health Group
Department of Infectious Disease Epidemiology
Faculty of Epidemiology and Population Health,
London School of Hygiene & Tropical Medicine.

Advisory Committee:

Professor Joy Lawn
Director Centre for Maternal, Adolescent, Reproductive and Child Health (MARCH)
Faculty of Epidemiology and Population Health,
London School of Hygiene & Tropical Medicine.

Assistant Professor Paul Mee
Department of Infectious Disease Epidemiology
Faculty of Epidemiology and Population Health,
London School of Hygiene & Tropical Medicine.

Declaration of own work

I, Louise Tina Day, 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: [redacted]

Date: 23rd January 2024

Abstract

Background

To end preventable maternal and newborn mortality and stillbirths, attention is shifting to improving quality of care, both provision and experience. Health information systems can improve health system performance when data are used for decision-making. High quality data are urgently needed to assess progress and accelerate efforts towards the globally agreed 2030 Sustainable Development Goals. Improving data for action is a focus of recent global movements to improve outcomes for women and newborns and reduce stillbirths.

Most preventable deaths of newborns, stillbirths and women occur in settings with data gaps that impede progress for coverage, equity, and quality of care. Routine health information systems (RHIS) make use of data documented by health professionals as part of the treatment they offer in health facilities. With the proportion of births occurring in health facilities rising, there is increasing interest in using routine health facility data to track intrapartum care provided for women and newborns, especially as this is the time during pregnancy that contributes most to women and newborns surviving and thriving. The quality of data in routine labour and delivery registers, including how accurately it captures the care provided, is currently understudied.

Methods

My thesis describes a completed body of work that explored the quality of routine labour and delivery register data. Analyses used the 'Every Newborn – Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) dataset from five comprehensive Emergency Obstetric and Newborn Care hospitals in Bangladesh, Nepal and Tanzania.

My focus is on the data at the foundations of the data-information pyramid, with a particular emphasis on core indicator data elements collected in routine labour and delivery registers. These core indicator data for RHIS tracking are explored for the maternal-newborn dyad, structured by two sections:

Section A: Assessing labour and delivery routine register indicator data quality for hospital births.
Section B: Identifying opportunities to improve labour and delivery routine register data quality for hospital births.

Results

Labour and delivery register data were available, legible, and complete however data quality was very mixed, varying by indicator and hospital. Frontline health professionals in high-mortality settings face the tension of dual demands: providing high quality of care for women and newborns, while also documenting routine data on care and outcomes.

Opportunities to enable a virtuous cycle of data use and data quality from labour and delivery ward registers were identified.

A novel '**Quality of Care and Quality of Data conceptual framework'** is presented, linking delivery register data quality to the WHO domains of quality of care. Central to this is a missing link expressed as the **'Data Quality Continuum'** – that the hospital routine data culture determines quality of neonatal data used both for clinical care and to track outcomes.

Conclusion

Improving the quality of care is impeded by a lack of high quality data that can be used both for clinical decisions and by policy makers for planning and investment. RHIS strengthening needs to overcome the tension for frontline health workers between care and data and create traction to enable high quality data for use to improve quality of care.

Acknowledgements

This PhD seeks to contribute to ending preventable maternal and newborn mortality and morbidity:

"See, I will create new heavens and a new earth……
the sound of weeping and of crying will be heard in it no more….
Never again will there be in it an infant who lives but a few days,
or an old woman or man who does not live out his years….."

Isaiah 65 v 17-20

~~~

With thanks to:

My PhD supervisors: Associate Professor Cally Tann, Professor Carine Ronsmans and Associate Professor Hannah Blencowe and advisory committee Professor Joy Lawn and Assistant Professor Paul Mee.

The 'Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) study group:
Principal Investigators: Professor Joy Lawn (London School of Hygiene & Tropical Medicine (LSHTM), Dr Shams El Arifeen (icddr,b Bangladesh), Dr Ashish KC (Golden Community, Nepal) and Dr Honorati Masanja (Ifakara Health Institute, Tanzania).
Co-investigators - including joint first authors: Harriet Ruysen, Georgia Gore-Langton, Qazi Sadeq-ur Rahman, Donat Shamba, Stefanie Kong, Ashish KC, Kimberly Peven.

~~~

This PhD is dedicated to:
1) my colleagues in LAMB Integrated Rural Health and Development Management Information Systems-Research Department and Hospital, Bangladesh where my journey for reducing the tension between data and care in high mortality / morbidity settings began. Thank you for our learning together, generating traction between high quality data and high quality care.
2) my parents and family, thank you….

~~~

Appreciating many colleagues and collaborators for demonstrating kindness: listening, encouraging, sharing, R teaching, reviewing and caring…….

"There is no safe investment. To love at all is to be vulnerable.
Love anything, and your heart will certainly be wrung and possibly be broken.
If you want to make sure of keeping it intact,
you must give your heart to no one, not even to an animal.
Wrap it carefully round with hobbies and little luxuries; avoid all entanglements;
lock it up safe in the casket or coffin of your selfishness.
But in that casket – safe, dark, motionless, airless – it will change.
It will not be broken; it will become unbreakable, impenetrable, irredeemable."

C.S. Lewis, The Four Loves
~~~

Table of Contents

Contents

List of figures

List of tables

Table of Abbreviations

ACS	Antenatal corticosteroids
BD	Bangladesh
BMV	Bag-mask ventilation
CEmONC	Comprehensive Emergency Obstetric and Newborn Care
CIFF	Children's Investment Fund Foundation
CRVS	Civil Registration and Vital Statistics
DALYS	Disability Adjusted Life Years
DHIS2	DHIS2
DHS	The DHS program, Demographic Health Surveys
DTP/ Penta 3	Diphtheria, Tetanus, Pertussis, Hepatitis B and Hib vaccines combined as "Pentavalent vaccine" reported as DTP/Penta3 in the SCORE report.[7]
EmONC	Emergency Obstetric and Newborn Care
ENAP	Every Newborn Action Plan now branded as Every Newborn
EN-BIRTH	'Every Newborn-Birth Indicators Research Tracking in Hospitals' study
EPMM	Ending Preventable Maternal Mortality
HIS	Health Information Systems
HMIS	Health Management Information Systems
Icddr,b	International Centre for Diarrheal Disease Research in Bangladesh
BD	Bangladesh
IHI	Ifakara Health Institute in Tanzania
KMC	Kangaroo Mother Care
LBW	Low Birthweight
LMIC	Low- and Middle-Income Countries
LSHTM	London School of Hygiene &Tropical Medicine in UK
MICS	Multiple Indicator Cluster Surveys
MMR	Maternal Mortality Ratio
MUHAS	Muhimbili University of Health and Allied Sciences
NMR	Neonatal Mortality Rate
NP	Nepal
PI	Principal Investigator
PPH	Post-partum haemorrhage
PRISM	Performance of Routine Information System Management
RHIS	Routine Health Information Systems
SBR	Stillbirth Rate
TZ	Tanzania
UNICEF	United Nations International Children's Emergency Fund
WHO	World Health Organization

Glossary

Operational definitions for the characteristics of data quality

Dimension of data quality	Description and notes
Accessibility of data	Data are available to authorized person when and where needed.
Accuracy of data	'Every Newborn-Birth Indicators Research Tracking in Hospitals' EN-BIRTH protocol paper definition[1]: Accuracy is the closeness of a measured value to a standard value – in the context of this study the standard was observation of practice/ intervention).[1] Other definitions: Data faithfully reflect the actual level of service delivery that was conducted in the health facility.[8] Accurate data are considered correct: the data measure what they are intended to measure. Accurate data minimize errors to a point of being negligible. In the context of data quality, accuracy means that data transmitted from one level of the health system to another match. For example, data in client records at the facility level should match data in the monthly health information system (HIS) summary report transmitted to the district level.
Behavioural determinants of RHIS performance	Factors affecting health information systems (HIS) performance that are related to individual behaviour of health professionals and data professionals such as motivation, attitude, empowerment, and confidence.[6]
Bias	Any effect during the collection or interpretation of information that leads to a systematic error in one direction.
Completeness of information systems	In the context of data quality, completeness means the degree to which HIS data: (1) cover all geographical areas, services, and facilities (2) are filled out in full in data collection process documents including forms and registers.
(1) Completeness of geography, services, facilities	The results are derived from a system that is appropriately inclusive: it represents the complete list of eligible persons or units and not just a fraction of the list. *(e.g., answers the question:* *Are the data complete enough to determine whether the health programme is effective and is achieving the desired results?)*
(2) Completeness of data collection process documents	A measure of the proportion of entries (data elements/ indicators) in data collection process documents including forms and register that had any data recorded for the specified data element/ indicators.[8]

Confidence in RHIS tasks	How comfortable a person feels performing a certain RHIS task (e.g., data capture/ collection, data transmission, data process, data analysis, data quality check) competently.
Coverage indicator	Number of individuals receiving an intervention or service (numerator), from among the population in need of the intervention or service (denominator). For institutional births the population in need is defined as the women/ babies in the health facility.
Culture of information	The promotion of a culture of information is operationally defined as an organization having the capacity and control to promote values and beliefs among its members to promote collection, analysis and use of information to accomplish its goals and mission. For assessing whether health facilities promote a culture of information, the construct is operationalized as having dimensions that include the promotion of: a) commitment and support for high quality data b) commitment and support of information use c) evidence-based decision-making culture d) promotion of problem-solving culture e) sharing information between level f) sense of responsibility g) empowerment and accountability h) rewarding good performance[6,9]
Data	An unprocessed set of values of qualitative or quantitative variables. (Note: “information” and “data” tend to be used inter-changeably in policy documents and I mirror this in my thesis by using the term “data-information”).
Data demand and use model	This model for understanding health information systems (HIS) performance examines the data use cycle, from demand for information to data collection and availability, to the use of information, and then to feedback, which in turn increases the demand for information.[10]
Data-information	Term used in these to capture the concept of “information” and “data” being used inter-changeably in policy documents.
Data-information Pyramid	A schematic way of looking at the number of data items to be collected at each level of the health system allowing each level to gather data of importance and relevance to their daily work while avoiding excessive data where no action is taken. The pyramid illustrates how most data are collected at the base of the pyramid in the health facility, where most health service action takes place. Data are processed, filtered and streamlined as data sets are passed up the health system.[11]

Data quality	Data element level: Quantifies problems of data completeness, timeliness, consistency and accuracy in order to ascertain the extent to which the health-facility data are fit for purpose.[8] Health Information systems level: The degree to which health information systems (HIS) data are accurate, timely, complete, and relevant.
Evaluation	An assessment of whether a program's objectives have been achieved.
Evidence-based (or evidence-informed) decision making	An approach based on the use of reliable quantitative and qualitative information to guide decisions about the efficient targeting of resources.
External consistency of data	An assessment of the level of agreement between two sources of data measuring the same health indicator is. The two sources of data that are usually compared are data flowing through the HMIS or the programme-specific information system and data from a periodic population-based survey. The HMIS data can also be compared to pharmacy records or other types of data to ensure that the two sources fall within a similar range.[8]
Health Information system (HIS)	A system that integrates data collection, processing, reporting, and use of the information necessary for improving health service effectiveness and efficiency through better management at all levels of health services.[12]
Health Management Information System	An information system specially designed to assist in the management and planning of health programmes, as opposed to delivery of care.[12] A system that ensures the production, analysis, dissemination, and use of reliable and timely information on health determinants, health system performance, and health status.[13]
Health system	All actors, institutions, and resources that undertake health actions (i.e., actions whose primary purpose is to promote, restore, or maintain health). The World Health Organisation describes health systems in terms of six core components or "building blocks": (i) service delivery, (ii) health workforce, (iii) health information systems, (iv) access to essential medicines, (v) financing, and (vi) leadership/governance
Heaping of data	A measure of the proportion of values falling on specific values (e.g., Birthweight on 2000g or 2500g) or rounded (i.e., ending in "00" or "50").
Impact indicator	A measure of the extent to which health status is being achieved (e.g., maternal and newborn mortality); used for global tracking.

Implausibility of data	A measure of whether individual data are outside pre-defined ranges of biological credibility
Indicator	Defined, measurable data indicating progress in the achievement of one or more objectives.
Information	Data that have been processed and interpreted so that they have meaning and can be used for decision making. (Note: "information" and "data" tend to be used inter-changeably in policy documents and I mirror this in my thesis by using the term "data-information").
Information system	A collection of components that work together to the common objective of providing information support to the decision-making process at each level of an organisation.[12]
Information-generating process	A process by which health information systems (HIS) data are transformed into information that is used for decision making. The process includes the steps of defining information needs/indicators, data collection, data transmission, data processing, data analysis, and the management issues affecting this process (resources and organizational rules).
Integrity of data	Data have integrity when the system used to generate them is protected from deliberate bias or manipulation for political or personal reasons.
Internal consistency of data	An assessment of the level of agreement between two variables in the same source documents or document flow. Typically this involves assessing reporting accuracy of selected indicators through the review of source documents in health facilities and district offices. This element of internal consistency is measured by a data verification exercise.[8] *(e.g., Answers the question:* *Is the observed relationship between the indicators, as reflected in the reported data, that which we would expect?)*
Legibility of data	Data are readable.
Meaning or usefulness of data/ information	Information is pertinent and beneficial.
Objective	A specific outcome to be accomplished to achieve a goal. It may be a milestone along the way in the implementation of a strategy.
Organizational determinants of RHIS performance	Factors affecting health information systems (HIS) performance related to environmental or systemic issues, or the context in which the HIS functions. They could include resources, health system structure, roles and responsibilities of personnel, organizational culture, and budget control.[6]
Precision of data elements	Agreement among repeated measurements of the same data element.
Precision of health information systems	Sufficient detail is captured. For example, an indicator requires the number of individuals who received HIV counselling and testing and received their test results by sex of the individual. An information

	system lacks precision if it is not designed to record the sex of the individual who received counselling and testing.
PRISM Framework	Performance of Routine Information System Management (PRISM). A model for understanding the technical, organizational, and behavioural factors that drive health information systems (HIS) performance. The framework helps HIS professionals with needs assessment, strategy planning, and improvement processes.
Proxy indicator	Indicator used to describe a situation, phenomenon, or condition for which no direct information is available.
Reliability of data	Consistency of the data when collected repeatedly using the same procedures and under the same circumstances.[8]
Resource	The input needed to perform a task, such as funds, personnel, infrastructure, or materials.
Routine health information system (RHIS)	Systems that generate data collected at public and private health facilities, institutions and community-level healthcare posts and clinics, at regular intervals of a year at least. The data give a picture of health status, health services, and health resources. Most of the data are gathered by healthcare professionals as they go about their work, by supervisors, and through routine health facility surveys. The sources of those data are generally individual health records, records of services delivered, and resource health records.[14]
RHIS performance	The effectiveness of a Routine Health Information Systems (RHIS), defined in terms of data quality and use of information.
Stakeholder	A person or organization affected in a significant way by the outcome of a process who can in turn affect the outcome of that process.
Strategy	A method, set of activities, and/or process(es) required to achieve a goal.
System	A collection of components that work together to achieve a common objective.[12]
Target	A specific, measurable figure to be achieved for a given indicator, as part of a goal or objective (e.g., 90% on-time reporting rate).
Task	A defined action that is required as part of the implementation of a plan (also "activity").
Technical determinants of RHIS performance	Factors affecting health information systems (HIS) performance that are related to system components, such as indicators, personnel training, technology, forms, data submission, and reporting.[6]
Tension	State of being stretched or extended or in opposition between tasks/ responsibilities (adapted from[15]).
Timeliness of data	In the context of data quality, the degree to which reports are submitted on time according to established deadlines. Health information systems (HIS) data are timely when they are up-to-date (current), and when the information (i.e. processing of data) is available when required to make decisions about the health of the

	population and to target resources to improve health-system coverage, efficiency and quality.[8] *(e.g., Answers the question:* *Are the data sufficiently recent that achievement or gaps indicated by the data actually reflect the current level of achievement of health indicators?)*
Traction	The support or interest that is needed for something to make progress or succeed.[15]
Use of information	When a decision maker is explicitly aware of a decision and alternatives and considers relevant information in the process of decision-making.
Utility	The transformation of count data into indicators by using them as numerators and denominators or cross-tabulation.
Validity of data element/ indicator	The extent to which a measurement or test accurately measures what is intended to be.
Validation	EN-BIRTH protocol paper definition[1]: The process whereby the ability of health indicators to measure what they are supposed to measure is determined. Criterion "validity testing" assessment compares measurement against an objective gold standard to assess if indicators accurately measure what they intend to.[16]

Glossary definitions from references [17-19]

Operational definitions for the characteristics of quality of care

Dimension of quality of care	Description and notes
Safe	Delivering health care which minimises risks and harm to service users, including avoiding preventable injuries and reducing medical errors.
Effective	Providing services based on scientific knowledge and evidence-based guidelines.
Timely	Reducing delays in providing/receiving health care.
Efficient	Delivering health care in a manner which maximises resource use and avoids wastage.
Equitable	Delivering health care which does not vary in quality because of personal characteristics such as gender, race, ethnicity, geographical location, or socioeconomic status.
People-centred	Providing care which considers the preferences and aspirations of individual service users and the cultures of their communities.

Glossary definitions from Reference:[5,20,21]

Chapter 1 – Aims and Objectives of the PhD

Background

Most preventable deaths of newborns. stillbirths and women occur in settings with data gaps that impede progress for coverage, equity, and quality of care. With rising facility births, there is increasing interest regarding routine health facility data to track intrapartum care provided for women and newborns, especially as this is the time during pregnancy contributing most to mortality. Routine health information systems (RHIS) typically use aggregate data documented by health professionals in health facility routine registers, the quality of data from routine labour and delivery registers is currently understudied. Improving data for action is also a focus of recent global movements for women, newborns, and stillbirths.

Aim

This PhD thesis aims to explore hospital labour and delivery routine register data quality and opportunities to improve measurement of newborn indicators in high mortality settings.

Objectives

The overall aim of my PhD thesis will be addressed by six objectives covering two main sections:

Section A – Assessing existing labour and delivery routine register data quality for hospital births

- Objective 1 – To describe the protocol for the 'Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) observational study for coverage and quality of maternal and newborn health care in Bangladesh, Nepal, and Tanzania.
- Objective 2 – To assess labour and delivery ward baseline routine register data availability, quality, and utility in five EN-BIRTH study hospitals.
- Objective 3 – To assess validity of coverage indicator measurement of newborn and maternal health care coverage in EN-BIRTH study hospitals.

Section B – Identifying opportunities to improve labour and delivery routine register data quality for hospital births

- Objective 4 – To explore barriers and enablers for health professionals to record high quality data for newborn and maternal health indicator measurement from labour ward routine registers in five EN-BIRTH study hospitals.
- Objective 5 – To assess routine birthweight in EN-BIRTH study hospitals: accuracy, gaps and opportunities to measure coverage and quality of care.
- Objective 6 – To assess measurement opportunities for neonatal resuscitation: indicator definitions and quality of care.

Positionality Statement

This thesis has been written in my role as an academic researcher at the London School of Hygiene & Tropical Medicine (LSHTM), UK and I include a positionality statement for this mixed-methods research.[22]

I am a UK trained paediatrician and obstetrician and was privileged to live in rural south Asia for 15 years working as a clinician scientist across the mother-newborn-child-adolescent health and nutrition continuum of care. In November 2017 I joined the LSHTM to manage the multi-country EN-BIRTH study.

This PhD thesis has enabled me to apply a dual perspective as an insider/outsider to this topic, now as a researcher and previously as a practitioner in the types of settings which are the focus on my research.

My personal interest in maternal and neonatal hospital data availability in high-mortality settings began in 2003 as a "a story of survival". Transitioning from UK clinical practice in obstetrics/ paediatrics to rural south Asia, I found data were often not available for day-by-day clinical decisions and quality improvement including clinical audit and regular maternal and perinatal death audit meetings. Whether high quality data was available impacted a woman and baby's survival as well as contributing to health professional stress needing to make clinical decisions with insufficient information. As a rural hospital we faced the typical limited personnel, time and funds and seemingly unlimited monthly reporting requirements for hospital management, district/national RHIS systems and individual programme requirements. In response, between 2004-2017, I co-created an actionable hospital information system to enable data for clinical use, audit and research.[23]

From January 2020, I was the LSHTM principal investigator for **Advancing Routine Health Information Systems (RHIS) to Deliver for Every Newborn - EN-BIRTH 2 study** which contributes content in my PhD thesis introduction and general discussion chapters. I collaboratively designed the EN-BIRTH 2 study, drafted the concept note, designed the conceptual framework, and co-led the global protocol and LSHTM ethics submission. I conceptualised the current major output - **Every Newborn-Measurement Improvement for Newborn & Stillbirth Indicators (EN-MINI) Tools for Routine Health Information Systems**, co-designed the EN-MINI-PRISM Analysis Tool and drafted the results report from Tanzania (Appendix 4).

Outline of Work

My PhD thesis brings together my work completed as part of the '**Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) study** (Figure 1). The study was conducted in five comprehensive Emergency Obstetric and Newborn hospitals in Bangladesh, Nepal, and Tanzania. The principal investigators were Professor Joy Lawn (London School of Hygiene & Tropical Medicine (LSHTM), Dr Shams El Arifeen (icddr,b Bangladesh), Dr Ashish KC (Golden Community, Nepal) and Dr Honorati Masanja (Ifakara Health Institute, Tanzania). I joined the LSHTM in November 2017, mid-way through EN-BIRTH data collection and thereafter managed the multi-country study.

Ethical approval for this research is shown in Appendix 1.

Figure 1: Panel Overview of EN-BIRTH study

'Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) study

The EN-BIRTH study aimed to assess validity of measurement of selected newborn and maternal care health intervention indicators (coverage/ quality aspects and/or safety) in hospitals to inform routine use in national HMIS and global tracking [1].

The EN-BIRTH study had 4 objectives:

Objective 1 – **Numerators:** To determine validity (accuracy) of both routine hospital register and women's report exit surveys, compared to direct observation for selected maternal and newborn care interventions: uterotonics for post-partum haemorrhage (PPH) prevention, early initiation of breastfeeding, neonatal bag-mask ventilation, Kangaroo Mother Care (KMC); and, verification with individual case records, antibiotic treatment for neonatal infection.

Objective 2 – **Denominators:** To compare different denominator options including proxies, and assess feasibility of their use in routine data platforms, including:
-Target population requiring intervention (clinical need) in the hospital ("true" denominator).
-Live births in the hospital.
-Total births (live births and stillbirths) in the hospital.
-Estimated population births (live or total): hospital births and home births.

Objective 3 – **Content /quality of care:** To evaluate different domains of coverage (e.g., timing, completion rates, safety) for selected interventions.

Objective 4 – **Barriers and enablers:** To evaluate barriers and enablers to routine recording of selected indicators, and to explore perceived utility of these data to improve decision-making, healthcare coverage, and quality of care at all levels.

Table 1 outlines the six objectives of my PhD thesis and my specific contribution to each research chapter: protocol paper (PhD objective 1), baseline register analysis (PhD objective 2), overall results paper (PhD objective 3), barriers and enablers to routine register recording (PhD objective 4), birthweight measurement (PhD objective 5), neonatal resuscitation measurement (PhD objective 6).

Details of the six publications included in my PhD thesis are listed in Appendix 2 and my additional contributions to manuscripts in the EN-BIRTH supplement in Appendix 3.

Table 1: Outline of work structured in two sections and six objectives, my contribution, status, outputs and progress

Objectives:		Status and outputs completed	Personal share in the investigation by candidate Louise Tina Day
Section A – Existing labour and delivery routine register data quality for hospital births			
1.	To describe the protocol for the 'Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) observational study for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania.	Published co-first author paper: **Day LT, Ruysen H**, et al. "Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. *Journal of Global Health* 2019; **9**(1). https://doi.org/10.7189/jogh.09.010902 [1]	• I drafted the first full manuscript, designed data flow figure (5), incorporated collaborative inputs from PIs, country teams, national/global expert advisory groups. • I submitted the manuscript on behalf of the authors and corresponding PI, revised the manuscript in response to peer-review, including co-ordination of collaborative inputs from co-authors. • I co-managed proofs with my joint first author.
2.	To assess labour and delivery ward baseline routine register data availability, quality, and utility in five EN-BIRTH study hospitals.	Published co-first author paper: **Day LT, Gore-Langton GR**, et al Labour and delivery ward register data availability, quality, and utility – Every Newborn – birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries. *BMC Health Serv Res* 2020, **20**(1):737. https://doi.org/10.1186/s12913-020-5028-7 [24]	• I jointly designed the objectives with the PI and co-first author. • I designed and refined matrices for routine labour ward registers. • I designed data cleaning and analysis plan, my co-first author colleague ran the statistical analysis under my leadership. • I jointly drafted the manuscript with my co-first author. In particular, I led on results, discussion and conclusion. • I am the corresponding author and jointly revised the manuscript in response to peer-review.

Objectives:		Status and outputs completed	Personal share in the investigation by candidate Louise Tina Day
3.	To assess validity of measurement of newborn and maternal health-care coverage in EN-BIRTH study hospitals.	Published co-first author paper: **Day LT, Rahman QS**, et al. Assessment of the validity of the measurement of newborn and maternal health-care coverage in hospitals (EN-BIRTH): an observational study. *The Lancet Global Health* 2021; **9**(3): E267-79. https://doi.org/10.1016/S2214-109X(20)30504-0.[25]	• I managed the collaborative multi-country/site study from mid-way through data collection. Specific contributions included designing the duplicate observation/ extraction process and the variable renaming and collated the code book. • I jointly designed the analysis plan with the PIs and co-authors. I co-led the data cleaning process with my co-first author. I designed the variable matrix for population-level coverage comparison and individual-level validity testing. The statistical analysis was run by my joint first author. • I conceptualised the additional research questions to stratify by mode of birth (all sites) and the natural experiment of original and revised registers for the Bangladesh study hospitals. • I drafted the manuscript, conceptualised the validity ratios and designed all figures. • I incorporated collaborative inputs from PIs, country teams, national/global expert advisory groups. • I am the corresponding author, lead on the two rounds of manuscript revision in response to peer-review, including co-ordinating collaborative inputs from co-authors. • I managed proofs and presentation of results at dissemination activities. • I presented results at dissemination activities including preliminary results at MoNITOR, study results launch.

Objectives:		Status and outputs completed	Personal share in the investigation by candidate Louise Tina Day
Section B – Opportunities to improve labour and delivery routine register data quality for hospital births			
4.	To investigate the barriers and enablers for health professionals to hospital routine register documentation of coverage of care indicators for women and newborns in five EN-BIRTH study hospitals.	Published co-first author paper: **Shamba D, Day LT**, Zaman SB, et al. Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth*, 2021; **21**(Suppl 1): 233. https://doi.org/10.1186/s12884-020-03517-3.[26]	• I jointly co-ordinated multi-country technical working group regular multi-site meetings. • I jointly led collaborative design of data collection tools, thematic analyses and synthesis. • I led the routine register design synthesis and the main results figure 1. • I jointly drafted the manuscript with my co-first author. • I am the corresponding author, led on the manuscript revision in response to peer-review, including co-ordinating collaborative inputs from co-authors. • I managed proofs and co-presented results at dissemination activities.
5.	To assess routine birthweight in EN-BIRTH study hospitals: accuracy, gaps and opportunities to measure coverage and quality of care.	Published co-first author paper: **Kong S, Day LT**, Zaman SB, et al. Birthweight: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* 2021; **21**(Suppl 1): 240. https://doi.org/10.1186/s12884-020-03355-3.[27]	• I jointly designed the objectives with my co-first author. • I co-designed data cleaning and analysis plan working with my co-first author colleague who ran the statistical analysis under my leadership. • I jointly drafted the manuscript whilst mentoring co-first author. I led on results, discussion and conclusion. • I am the corresponding author, lead on the manuscript revision in response to peer-review, including co-ordinating collaborative inputs from co-authors.

Objectives:		Status and outputs completed	Personal share in the investigation by candidate Louise Tina Day
6	To assess measurement opportunities for neonatal resuscitation: indicator definitions and quality of care.	Published co-senior author paper: Kc, A., Peven, K., Ameen, S., Msemo, G., Basnet, O., Ruysen, H., Zaman, S. B., Mkony, M., Sunny, A. K., Rahman, Q. S., Shabani, J., Bastola, R. C., Assenga, E., Kc, N. P., El Arifeen, S., Kija, E., Malla, H., Kong, S., Singhal, N., Niermeyer, S., Lincetto, O., **Day, L. T., Lawn, J. E**. and EN-BIRTH Study Group. Neonatal resuscitation: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth.* Vol. 21, 2021/03/27 edn; 2021: 235. https://doi.org/10.1186/s12884-020-03422-9.[28]	• I jointly designed the objectives with the co-first authors and co-senior authors. • I co-designed data cleaning and analysis plan linked to my knowledge and experience of neonatal resuscitation algorithms, working with my co-first author colleagues who ran the statistical analysis. • I made substantial contributions to the drafted manuscript especially the background, results, discussion and conclusion whilst mentoring my colleague one of the co-first authors, Kimberly Peven. • I made substantial contributions to the paper revisions in response to peer-review.

Chapter 2 – Background and development of conceptual framework

The wellbeing of the worlds' smallest and most powerless human beings reflects our humanity for all global citizens. Every newborn has the right to survive and thrive, yet each year an estimated 4.2 million die as newborns and stillbirths; nearly all (98%) are in low- and middle-income countries (LMIC).[29-31] The first global endorsement of national targets for newborns and stillbirths was outlined in the Every Newborn Action Plan (ENAP) at the World Health Assembly in 2014.[32] The Global Strategy for Women's, Children's and Adolescents' Health (Global Strategy), 2016-2030 adopted these targets but only the newborn target was included in the Sustainable Development Goals (SDG) as Target 3.2.[33,34] In 2021, the global neonatal mortality rate (NMR) was estimated at 18/1000 live births.[35] In 2022, the global average stillbirth rate (SBR) was estimated at 14/ 1000 total births.[36] Both NMR and SBR range widely between geographies and many countries are off track for the NMR target of fewer than 12/1000 live births and the SBR target of fewer than 12/1000 total births.[37]

Maternal deaths also continue at unacceptably high levels, most often in the same geographies where stillbirth and neonatal mortality rates are high.[38] Each year an estimated 0.29 million women die and projections indicate nearly a ten-fold increase in annual rate of reduction is needed to meet the global SDG target of 70/100,000 live births.[37] The Ending Preventable Maternal Mortality (EPMM) and ENAP movements have recently come together and published their first joint report: *Improving maternal and newborn health and survival and reduce stillbirth.*[37]

Timely accurate data are needed to track the progress of accelerating action to meet these ambitious goals towards ending preventable stillbirths, newborn and maternal deaths.[32-34] Yet the settings with the highest burden of mortality have the least data on coverage and quality of care – known as the "inverse data law".[39] Key global initiatives for women, newborns and stillbirths acknowledge the importance of improving data availability, quality and use for action to achieve universal health coverage (UHC).[32,34,40] These data initiatives include:

- SDG 17 "Revitalize the global partnership for sustainable development" which states a specific target to increase the availability of high quality, timely and reliable data;[41]
- ENAP fifth objective strategy to "Count Every Newborn";[32]
- Global Strategy stated focus includes producing disaggregated data age, sex, socioeconomic status and other dimensions to ensure that no one is left behind, including in humanitarian and other fragile settings.[34]

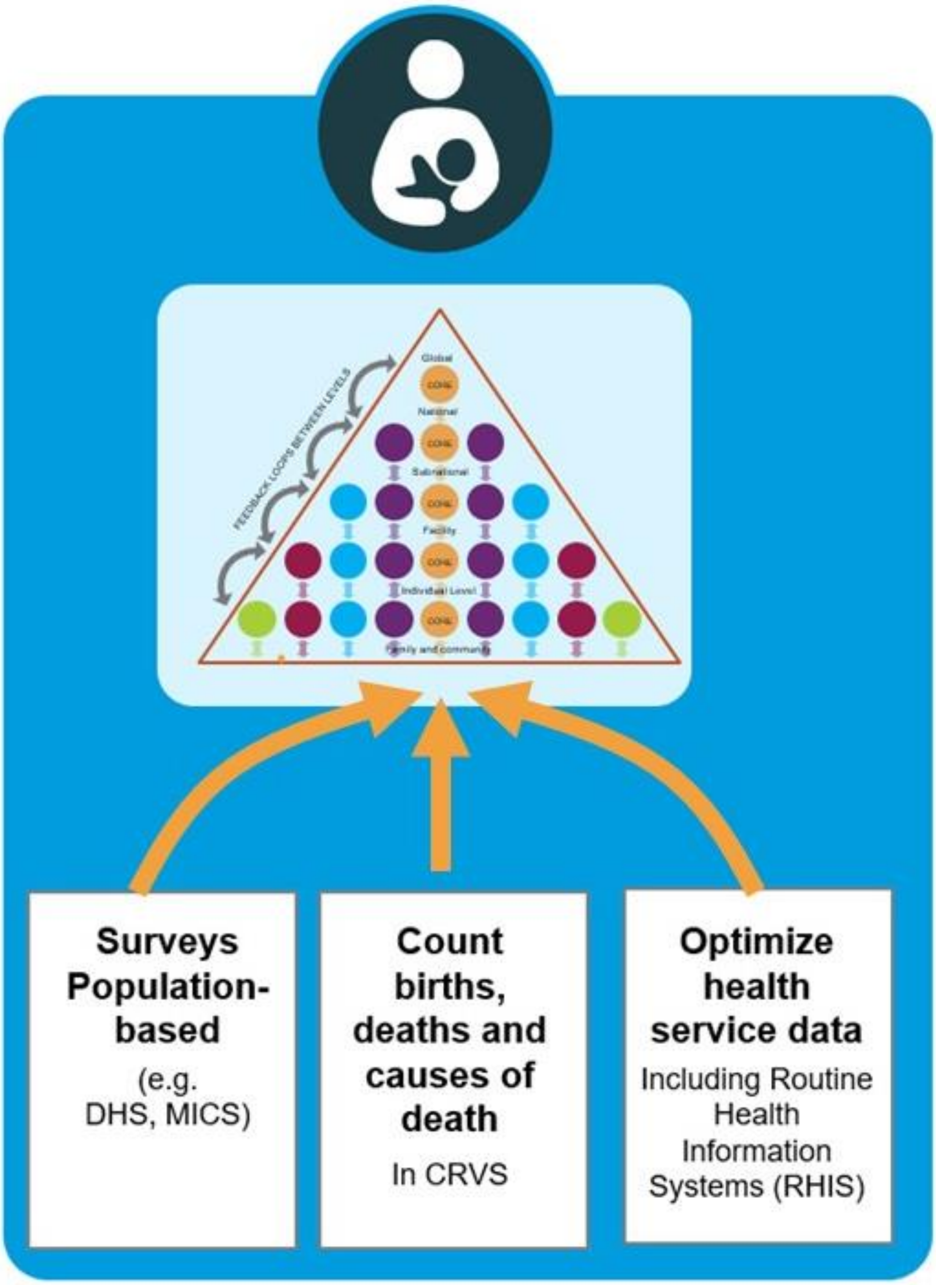
Settings with high mortality for the maternal-newborn dyad are also the settings with high morbidity. Neonatal disorders remain the top cause and maternal disorders remain among the 30 leading causes of global disability adjusted life years (DALYS).[42] Improving data for action is needed to end preventable mortality and morbidity as articulated in the Survive and Thrive agenda.[34] , When "high mortality" is used throughout this thesis, it implies both high mortality and high morbidity.

The aspirations to improve maternal and newborn health measurement in global initiatives (e.g., ENAP, the SDGs and the Global Strategy) tend to use the terms "data" and "information" interchangeably. Elsewhere data and information are distinguished: "data" as an unprocessed set of values of qualitative or quantitative variables and "information" as data that have been processed and interpreted so that they have meaning and can be used for decision making.[43,44] For my PhD thesis I mirror the pattern of using the terms interchangeably and often use a combined term "data-information".

2.1 Indicators and the information-data pyramid

Health information systems (HIS) exist to produce relevant and quality information to support health interventions.[45] Newborn HIS data sources include civil registration and vital statistics (CRVS), census, population-based surveys, and routine data generated from health facilities and administrative information systems (Figure 2).

Figure 2: Data sources for the maternal and newborn data-information pyramid



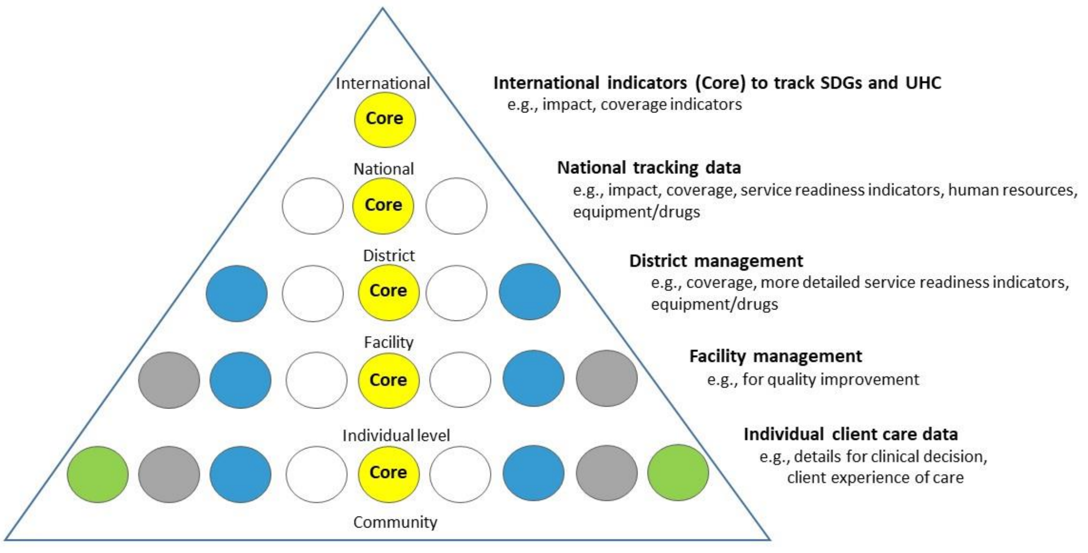
Key: DHS = The Demographic Health Survey[2], MICS = Multiple Indicator Cluster Surveys[4], CRVS = Civil Registration and Vital Statistics, RHIS = Routine Health Information Systems.

Among the six health system building blocks, the HIS is central to the other five: service delivery, health workforce, essential medicines/technology, finance, and leadership/governance.[46] Multiple initiatives over recent decades have highlighted the urgency to strengthen the HIS, including for surveillance and early warning of a potential global pandemic.[45] The COVID-19 pandemic re-emphasized the importance of timely, reliable and actionable HIS data, and further accelerated efforts for standardisation of HIS for Reproductive Maternal Newborn Child and Adolescent Health.[47]

Indicators are standardized, quantitative measures that provide critical information to monitor performance, measure achievement and determine accountability. Indicators require defined components: title, definition, purpose, rationale, method of measurement, numerator, denominator, data collection method (tool/ frequency), suggestions for data disaggregation, and guidelines to interpret and use data.[48] Indicator measurement is only valuable if the data are valid, reliable and non-biased.[48]

Health indicators form an important part of the actionable information system and allow for comparisons over time and across geographies. The data-information pyramid is a schematic way of looking at the number of data items to be collected at each level of the health system allowing each level to gather data of importance and relevance to their daily work, while avoiding excessive data where no action is taken (Figure 3). The pyramid illustrates how most data are collected at the base of the pyramid, in the health facility, where most health service action takes place. Data are processed, filtered and streamlined as data sets are passed up the health system.[11]

Figure 3: Health System Data-Information pyramid



Source: Figure 3 in "Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal, and Tanzania. [1]

Applying this data-information pyramid visualisation for women and newborn indicators emphasises the importance of the right data reaching the appropriate level of the health system for use and action. Many types of health data for each individual woman and newborn are generated at the base of the pyramid (shown in Figure 3 as green, grey, blue, white circles) with a smaller subset of relevant core data-information rising to the higher levels (shown in Figure 3 as yellow circles).

Typical maternal and newborn indicator users at different levels of the health system include health care professionals, sub-national and national level programme managers, and policymakers. Less frequently mentioned as data users are the parents, families, caregivers, and community surrounding the woman and newborn.

At the point of service delivery, individual-level clinical information (or data) are used by health professionals to inform clinical decisions e.g., newborn temperature, daily weight gain. At facility level, aggregate data are collated to inform administrative and managerial decisions for maternal and newborn planning, budgeting and procurement of equipment and drugs as well as local quality improvement including survival/ mortality audit. Similarly, at subnational level, data are required for planning including human resources for health, budgeting, and procurement. At the national and global level, a few standardised measures or "core indicators" are tracked for accountability, shown in yellow at the centre of the data-information pyramid.

These opening paragraphs of my PhD thesis have discussed routine measurement around the time of birth for women, newborns, and stillbirths as these use similar processes in high mortality settings. The word "newborn" in the Every Newborn Action Plan (ENAP) has a dual meaning – for live births and stillbirths.[32] The main focus of my PhD is newborn and stillbirth core and additional indicators which was the topic of the EN-BIRTH study and maternal health indicators are included to a lesser extent. Throughout this PhD I will mainly describe newborns and stillbirth indicators measurement, but at times link to maternal indicators, to emphasise the centrality of the maternal-fetal dyad.

2.2 Core and additional newborn and stillbirth indicators

WHO maintains a list of 100 core health indicators, prioritised to monitor health trends towards the SDGs and UHC.[49] Among these, only five indicators are specific for newborns and stillbirths: neonatal mortality rate, stillbirth rate, low birthweight rate, early initiation of breastfeeding and postnatal care coverage - newborn. A wider list of core and additional newborn and stillbirth indicators were prioritised by the ENAP (Figure 4).[32,50,51]

Figure 4: Every Newborn Action Plan (ENAP) Core and additional indicators

Current Status		**Core Indicators**	**Additional indicators**
Definitions clear – but quantity and consistency of data lacking	***Impact***	1. Maternal mortality ratio* 2. Stillbirth rate* 3. Neonatal mortality rate*	Intrapartum stillbirth rate Low birthweight rate Preterm birth rate Small for gestational age Neonatal morbidity rates Disability after neonatal conditions
Contact point definitions clear but data on content of care are lacking	***Coverage:*** **Care for All Mothers and Newborns**	4. Skilled attendant at birth* 5. Early postnatal care for mothers and babies* 6. Essential newborn care (tracer is early breastfeeding*)	Antenatal Care* Exclusive breastfeeding up to 6 months*
Gaps in coverage definitions, and requiring validation and feasibility testing for RHIS use	***Coverage:*** **Complications and Extra Care**	7. Neonatal resuscitation 8. Kangaroo mother care 9. Treatment of serious neonatal infections 10. Antenatal corticosteroid use	Caesarean section rate Chlorhexidine cord cleansing
	Input: **Service Delivery Packages for Quality of Care**	Emergency Obstetric Care Care of Small and Sick Newborns Every Mother Every Newborn Quality Initiative with measurable norms and standards	
	Input: **Counting**	Birth Registration	Death registration, cause of death

**Core indicator for 100 Core health indicators and/or Sustainable Develop Goals*

Source: Figure 1 in "Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania [1]

These ENAP indicators were categorised in the newborn measurement improvement roadmap into three groups (highlighted in green, yellow and red in Figure 4) according to their "current measurement status" in 2014.[51] The EN-BIRTH study was designed to address the evidence gap regarding validity of coverage indicator measurement in health facilities – highlighted in red in Figure 4.

The intention of the ENAP coverage indicators (numbered 7, 8, 9 and10 in Figure 4) is to track cost-effective clinical interventions to reduce newborn mortality and morbidity. Evidence for the potential impact of such interventions has been available for decades, yet global scale-up has varied.[52] Uptake of Kangaroo Mother Care (KMC) for low birthweight babies has been slow, including in high-income settings. [53-55] By contrast, dissemination of neonatal resuscitation has accelerated over the last decade with focused initiatives such as the 'Helping Babies Breathe' collaborative partnership.[56,57] However, the actual service coverage for neonatal interventions including KMC and bag-mask ventilation in high mortality settings, remains uncertain due lack of high quality routine data.[58]

My PhD thesis focuses on health facility data quality for a subset of these core and additional ENAP coverage indicators, specifically those captured in the labour and delivery register. Labour and delivery registers are typically found on childbirth ward, capturing data for stillborn babies, livebirths and women which are aggregated for RHIS use. By contrast, postnatal and neonatal ward registers availability and RHIS use are more variable. Therefore, I chose to focus this PhD thesis exploring source RHIS data in the labour and delivery ward register because of potential for generalisability.

2.3 Maternal and Newborn data availability

Population-based household survey data, e.g., Demographic Health Survey (DHS) and Multiple Indicator Cluster Surveys (MICS) have been the main source for tracking maternal and newborn health indicators in LMIC since the Millennium Development Goals (MDGs) era.[2] The potential of RHIS health-facility data is being realised during the current SDG era for the measurement of prioritised newborn indicators for four reasons[59,60]. First, the proportion of births occurring in health facility birth rates have risen in recent decades to 2023 estimates of: 78% globally, south Asia 81% and Africa 58-65%. Thus more newborn and maternal data are potentially available for use around the time of birth.[61] Second, inpatient hospital care for small and/or sick newborns is being expanded in many settings again with implications of increased data availability for this subset of babies.[62,63] Third, investments in RHIS through other disease areas that touched on maternal and newborn health e.g., HIV/AIDS. Finally, open-source digital data platforms are increasing in higher mortality settings e.g., District Health Information Software 2 (DHIS2) is now being used by more than 130 countries.[64] The synergy of increasing childbirth and newborn care in health facilities, coupled with advancing RHIS and digital platforms could transform routine newborn data availability for use in settings that currently have the largest data gaps.

Newborn indicator data availability presently lags behind other health indicators as illustrated in the *2020 global report of health data systems and capacity*. [7] Among 133 countries measuring health facility-based indicators, only 74% report neonatal low birthweight prevalence nationally, reducing to 51% sub-nationally and only 25% disaggregated by sex. In comparison, for immunization programmes which run vertically with funding allocated dependent on coverage rates, data for childhood immunisation rates for Diphtheria, Tetanus, Pertussis, Hepatitis B and Hib vaccines (DTP/Penta 3) are reported by 97% of these 133 countries nationally and 83% sub-nationally. [7]

2.4 Newborn data quality and use

RHIS newborn data availability is necessary but not sufficient. For the data to be meaningful there must be trust in its quality. Previous publications on newborn health indicators captured via facility reporting have raised concerns regarding poor data quality for completeness, timeliness, and accuracy.[65,66]

The Performance Information System Management (PRISM) framework elegantly depicts determinants and causal pathways of RHIS performance that are relevant for newborn indicator data (Figure 5).[6] The PRISM framework shows the connection between **RHIS inputs** and **RHIS outputs** of high quality data for use to improve health system performance and health outcomes. For RHIS input determinants, this framework acknowledges three categories of determinants:

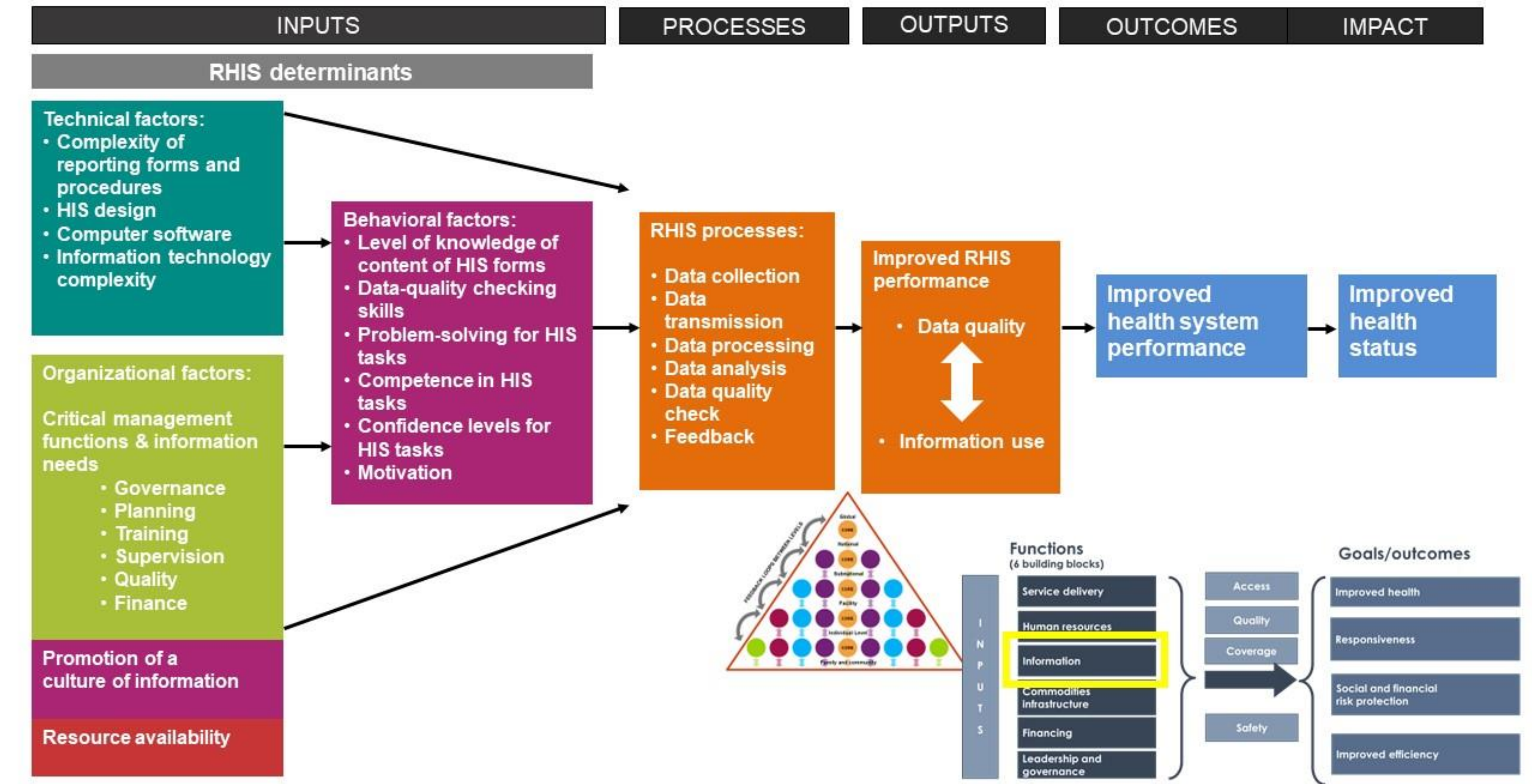
• *Technical determinants*: Complexity of reporting forms & procedures, HIS design, computer software, information technology complexity.

• *Organizational determinants*: Critical management functions & information needs: Governance, planning, training, supervision, quality, finance, promotion of culture of information, availability of resources. Information culture, structure, resources, roles, and responsibilities of key contributors at each level of the health system.

• *Behavioural determinants*: level of knowledge of content of HIS forms, data quality checking skills, problem solving for HIS tasks, competence in HIS tasks, confidence levels for HIS tasks, motivation, demand.

Global and national investments to advance RHIS in high burden settings have mainly focused on technical and organizational at national and sub-national levels (e.g., digital platforms) in recent years.[67,68] The health facility source input data have received less attention, and specifically the technical, and organizational and behavioural determinants that affect data quality, which is the subject of my PhD thesis.

Figure 5: The Performance of Routine Information System Management (PRISM) - Conceptual Model – adapted to show the data pyramid and the health system[6]

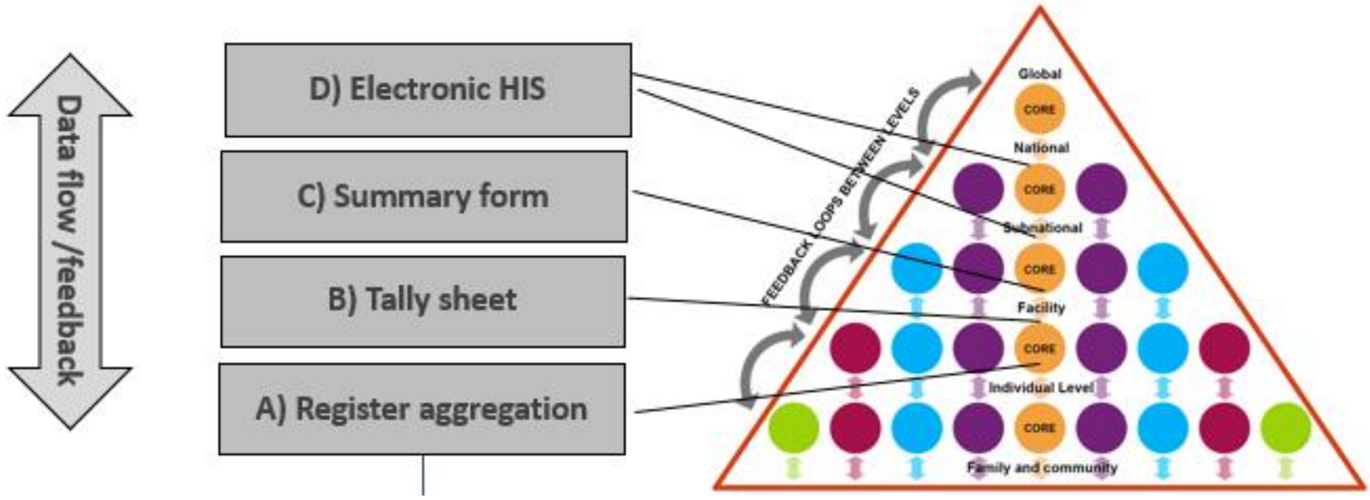


Adapted from: Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan 2009; 24(3): 217-28. Published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed.
Note: Shading reflects original PRISM colours (MEASURE Evaluation. PRISM Training Kit. 2019. https://www.measureevaluation.org/prism (accessed 5 November 2023)). Within the RHIS input determinant "Organisational factors", "Critical management functions and information needs" (shaded green) contribute to the "Promotion of a culture of information" (shaded pink) and "Resource availability" (shaded red). The RHIS input determinant "Behavioural factors" is shaded in the same pink as "Promotion of a culture of information" to illustrate overlapping components.

2.5 Health Facility Routine Data Collection

Data from health facilities typically flow up the data-information pyramid though several steps: A) routine **register aggregation** typically monthly using B) a **tally sheet** and/or C) **Summary form** for entry into D) **Electronic HIS** (Figure 6).

Figure 6: Data aggregation steps for RHIS newborn data transmission up the data-information pyramid



Health facility routine registers (step A) in Figure 6) in high mortality settings are typically paper-based registers each situated in a clinical location e.g., the labour and delivery ward. Health professionals document individual women or newborns in register rows and admission details in the columns, including birth outcomes, clinical care practices and interventions. Routine registers form a parallel system to individual patient case notes which are used by health professionals to document clinical care and progress during admission including ward round notes. Frontline health professionals have the responsibility: to *provide* high quality care, and to *document* that care in registers and case notes. In many settings, health professionals also have the responsibility to *aggregate* routine register data for monthly paper-based reporting.

2.6 Routine register data availability and quality

Limited research has been conducted regarding the accuracy health facility routine register data for newborns and stillbirths in high mortality settings.[69-71] A review maternal and newborn registers in 24 countries found data elements were not consistently available, , labour ward registers were missing and completeness varied.[72] Standardisation is lacking between countries for routine registers across the continuum of antenatal care (ANC), intrapartum care (IPC) and postnatal care (PNC). [73] Improving availability of high quality routine health facility neonatal data is essential, otherwise mistrust in data quality will hamper data use.

Assessments of RHIS data quality typically focus on comparing the summary form (step C) in Figure 6) with the electronic HIS (step D) in Figure 6) generated from the register aggregation (step A) in Figure 6). These assessments measure the data quality dimensions of completeness, timeliness, and internal consistency. In this context, completeness is defined as a measure of the proportion of entries in the registers and summary form that had any data recorded for the specified data element/ indicators.[8] Timeliness is defined as the degree to which reports are submitted on time according to established deadlines. Internal consistency is defined as the level of agreement between the registers as the source documents and the summary forms.

The PRISM conceptual framework can be used to interpret the results of RHIS data quality assessments to plan actions to improve data quality for use.[6,74] Factors known to negatively impact routine data quality in other health sectors include: RHIS outputs (poor use of data), RHIS processes (lack of feedback) and RHIS input determinants: organisational (low management support); behavioural (lack of health professional competence and confidence, low motivation, and low perceived utility of routine recording tasks).[6,75-77] Less is known regarding the RHIS input determinant of technical factors in health facilities in high mortality settings.[78]

2.7 Assessing criterion validity of newborn indicator measurement

Criterion "validity testing" assessment compares measurement against an objective gold standard to assess if indicators accurately measure what they intend to.[16] Prior to the EN-BIRTH study, maternal and newborn indicator measurement criterion validation research had primarily focused on population-based surveys. These assessments had generally shown low accuracy for women's report of measures regarding interventions around the time of birth.[79-82] Limited studies had assessed criterion validity of routine labour ward register data for core/ additional maternal and newborn indicators among which are very few observational studies.[69-71] Thus the validity of measurement for health facility coverage newborn indicators was a priority evidence gap identified by the ENAP measurement improvement roadmap.[51] Understanding routine register data quality measurement for core newborn and maternal indicators could contribute to understanding the existing usefulness of RHIS data, identify areas for improvement, and guide investment.

2.8 The relationship between routine data quality and quality of care

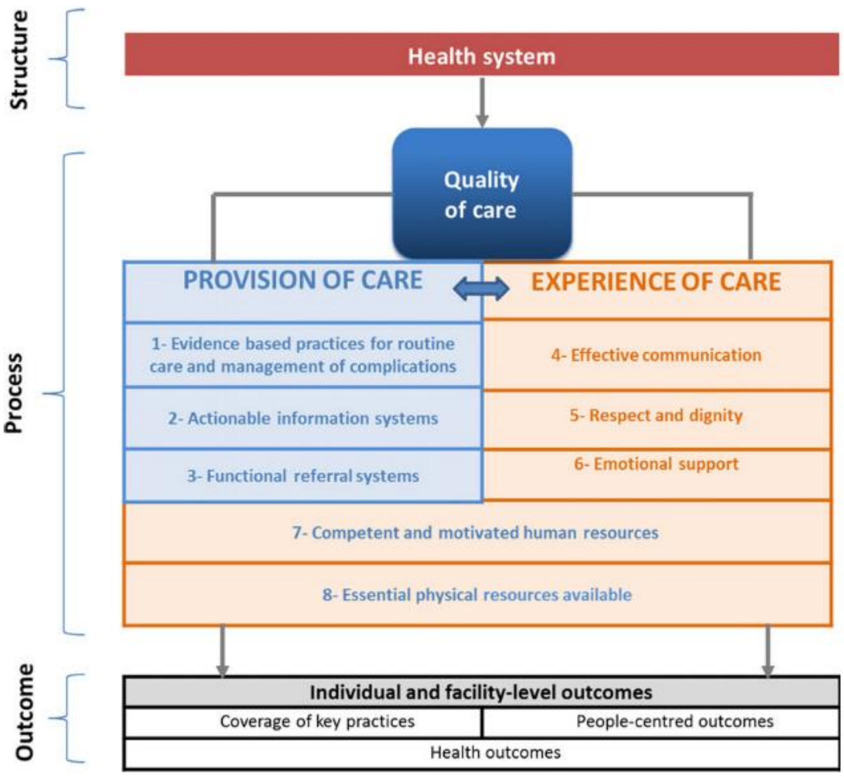
The collection of an abundance of data poorly aligned to specific goals and objectives and/or poorly processed into information (e.g., indicators) has been described as the Data Rich Information Poor (DRIP) syndrome.[83] This generates a high burden across the data-information pyramid (Figure 3) for data collection, processing and reporting affecting data quality and use. DRIP syndrome is common in high mortality settings with multiple programmes, donors, and indicators.[83]

The DRIP syndrome data burden on frontline health professionals will compete with their capacity to provide high quality care to women and newborns. Improving quality of care for women and newborns is a priority, in recognition that poor quality care makes a major contribution to preventable maternal and newborn mortality and stillbirths.[5,84,85]

Quality care for pregnant women and newborns has been articulated by WHO as safe, effective, timely, efficient, equitable and people-centred.[5,20] The linked WHO Quality of Care Framework uses a Donabedian model of **structure, process and outcomes** (Figure 7).[86]

The **structure** is shown as the Health System, previously described with six building blocks: service delivery, health workforce, information systems, access essential medicines, finance, leadership and governance.[46] Quality of care is conceptualised as the **process** with two dimensions PROVISION OF CARE and EXPERIENCE OF CARE, captured by eight quality domains "*1. Evidence-based practices for routine care and management of complications*", "*2. Actionable information systems*", "*3. Functional Referral Systems", "4. Effective communication", "5. Respect and preservation of dignity*", *"6. Emotional support"*, "*7. Competent motivated human resources*", "*8. Essential physical resources available".* The **outcomes** of the framework are individual and facility-level outcomes – coverage of key practices and people-centred outcomes – for pregnant women and newborns.

Figure 7: WHO Quality of Care Framework for maternal and newborn health[5]



This thesis seeks to explore hospital labour and delivery routine register data quality and opportunities to improve measurement of newborn indicators in high mortality settings. These data are used to track coverage of key care practices for the maternal-fetal dyad, shown as **outcomes** in the WHO Quality of Care framework (Figure 7).

I conceptualise my exploration of data quality as a relationship between *quality of data* and *quality of care,* shown in a novel adaptation of the WHO Quality of care framework, the **'Quality of Care and Quality of Data Conceptual Framework'** which evolved during the research for this thesis and now acts as the organising framework (Figure 8).[5] This conceptual framework is introduced in this chapter, and further explored in the integrated discussion of this PhD thesis in Chapter 9.   Figure 8 shows the two sections and six research chapters in my PhD on the left of the framework:

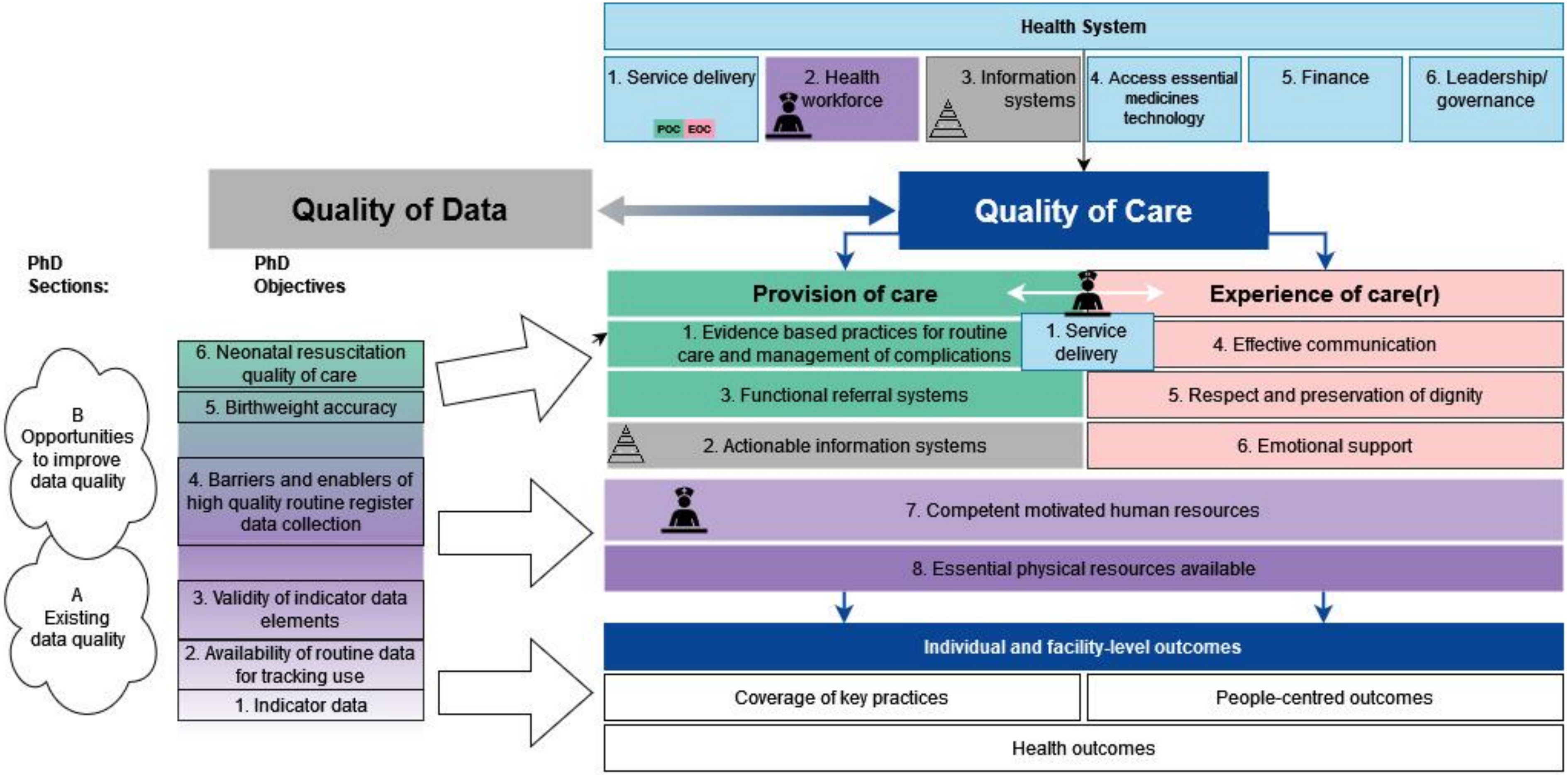
Section A assesses existing labour and delivery routine register data quality for hospital births.

Section B identifies opportunities to improve labour and delivery routine register data quality.

The **'Quality of Care and Quality of Data Conceptual Framework'** links quality of data with quality of care. Specifically, the framework focuses on three of the eight quality domains, namely: evidence-based practices for routine care and management of complications (domain 1), actionable information systems (domain 2), and competent and motivated human resources (domain 7). The relationship between quality of data and quality of care, and specifically the tension between them or the traction that can be developed, is the framing for this thesis. This is grounded by identifying opportunities to improve routine health facility data quality based on the assessment of existing data quality in five hospitals in Bangladesh, Nepal, and Tanzania.

At the start of section A and section B in this thesis I use the '**Quality of Care and Quality of Data Conceptual Framework'** to show how the objectives connect within section A and B (Figure 8).

Figure 8: PhD structure linking to the 'Quality of Care and Quality of Data conceptual framework' – full framework shown in Figure 24



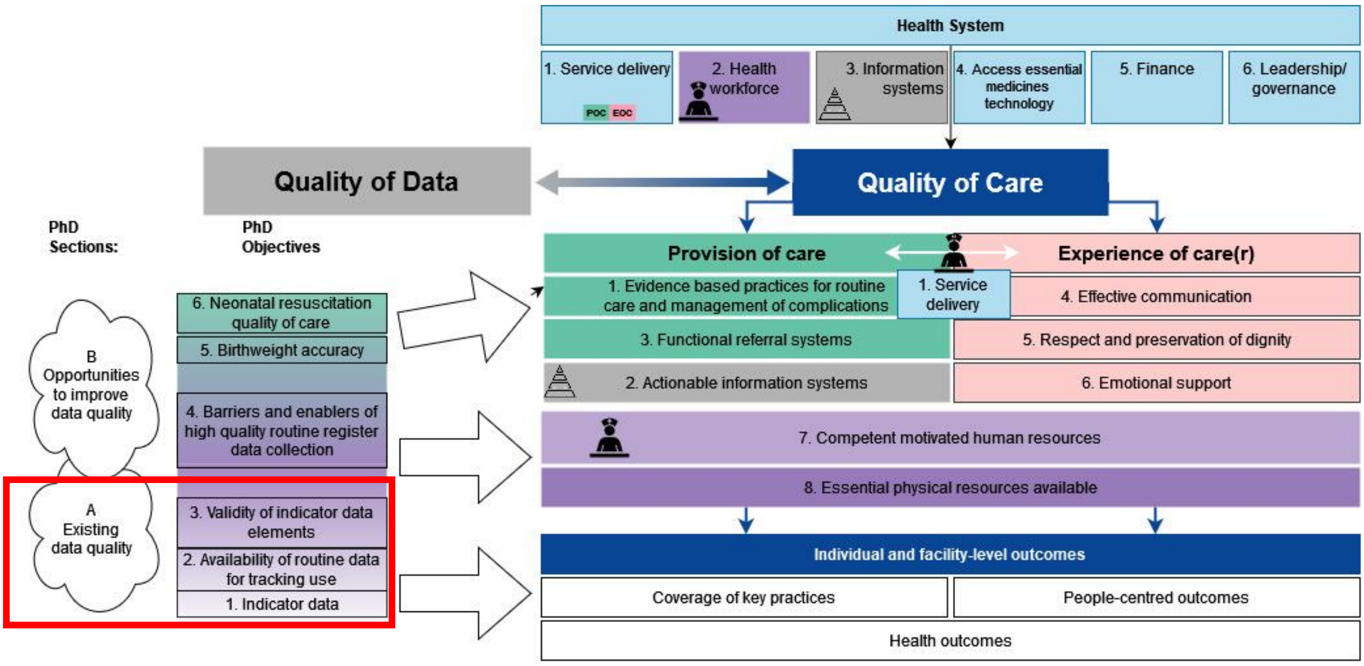
Note: My PhD structure links Quality of data to three of the eight domains of the WHO Quality of Care framework (Tunçalp Ö, Were W, MacLennan C, et al. Quality of care for pregnant women and newborns—the WHO vision. Bjog 2015; 122(8): 1045-9): *Domain 1* - evidence based practices for routine care and management of complications, *Domain 2* - actionable information systems, *Domain 7* - competent and motivated human resources.

SECTION A: Assessing existing labour and delivery routine register data quality for hospital births

Section A – explores objectives 1, 2, and 3 as highlighted in the red box in Figure 9**:**

- Objective 1 – To describe the protocol of the 'Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) observational study for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania.
- Objective 2 – To assess labour and delivery ward baseline routine register data availability, quality, and utility in five EN-BIRTH study hospitals.
- Objective 3 – To assess validity of coverage indicator measurement of newborn and maternal health care coverage in EN-BIRTH study hospitals.

Figure 9: 'Quality of Care and Quality of Data conceptual framework' – highlighting PhD Thesis Section A – Assessing existing labour and delivery routine register data quality for hospital births.



Chapter 3 - Objective 1: Indicator data

This chapter describes the protocol of the 'Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) observational study for coverage and quality of maternal and newborn health care in Bangladesh, Nepal, and Tanzania

The chapter was published in June 2019 in the Journal of Global Health. The manuscript was published under a creative commons license (4.0 International (CC BY 4.0) and no further permission are needed.

The published manuscript is included in full below and the supplementary material referenced in the paper is available at http://jogha.org/documents/issue201901/jogh-09-010902-s001.pdf

3.1 List of Figures

3.2 List of Tables

3.3 Citation

Day LT, Ruysen H, et al.
"Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania.
Journal of Global Health 2019; **9**(1). https://doi.org/10.7189/jogh.09.010902 [1]

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	034282	**Title**	Dr
First Name(s)	Louise Tina		
Surname/Family Name	Day		
Thesis Title	Quality of care and quality of data for hospital births – tension or traction?		
Primary Supervisor	Associate Professor Cally Tann		

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?	"Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. Journal of Global Health 2019; 9(1). https://doi.org/10.7189/jogh.09.010902 1 Day, Louise T., Ruysen, Harriet, Gordeev, Vladimir S., Gore-Langton, Georgia R., Boggs, Dorothy, Cousens, Simon, Moxon, Sarah G., Blencowe, Hannah, Baschieri, Angela, Rahman, Ahmed Ehsanur, Tahsina, Tazeen, Zaman, Sojib Bin, Hossain, Tanvir, Rahman, Qazi Sadeq-ur, Ameen, Shafiqul, El Arifeen, Shams, Kc, Ashish, Shrestha, Shree Krishna, Kc, Naresh P., Singh, Dela, Jha, Anjani Kumar, Jha, Bijay, Rana, Nisha, Basnet, Omkar, Joshi, Elisha, Paudel, Asmita, Shrestha, Parashu Ram, Jha, Deepak, Bastola, Ram Chandra, Ghimire, Jagat Jeevan, Paudel, Rajendra, Salim, Nahya, Shamb, Donat, Manji, Karim, Shabani, Josephine, Shirima, Kizito, Mkopi, Namala, Mrisho, Mwifadhi, Manzi, Fatuma, Jaribu, Jennie, Kija, Edward, Assenga, Evelyne, Kisenge, Rodrick, Pembe, Andrea, Hanson, Claudia, Mbaruku, Godfrey, Masanja, Honorati, Amouzou, Agbessi, Azim, Tariq, Jackson, Debra, Kabuteni, Theopista John, Mathai, Matthews, Monet, Jean-Pierre, Moran, Allisyn, Ram, Pavani, Rawlins, Barbara, Sæbø, Johan Ivar, Serbanescu, Florina, Vaz, Lara, Zaka,

	Nabila and Lawn, Joy E.		
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.

<u>SECTION C – Prepared for publication, but not yet published</u>

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.

<u>SECTION D – Multi-authored work</u>

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 drafted the first full manuscript, designed the data flow figure (5), incorporated collaborative inputs from PIs, country teams, national/global expert advisory groups. I submitted the manuscript on behalf of the authors and corresponding PI, revised the manuscript in response to peer-review, including co-ordination of collaborative inputs from co-authors. I co-managed proofs with my joint first author.

<u>SECTION E</u>

Student Signature	[redacted]
Date	[redacted]

Supervisor Signature	[redacted]
Date	[redacted]

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

journal of global health

"*Every Newborn*-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania

Louise T Day[*,1], Harriet Ruysen[*,1], Vladimir S Gordeev[1], Georgia R Gore-Langton[1], Dorothy Boggs[1], Simon Cousens[1], Sarah G Moxon[1], Hannah Blencowe[1], Angela Baschieri[1], Ahmed Ehsanur Rahman[2], Tazeen Tahsina[2], Sojib Bin Zaman[2], Tanvir Hossain[2], Qazi Sadeq-ur Rahman[2], Shafiqul Ameen[2], Shams El Arifeen[2], Ashish KC[3], Shree Krishna Shrestha[4], Naresh P KC[5], Dela Singh[4], Anjani Kumar Jha[6], Bijay Jha[6], Nisha Rana[3], Omkar Basnet[7], Elisha Joshi[8], Asmita Paudel[10], Parashu Ram Shrestha[5], Deepak Jha[5], Ram Chandra Bastola[9], Jagat Jeevan Ghimire[6], Rajendra Paudel[10], Nahya Salim[11], Donat Shamba[12], Karim Manji[11], Josephine Shabani[12], Kizito Shirima[12], Namala Mkopi[11], Mwifadhi Mrisho[12], Fatuma Manzi[12], Jennie Jaribu[12], Edward Kija[11], Evelyne Assenga[11], Rodrick Kisenge[11], Andrea Pembe[11], Claudia Hanson[13], Godfrey Mbaruku[12†], Honorati Masanja[12], Agbessi Amouzou[14], Tariq Azim[15], Debra Jackson[16], Theopista John Kabuteni[17], Matthews Mathai[18], Jean-Pierre Monet[19], Allisyn Moran[20], Pavani Ram[21], Barbara Rawlins[22], Johan Ivar Sæbø[23], Florina Serbanescu[24], Lara Vaz[25], Nabila Zaka[16], Joy E Lawn[1]

Authors' affiliations are at the end of the manuscript
*Joint first authors
†Deceased 2 September 2018

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

Background To achieve Sustainable Development Goals and Universal Health Coverage, programmatic data are essential. The *Every Newborn* Action Plan, agreed by all United Nations member states and >80 development partners, includes an ambitious Measurement Improvement Roadmap. Quality of care at birth is prioritised by both *Every Newborn* and Ending Preventable Maternal Mortality strategies, hence metrics need to advance from health service contact alone, to content of care. As facility births increase, monitoring using routine facility data in DHIS2 has potential, yet validation research has mainly focussed on maternal recall surveys. The *Every Newborn* – Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study aims to validate selected newborn and maternal indicators for routine tracking of coverage and quality of facility-based care for use at district, national and global levels.

Methods EN-BIRTH is an observational study including >20 000 facility births in three countries (Tanzania, Bangladesh and Nepal) to validate selected indicators. Direct clinical observation will be compared with facility register data and a pre-discharge maternal recall survey for indicators including: uterotonic administration, immediate newborn care, neonatal resuscitation and Kangaroo mother care. Indicators including neonatal infection management and antenatal corticosteroid administration, which cannot be easily observed, will be validated using inpatient records. Trained clinical observers in Labour/Delivery ward, Operation theatre, and Kangaroo mother care ward/areas will collect data using a tablet-based customised data capturing application. Sensitivity will be calculated for numerators of all indicators and specificity for those numerators with adequate information. Other objectives include comparison of denominator options (ie, true target population or surrogates) and quality of care analyses, especially regarding intervention timing. Barriers and enablers to routine recording and data usage will be assessed by data flow assessments, quantitative and qualitative analyses.

Conclusions To our knowledge, this is the first large, multi-country study validating facility-based routine data compared to direct observation for maternal and newborn care, designed to provide evidence to inform selection of a core list of indicators recommended for inclusion in national DHIS2. Availability and use of such data are fundamental to drive progress towards ending the annual 5.5 million preventable stillbirths, maternal and newborn deaths.

Valid data and measurement are central to achieving the Sustainable Development Goal (SDG) aspiration of "no-one left behind" [1]. In the United Nation's Global Strategy for Women's Children's and Adolescent's Health the ongoing imperative for the right to survive, is joined by a new focus on thriving, with wider transformation [2]. Progress for survival has been slowest for the 5.5 million deaths of women and babies around the time of birth each year, including an estimated 2.5 million newborns dying in the first 28 days of life, 2.6 million babies stillborn and 303 000 maternal deaths [3-5]. Most of these deaths happen to the poorest families in the poorest countries, and most are preventable [6]. Opportunity exists to save an estimated 3 million lives per year by improving quality of care at birth and care of small and sick newborns [7,8]. Based on this evidence, the *Every Newborn* Action Plan (ENAP) was launched in 2014 and endorsed by all member states in a World Health Assembly resolution [9]. The plan outlines 2030 country targets of 12 or fewer newborn deaths per 1000 live births and 12 or fewer stillbirths per 1000 total births. *Every Newborn* is closely aligned with the World Health Organization (WHO) Strategy for Ending Preventable Maternal Mortality (EPMM) [10] since both include a priority for quality of care at birth alongside the Quality, Equity, Dignity movement led by WHO, UNICEF and UNFPA in 11 countries, aiming to halve facility deaths by 2020 [11].

Accurate data are essential to drive progress towards these targets. However, at the dawn of the SDG era, most deaths around the time of birth still occur in settings with the least data on coverage and quality of care – the "inverse data law" [12]. One of five strategic objectives of *Every Newborn* is to transform measurement and use of data to track coverage and quality of care [8,9,13]. A top priority has been to develop and implement a time-limited plan to ensure required core indicators are validated and feasible to measure at scale. In support, WHO and the London School of Hygiene & Tropical Medicine (LSHTM) have coordinated an ambitious Measurement Improvement Roadmap which reviews specific measurement gaps and provides a multi-year, multi-partner pathway to define specific indicators, test validity if needed, develop tools, and promote use of data by 2020 [14-16].

Ten core indicators were prioritised as part of the *Every Newborn* multi-country consultation process including those for impact, coverage and input (**Figure 1**) [9,16,17]. This protocol relates to the coverage indicators shown in the middle of **Figure 1**. Indicators of coverage of care for all women and newborns are shaded amber, because whilst definitions are clear, content and quality of care data requires improvement. The greatest metrics gap is core coverage indicators for specific, high impact interventions, shown in red in **Figure 1**. The combination of core indicators for *Every Newborn* and EPMM is illustrated in **Figure 2** and approximately half of these indicators are the same [10]. Validating the highest priority indicators, highlighted in red in **Figure 2**, is the topic of this research: all women to receive uterotonics and

Current Status		Core Indicators	Additional indicators
Definitions clear – but quantity and consistency of data lacking	*Impact*	1. Maternal mortality ratio* 2. Stillbirth rate* 3. Neonatal mortality rate*	Intrapartum stillbirth rate Low birth weight rate Preterm birth rate Small for gestational age **Neonatal morbidity rates** **Disability after neonatal conditions**
Contact point definitions clear but data on content of care are lacking	*Coverage:* Care for All Mothers and Newborns	4. Skilled attendant at birth* 5. Early postnatal care for mothers and babies* 6. Essential newborn care (tracer is early breastfeeding)	Antenatal Care* Exclusive breastfeeding up to 6 months*
Gaps in coverage definitions, and requiring validation and feasibility testing for HMIS use	*Coverage:* Complications and Extra Care	7. **Neonatal resuscitation** 8. **Kangaroo mother care** 9. **Treatment of serious neonatal infections** 10. **Antenatal corticosteroid use**	Caesarean section rate **Chlorhexidine cord cleansing**
	Input: Service Delivery Packages for Quality of Care	Emergency Obstetric Care **Care of Small and Sick Newborns** **Every Mother Every Newborn Quality Initiative with measurable norms and standards**	
	Input: Counting	Birth Registration	Death registration, cause of death

Figure 1. *Every Newborn* Action Plan core and additional indicators. Shaded – not currently routinely tracked at global level. Bold red – indicator requiring additional testing to inform consistent measurement. Asterisk – also SDG core or complementary indicator. Indicators disaggregated by equity such as urban/rural, income, and education. Adapted from references [9,16,17].

	Ending Preventable Maternal Mortality (Phase 1)	Every Newborn	
IMPACT	**1.Maternal mortality ratio** Maternal cause of death with ICD-MM 4.Adolescent birth rate	**1.Maternal mortality ratio** 2.Stillbirth rate + intrapartum stillbirth rate 3.Neonatal mortality rate	
COVERAGE Care for all women and newborns	5.Antenatal visits – four or more **6.Skilled birth attendance** 7.Institutional Delivery **8.Early postnatal care** 9.Met need for family planning	**6.Skilled birth attendance** **8.Early postnatal care**	
	10.Uterotonic immediately after birth	11.Immediate newborn care (tracer immediate breastfeeding)	FOCUS OF EN-BIRTH STUDY
COVERAGE Care for all women and newborns with complications	**12.Caesarean section rate**	**12.Caesarean section rate** 13.Newborn resuscitation 14.Kangaroo mother care 15.Treatment of serious newborn infection Antenatal corticosteroid use	
INPUTS Counting	Maternal death registration	Birth registration, death registration, cause of death	
Availability of services and care	Availability of functional EmONC facilities Respectful maternity care	EmONC + service readiness for newborns (including small and sick newborns) Quality of Care- measurable norms and standards	

Figure 2. Combined priority indicator table for relevant plans: Ending Preventable Maternal Mortality and *Every Newborn* [10]. Highlighted in red with box is the priority for measurement improvement and the focus of this research.

newborns with complications to receive neonatal resuscitation, Kangaroo mother care (KMC), treatment for possible serious infections and maternal antenatal corticosteroids (ACS)[16]. The assumed need for these interventions, likely coverage and expected prevalence is shown in the Appendix S2, Table S1 in **Online Supplementary Document**.

Coverage is defined as the number of individuals receiving an intervention or service (numerator), from among the population in need of the intervention or service (denominator). To date the main source of coverage and impact data in high-burden countries has been intermittent household surveys, including: Demographic Health Survey (DHS) and Multiple Indicator Cluster Survey (MICS) [18,19]. Currently monitored coverage indicators, including antenatal care, skilled birth attendance and postnatal care, mainly measure contact points with health care services but additional indicators are required to capture effective content of care [16,20,21]. Quality of care measurement requires definitions of characteristics for both provision (eg, safety, effectiveness, timeliness, equity, completeness) and experience of care (eg, client satisfaction) [22,23]. Household survey data accuracy depends first on the woman's interpretation of what took place at the time and second on recalling and reporting this understanding up to five years after the event. Evidence suggests that household surveys do not always accurately capture either numerator or denominator for some treatment interventions, such as pneumonia in young children [24] and events during labour [25]. In addition, since measurement of newborns with complications occur only for a subset of births (3%-15%, see Appendix S2, Table S1 in the **Online Supplementary Document**), the sample size required is higher than possible in most national DHS. Consequently, not all desired maternal and newborn intervention coverage indicators specifically relating to content and quality of care, can be captured through household surveys [16,26].

Globally more than 75% of babies are now born in facilities, and local count data from routine registers is increasingly available [27]. Whilst health-facility data can be used to track coverage more frequently than surveys, previous studies have demonstrated mixed data quality [28-30]. Health workers recording the care they deliver face many barriers in documentation [31,32]. Capturing denominators through routine data are also a major challenge. Firstly, for indicators regarding interventions for the whole population, disaggregated by equity criteria, facility births are not the "true" population denominators. Given the lack of specific and appropriate denominator data, a national health management information system (HMIS) typically use census-based data for deriving forecasts and key population calculations [28]. Secondly, the challenge is magnified if the "true" denominator for the intervention is based on clinical need, so targeted at a proportion of the total population eg, requiring treatment for possible serious bacterial infection. Measurement of the "true" denominator requires consistent and objective measure of clinical need. Yet clinical judgement and decision making, even using evidence based algorithms, is often still subjective [33,34]. Live births are often used as a proxy denominator when it is challenging to define and measure the "true" denominator. A benchmark "target coverage level" is required when proxy denominators are used, because 100% coverage is only a target for a "true" denominator. For example, the "true" denominator for Caesarean Section rate is "women in need for Caesarean section". Because this is challenging to define and measure, the proxy denominator per 100 live births is used, but benchmarking a "target Caesarean Section rate" has proved complex [35-39]. Large inequity within countries and over- and under-provision occurring in parallel [40] highlight the problem of constructing useful indicators to measure and compare met need for complications. Therefore, an important focus of this study will be to compare various denominator options and, if using a proxy denominator to consider benchmarking.

The hierarchy of data needs (**Figure 3**) illustrates scope and granularity of data use decreases at higher levels of the health system [41]. At the point of service delivery, data are needed for individual clinical decisions and to measure the client's perspective of care received. At facility level, aggregate data are collated to inform administrative and managerial decisions for planning and local quality improvement, mortality audit etc. At district level, data are required for planning (eg, human resources, equipment and drug availability). At national and global level, it is not possible or useful to collate all these data used at lower levels of the system. But it is crucial for accountability purposes to track a few core, standardised indicators to monitor SDGs and Universal Health Coverage at all levels – these "core indicators" are shown in the centre of the pyramid (**Figure 3**). WHO maintains a core list of 100 health indicators [42] and ENAP has prioritised 10 core indicators [9,16,17].

Improvements in civil and vital registration systems are enabling a more rapid transition to more timely denominator data on births and deaths. Data systems are transitioning to increasing use of HMIS to collect, collate, analyse and report routine data from health facilities up to district and national level. This has potential to be cost-efficient and generate more frequent coverage measurements [16,27]. Electronic

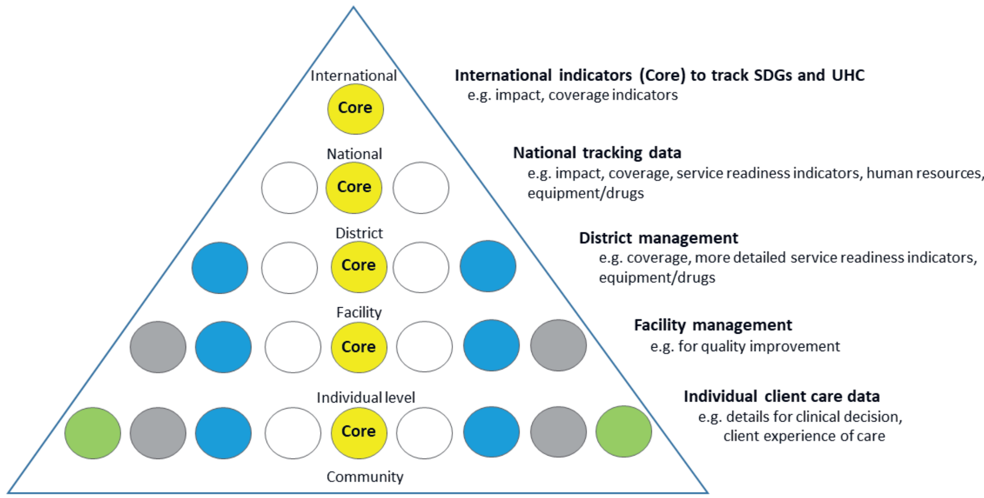


Figure 3. Data collection and use by level of health system. Adapted from [41].

HMIS platforms are increasingly being applied, offering great potential to harmonize traditionally fragmented information streams [43]. One such platform, the District Health Information System, version 2 (DHIS2) [44] is now being successfully implemented in >50 countries with high mortality burdens. Infrastructure and software development advances are currently driving a transition from predominately paper-based to mixed recording systems, even at clinical data level, ie, electronic patient records will increasingly be the basis of HMIS data in low-middle income country (LMIC) contexts.

Testing indicator validity is critical to improve measurement and inform decision makers of the likely accuracy of coverage collected by household survey and/or routine facility data [20]. Comparison of the reported indicator to an external data source "gold standard" is recommended [45]. Previous validation studies have mainly focused on population-based intervention coverage indicators for use in household surveys [25,46-48]. Observational studies to determine accuracy of facility registers in high burden settings have typically focused on outcome indicators [29]. The EN-BIRTH study seeks to address current evidence gaps by testing validity of priority coverage indicators for newborn and maternal health, in facilities in three high burden country settings.

Aim

This paper is the protocol paper for the *Every Newborn*-**B**irth **I**ndicators **R**esearch **T**racking in **H**ospitals (EN-BIRTH) Study, which aims to test validity of selected newborn and maternal care health intervention indicators (coverage/ quality aspects and/or safety) in facilities (**Table 1**). This study, as part of the *Every Newborn* Measurement Improvement Roadmap, and working closely with EPMM, aims to increase the evidence base to inform selection and use of maternal and newborn indicators in national HMIS (particularly DHIS2), and global tracking.

Research objectives

The research questions per objective, methods and analysis are detailed in **Table 2**.

Objective 1 – Numerators: To determine validity (accuracy) of both routine facility register and maternal recall surveys, compared to direct observation for selected maternal and newborn care interventions: uterotonics for 3rd stage labour, immediate breastfeeding, neonatal resuscitation, KMC; and, verification with patient case notes: neonatal infection management, and ACS administration (**Table 1**).

Objective 2 – Denominators: To compare different denominator options including proxies, and assess feasibility of their use in routine data platforms (**Table 1**), including:

- Target population requiring intervention (clinical need) in the facility ("true" denominator)
- Live births in the facility
- Total births (live births and stillbirths) in the facility
- Estimated population births (live or total): facility births *and* home births

Table 1. EN-BIRTH study selected indicators to be assessed for validity

Indicator	Place of care	Numerator	Denominator options
Uterotonic use for 3rd stage of labour	Labour/Delivery ward, or operating Theatre	Number of women who received a uterotonic immediately after birth	- Per 100 live births (currently used denominator) - Per 100 total births
Immediate newborn care		Number of babies who breastfed immediately after birth as possible surrogate for immediate newborn care	Per 100 live births (currently used denominator)
		Number of newborns who had chlorhexidine applied to the cord stump after birth (*Bangladesh and Nepal only*)	
Newborn resuscitation		Number of newborns for whom resuscitation actions (Bag and Mask Ventilation) were initiated	**To be compared for all 4 denominators options:** - Target population requiring the specific intervention (eg, admitted to the facility with presumed infection or at risk of preterm birth as per WHO guideline) - Live births in the facility - Total births in the facility (including stillbirths) - Estimated births in the population (live or total)
Kangaroo mother care (KMC)	KMC ward/ area	Number of eligible (<2000g) newborns initiated on facility-based KMC	
Treatment of neonatal infection	Newborn or postnatal wards	Number of neonates (<28 days old) who received at least one dose of antibiotic injection*	
Antenatal corticosteroid (ACS) use	Labour/delivery ward or antenatal ward	All women giving birth in a facility who are 24-34 weeks and received at least one dose of ACS†	

*Specific exclusions apply to exclude other primary diagnoses eg, congenital abnormalities, preterm births <32 weeks or <1500g and neonatal encephalopathy.
†ACS focus is to track safety, test methods to include gestational age and relevant safety outcomes.

Table 2. EN-BIRTH study summary of research questions, data collection and analysis by objective

Research questions	Data collection method	Data analysis approach
Objective 1 – Numerators		
- Do registers give a valid representation of observed maternal and newborn interventions? - Do maternal recall survey questions used in household surveys capture a valid representation of the observed maternal and newborn interventions? - What is the consistency between observers?	- Observation of clinical practice (or verification from inpatient records for neonatal infections and ACS) plus video film for neonatal resuscitation (Nepal only) - Maternal recall survey (all six indicators) - Extraction from routine data sources	- Sensitivity, positive predictive value - Specificity of numerator for those with all birth denominator or clearly measurable denominator - Inter-rater reliability (Cohen's Kappa)
Objective 2 – Denominators		
- How different are the coverage estimates when using alternative denominator options? - Which denominator options are feasible for use in each country HMIS?	Observation of clinical practice for measurement of "true" denominator Collection of hospital documentation for the denominator or alternative denominator options	- Descriptive statistics - Quantitative analysis with inflation factor for indicators with all-birth denominator
Objective 3 – Content and quality of care		
- What content of care are women and newborns observed to receive for each intervention, with focus on timing? - Which aspects of the content of care are already accurately recorded in registers? - Which aspects of the content of care are accurately recalled by women?	Observation of clinical practice (or verification from inpatient records for neonatal infections and ACS) plus video film for neonatal resuscitation (Nepal only) Maternal recall survey (all six indicators) Extraction from routine data sources	- Assessment of content/quality of care for specific aspects related to each intervention with emphasis on timing
Objective 4 – Barriers and enablers		
- Are some indicators recorded more completely than others? - Has routine recording changed during the time of the study? - What are the barriers and enablers to measurement of these indicators? - What are the barriers and enablers to perceived use of data regarding these indicators? - How can facility recording and flow of information into DHIS2 for these indicators be improved?	Quantitative – Register review for 12 months before and during study Qualitative FGD/IDI of study data collectors Qualitative FGD/IDI of health workers Qualitative FGD/IDI of other data users (policymakers etc) regarding data utility Process evaluation of data flow from patient level to DHIS2	- Quantitative comparison of registers applying data quality scores comparing before and after - Qualitative data for data collectors, health workers and data users - Process evaluation of data flow to DHIS2

FGD – focus group discussion, IDI – in-depth interview, DHIS2 – District Health Information System 2

Objective 3 – Content /quality of care: To evaluate different domains of coverage (eg, timing, completion rates, safety) for selected interventions (**Table 3**).

Objective 4 – Barriers and enablers: To evaluate barriers and enablers to routine recording of selected indicators, and to explore perceived utility of these data to improve decision-making, coverage and quality of care at all levels.

Table 3. EN-BIRTH study – Examples of indicator quality of care research questions, particularly regarding timing

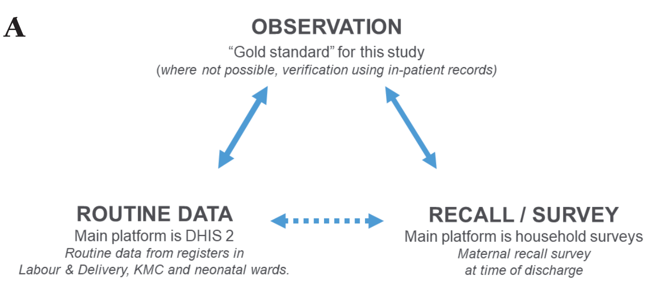
Intervention	Research question to answer using observation data
Uterotonic	Proportion of mothers who received oxytocin within recommended one minute after birth
Immediate breastfeeding	Proportion of babies whose breastfeeding was initiated within one hour of birth
Resuscitation	Proportion of non-breathing babies who had bag-and-mask initiated within one minute of birth
Kangaroo mother care	Proportion of babies receiving KMC, held in skin-to-skin position for 18 h or more, during the last 24 h
Neonatal infection	Proportion of cases with presumed sepsis, treated with antibiotics and for whom a blood culture result was available
Antenatal corticosteroids	Proportion of preterm labour cases who received antenatal corticosteroids according to WHO criteria for safety

METHODS

Study design

The EN-BIRTH study uses quantitative and qualitative methods across four objectives (**Table 2**). The validity of coverage indicators of selected maternal and newborn interventions as measured by routine facility registers and maternal recall surveys will be assessed by comparison with the "gold standard" of direct observation (**Figure 4**, panel A). Observation will be undertaken in three clinical settings (labour/

A



B

OBSERVED INTERVENTION (GOLD STANDARD)

RECORD OF INTERVENTION	Positive	Negative	
Positive	True positive (TP) Intervention occurred and recorded /reported*	False positive (FP) Intervention did not occur, but recorded /reported*	Total recorded /reported* as receiving intervention = TP + FP
Negative	False negative (FN) Intervention occurred but not recorded/reported *	True negative (TN) Intervention did not occur and not recorded/reported*	Total recorded /reported* as not receiving intervention = FN + TN
	Total receiving the intervention TP + FN	Total not receiving the intervention FP +TN	Total observed cases TP+FN+FP+TN
	Sensitivity =TP/ (TP+FN) Those receiving the intervention who were recorded /reported* as receiving the intervention i.e. expressing the level of true positives recorded /reported*	**Specificity =TN/ (FP +TN)** Those who did not receive the intervention who were recorded / reported* as having received the intervention i.e. the chances of misreporting /recording an intervention that did not happen	

TERMS	DEFINITION
Accuracy	Closeness of a measured value to a standard value For EN-BIRTH study, closeness of register recorded or maternal recall survey compared to direct clinical observation
Validation	The process whereby the ability of health indicators to measure what they are supposed to measure is determined compared to a standard
Positive Predictive Value (PPV)	The probability that an intervention recorded/reported as received had been observed as received = TP/ (TP + FP)
Negative predictive value (NPV)	The probability that an intervention not recorded/reported had been observed as not received = TN/ (TN+FN)

Figure 4. EN-BIRTH study validation and analysis approach. **Panel A.** Validation "gold standard" comparison to routine data (eg, HMIS/DHIS2) and to maternal recall survey data (eg, for household surveys). **Panel B.** Analysis for validation of sensitivity and specificity. Asterisk – recorded in facility L&D or KMC register / reported in maternal recall survey.

delivery ward, operation theatre, and KMC ward/area) by trained clinical observers. Data will be extracted from facility registers and verification of inpatient records carried out for newborns who received antibiotics for presumed infection, and for women who received ACS. Interviews to capture maternal recall will be conducted prior to discharge with all women whose births and/or their newborn's care were observed or case notes were verified. In addition, barriers and enablers to recording of selected indicators in routine facility registers will be evaluated. Data flow into national HMIS platforms and perceived utility of data will be documented.

Research questions were informed by consultation with many *Every Newborn* stakeholders [9,17] including WHO-led Measurement Improvement Roadmap meeting [15] and EN-BIRTH Expert Advisory Group (listed as author group). More than 60 participants in an EN-BIRTH study design workshop [49] provided representation from country partners, national stakeholders, UN agencies, leading academic and professional experts in the field, governmental and non-governmental organisations, clinicians, program managers, other key experts and donors (see Appendix S1 in **Online Supplementary Document**) and contributed to development of the research protocol (**Box 1**).

Box 1. Authorship teams for EN-BIRTH study

EN-BIRTH LSHTM Team: Louise T Day, Harriet Ruysen, Vladimir S Gordeev, Georgia R Gore-Langton, Dorothy Boggs, Simon Cousens, Sarah G Moxon, Hannah Blencowe, Angela Baschieri.

EN-BIRTH Co-PI and country teams

Bangladesh: Ahmed Ehsanur Rahman, Tazeen Tahsina, Sojib Bin Zaman, Tanvir Hossain, Qazi Sadeq-ur Rahman, Shafiqul Ameen, Shams El Arifeen.

Nepal: Ashish KC, Shree Krishna Shrestha, Naresh P KC, Dela Singh, Anjani Kumar Jha,

Bijay Jha, Nisha Rana, Omkar Basnet, Elisha Joshi, Asmita Paudel, Parashu Ram Shrestha, Deepak Jha, Ram Chandra Bastola, Jagat Jeevan Ghimire, Rajendra Paudel.

Tanzania: Nahya Salim, Donat Shamba, Karim Manji, Josephine Shabani, Kizito Shirima, Namala Mkopi, Mwifadhi Mrisho, Fatuma Manzi, Jennie Jaribu, Edward Kija, Evelyne Assenga, Rodrick Kisenge, Andrea Pembe, Claudia Hanson, Godfrey Mbaruku, Honorati Masanja.

Senior author/corresponding: Joy E Lawn

With the EN-BIRTH Expert Advisory group

Agbessi Amouzou, Tariq Azim, Debra Jackson, Theopista John Kabuteni, Matthews Mathai, Jean-Pierre Monet, Allisyn Moran, Pavani Ram, Barbara Rawlins, Johan Ivar Sæbø, Florina Serbanescu, Lara Vaz, Nabila Zaka.

On behalf of the EN-BIRTH study research design Windsor Workshop Invitees (not already names in above author groups

AI Ayede, Simon Azariah, Anne-Marie Bergh, Elahi Chowdhury, Olive Cocoman, Patricia Coffey, Jai Das, Ashok Deorari, Mary Drake, Queen Dube, Suzanne Fournier, John Grove, Rima Jolivet, Amira Khan, Dyson Likomwa, James Litch, Goldy Mazia, Kate Milner, Indira Narayanan, Susan Niermeyer, Alfred Osoti, Sayed Rubayet, Joanna Schellenberg, Wilfred Senyoni, Gaurav Sharma, Kavita Singh, Nalini Singhal, Cally Tann, Steve Wall.

Study settings

Tanzania, Bangladesh and Nepal were chosen as LMIC's currently implementing the selected maternal and newborn interventions within Sub-Saharan Africa and Asia [50]. Within these countries, research centres of excellence with a strong track record in maternal and newborn health were selected: Ifakara Health Institute (IHI) and Muhimbili University of Health and Allied Sciences (MUHAS) in Tanzania, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b); UNICEF-Nepal with Lifeline in Nepal. Criteria for selection of facilities were: providing the selected interventions in line with current WHO recommendations for improving quality of care; existing registers recording most interventions; and sufficient number of births to ensure sample size (except for ACS discussed under sample size section below).

Study populations

Inclusion / exclusion criteria for consenting women according to data collection methods (**Figure 5**) are:

- **Observation on labour and delivery, operating theatre:** All admitted women in active labour excluding those likely to deliver immediately. Women with a prior diagnosis of intrauterine death, were also excluded to avoid further maternal distress.

- **Observation KMC ward/area:** All in-born and out-born neonates admitted for KMC.
- **Verification from inpatient records for ACS administration**: All women being observed and reported to be <34 weeks' gestation at admission from Expected Date of Delivery (EDD).
- **Verification from inpatient records for neonatal infection cases:** All babies < 28 days old with a main diagnosis of infection (sepsis/meningitis) recorded in neonatal register or admission/discharge book. Babies will be excluded for major congenital abnormality, neonatal encephalopathy/severe asphyxia, <32 weeks' gestation and/or admission weight <1500 grammes.
- **Maternal recall survey:** All women whose birth and/or their newborn's KMC will be observed, or case notes verified for ACS or neonatal infection.
- **Routine register extraction:** All women whose birth and/or their newborn's KMC will be observed.

Sample size

Sample size was based on planned analysis for validity in objective one, by assuming 50% sensitivity ±10% precision, 50% specificity ±10% precision, with $\alpha = 0.05$ and then applying the lowest previously published rates for neonatal resuscitation [51] and for KMC initiation [52,53]. Since formative data suggested >80% coverage for uterotonic administration, this indicator will be well-powered (see Appendix S2, Tables S2-3 in **Online Supplementary Document**). Hence minimal sample size is 4850 observations in each country, increased to 5390 observations to allow for a non-consent rate of 10% (**Table 4**). As expected prevalence of ACS is less than 0.5%, the resulting very large sample size was not feasible for this study [54,55]. The 5390 observations will be collected from three countries. In Tanzania and Nepal, each facility will observe this number of births, and in Bangladesh observations will take place in two facilities (**Table 4**) [4,5,56]. We anticipate a total >20 000 observed births aiming to capture at least 106 observations per intervention per country, except for ACS (**Table 4** and Appendix S2, Table S3 in **Online Supplementary Document**).

Tool development

A formative research phase was undertaken from July – December 2016 including: health facility assessments [57], register reviews, data flow assessments, and interviews/focus group discussions (FGDs) with women, caregivers, health workers and senior facility-level staff. The results helped ensure study sites could meet inclusion criteria, achieve required sample size and informed refinement of observer checklists and data collection processes. Maternal Recall survey tools were translated into local languages and back-translated.

Table 4. EN-BIRTH study – national mortality rates, facility context and expected number of births and cases per indicator

Context		Facilities			Sample size		
Country	**National mortality rates***	**Name**	**Hospital type**	**Annual total births**	**Expected births in study**	**Uterotonic use†**	**Each for: resuscitation, Kangaroo mother care, neonatal infection management†**
Tanzania	MMR = 398 /100 000 NMR = 22/1000 SBR = 22/1000	Muhimbili National Hospital, Dar es Salaam	National Referral & University Teaching	9773	5390	>4310	>106
		Temeke Regional Hospital, Dar es Salaam	Regional Referral	14 655	5390	>4310	>106
Subtotal					**10780**	**>8620**	**>212**
Bangladesh	MMR = 176/100 000 NMR = 21/1000 SBR = 25/1000	Maternal and Child Health Training Institute (MCHTI), Dhaka	Tertiary	4488	2695	>2150	>53
		Kushtia District Hospital	Secondary	2581	2695	>2,150	>53
Subtotal					**5390**	**>4,310**	**>106**
Nepal	MMR = 258 /100 000 NMR = 22/1000 SBR = 18/1000	Pokhara Academy of Health Sciences	Tertiary	9427	5390	>4310	>106
TOTAL all				**40 924**	**21 560**	**>17 240**	**>424**

*MMR – maternal mortality ratio per 100 000 live births [5]; NMR – neonatal mortality rate per 1000 live births [54]; SBR – stillbirth rate per 1000 total births [4].

†Prevalence/incidence based on references [51-53,55,56]. More details in Appendix S2 of **Online Supplementary Document.**

Data collection software application

The development of a customised tablet-based software application (Android-based) for data collection and monitoring was undertaken by the icddr,b team supported by LSHTM (**Figure 5** and **Figure 6**) [58]. The software application has different permissions for various data collector cadres (observation, verification, maternal recall survey, and data extraction) and translated into local languages where relevant. Time-stamped data will be collected using this EN-BIRTH data collection software, stored locally on the tablet, and synchronised regularly to the local central secure database server.

Training of data collectors and supervisors

Data collector cadres include: tracker (responsible for consent, registration and assigning for observation/record verification and subsequent tracking); observer (direct observational data for assigned women and babies); interviewer (maternal recall survey interviews); data verifier/extractor (data from facility registers or case notes); and supervisor (responsible for all data collectors and quality assurance) (**Figure 5**). Observers with a clinical background (eg, nurses) will be recruited. Data collection staff will receive two weeks of training using classroom-based sessions, group activities and mock data collection within the health facility, detailed in the Data Collectors Training Handbook [58]. Observer training will include guidance on response to specific events, including managing maternal distress and when to pause data collection and assist in the care of the patient, if they perceive facility staff are responding inappropriately to a life-threatening situation. A minimum individual post-training assessment score of ≥80% is required before data collection can commence.

Procedures according to data collection method

Observation (Objectives 1, 2, and 3)

Informed written consent will be obtained prior to study registration and basic demographic data collected (**Figure 5**) by the tracker. Verbal consent will be obtained from the health workers. Observers working in Labour/Delivery ward, Operating theatre and KMC ward/areas will collect direct clinical observation data. These observers will not interact with participating pregnant women, her family members or attending health workers during observation (except to respond to a life-threatening event [58]).

Observations on Labour/Delivery ward will focus on specific aspects of: 1st, 2nd and 3rd stage of labour, postpartum haemorrhage, immediate newborn care and neonatal resuscitation. Multiple parameters will be recorded to assess content/quality of care, particularly related to intervention timing. KMC observations will focus on domains of initiation, position, feeding and other treatment administered. Mother and baby outcome at discharge from hospital will be documented [58].

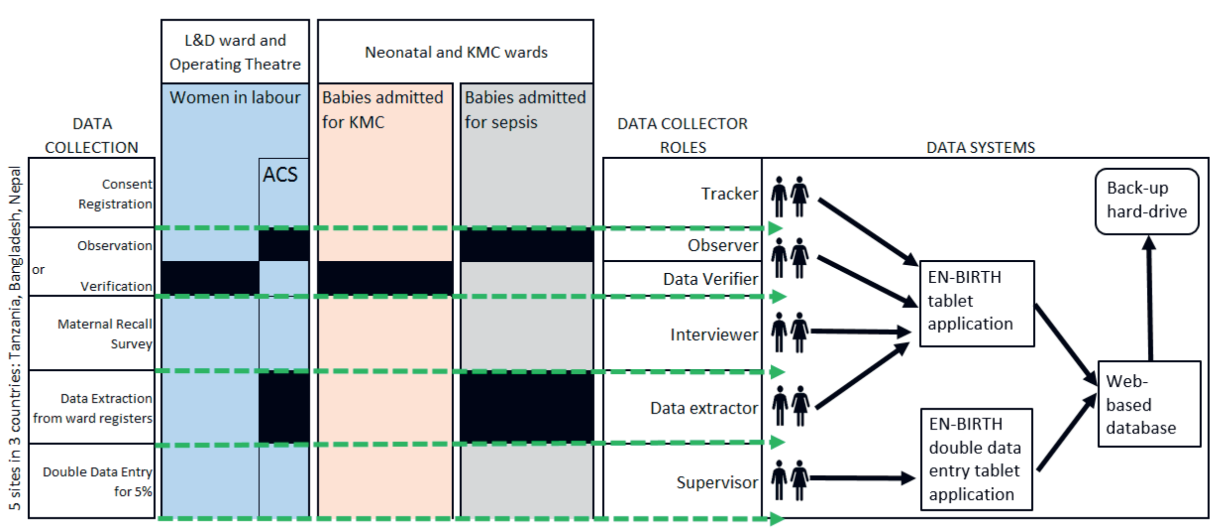


Figure 5. EN-BIRTH study – overview of data flow in study sites. Data Collection – "ward registers" on one line. Data collector roles revised with "Data Verifier" added. Data Systems needed "web based database" (word database was missing). ACS – antenatal corticosteroids.

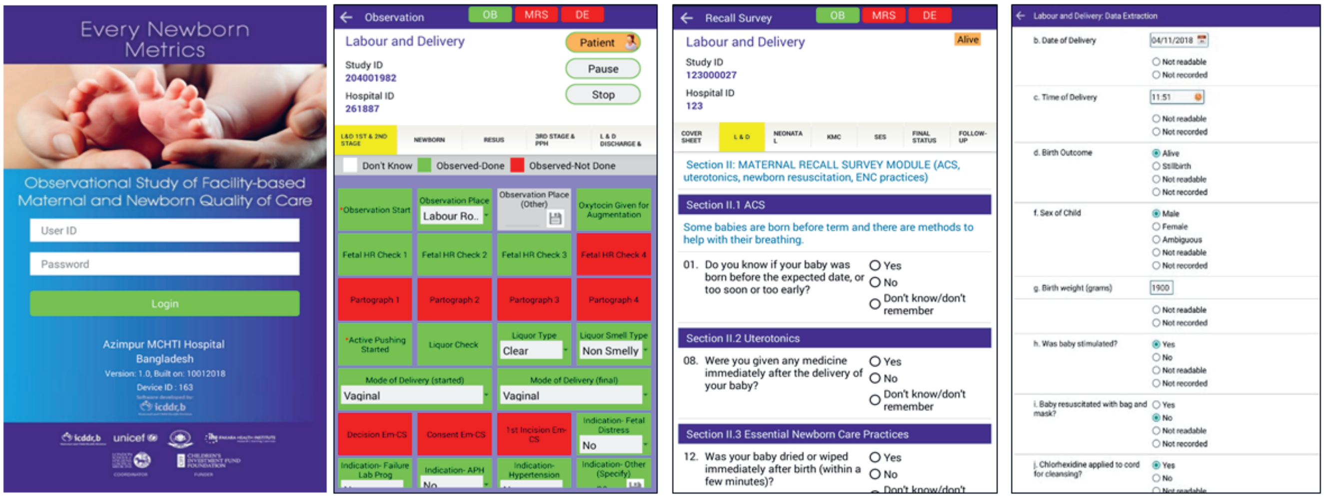


Figure 6. EN-BIRTH study software data collection showing examples of the tablet application screen shots.

Additionally in Nepal for neonatal resuscitation, observation video film recording and physiological assessment will be undertaken. Information regarding these additional processes will be provided separately to women and informed, signed consent taken [59]. Video cameras and pulse oximeters will be placed on resuscitation tables within Labour/Delivery ward and Operating theatres and research staff trained in this equipment operation and maintenance. A trained data collector will complete the observation checklist for resuscitation using the recorded video within 24 hours of birth [60,61]. If consent is subsequently withdrawn for video use, this data will be excluded, and the video deleted.

Verification using inpatient notes (Objectives 1, 2, and 3)

During the formative phase it was recognised that direct observation was not feasible for two of the selected interventions (neonatal infection and antenatal corticosteroids). For these interventions, data verifiers will use patient charts/ case notes, drug charts, laboratory reports and other relevant routine documentation to verify intervention and quality of care measurements. Supervisors will review/search for any missing or illegible documents before confirming data not readable/ not recorded [58].

Maternal Recall Survey (Objectives 1, 2, and 3)

Data collectors will interview mothers whose baby's birth or treatment is observed and/or verified prior to discharge from postnatal or KMC ward/areas The software programming of the structured questionnaires will automatically skip certain questions to minimise any risk of further emotional trauma if the mother has experienced a stillborn or neonatal death [58]. For multiple births the interview will be completed only for first-born babies. Consent will be repeated before this interview in recognition that the mother may have been in labour when she first consented to participation in this research. Consent will also be taken for repeat maternal recall surveys at different intervals after discharge, if funded for follow-up.

Routine register data extraction (Objectives 1, 2 and 3)

Data extractors will use routine labour/delivery registers, KMC registers and neonatal ward registers to extract participant data recorded by facility staff. If data are illegible or cannot be found, supervisors will review/search for these documents, before documenting data not readable/not recorded [58].

Assess barriers and enablers (Objective 4)

Mixed methods will be used to identify barriers and enablers to routine data recording and use of selected indicators (**Table 1**). Completeness and quality of existing documentation in routine registers (labour/delivery, KMC and/or neonatal) for 12 months prior to the study will be evaluated. In Bangladesh and Nepal, 100% of cases in these registers will be extracted. In the Tanzanian facilities, with a high number of births, a 20% sample randomly selected will be used for labour/delivery cases with 100% for KMC and neonatal infection cases.

Qualitative data collection tools for FGD, in-depth and key informant interviews will be informed by the MEASURE Evaluation Performance of Routine Information System Management (PRISM) conceptual framework and tools [62], including constructs for Technical, Organizational and Behavioural factors. Data will be collected from study data collectors and facility health workers. Data flow assessments will provide information on movement of data from registers, into DHIS2 and up to national level. Additionally, perceptions regarding indicators which are considered most valuable and most feasible to collect will be explored through interviews with policy makers and technical managers of DHIS2.

Data quality monitoring

The EN-BIRTH data collection software includes skip rules, and consistency checks as well as pre-defined value ranges for some variables. Progress will be monitored by an online data dashboard, providing real-time summary tables per site, including data capture cascade for selected coverage indicators at each step; registration, consent, observation/ verification, maternal recall survey and register data extraction. A traffic light system will indicate overall progress for each indicator using pre-defined thresholds. Bi-weekly all-site calls will provide an opportunity for country teams to review and discuss progress using these data dashboards, in addition to promoting collaborative quality improvement initiatives between countries and sites.

As part of the quality assurance process, for approximately 5% of cases in each site, simultaneous supervisor observation and duplicate data verification and extraction will also be conducted using EN-BIRTH data collection software. The supervisor data will be regarded as the standard, stored in a separate database, and variability between individual data collectors estimated by calculating inter-rater reliability using Cohen's kappa (κ) coefficient. Minimum agreement levels of ≥71% for observation and ≥91% for data extraction/case verification will be used [63].

Data management

EN-BIRTH tablet data will be synchronised, and uploaded to an in-country central server, regularly backed-up. Raw data will be encrypted, and access restricted to country data manager who will anonymise data before data sets are pooled. Server maintenance, data management, and cleaning will be coordinated according to agreed protocols including logical and completeness checks. A unified variable code book will contain description of variable names and answer options. Qualitative data will be digitally recorded, transcribed, and translated into English. All data will be stored on password-protected computers.

Analysis plan

Analyses will be coordinated, using a standard approach, both combining sites, and with site-specific and/or country-specific analyses. An overview of research objectives, main research questions and data analysis approach are summarised in **Table 2**. Quantitative analyses will be undertaken with Stata 15 (*Stata Statistical Software: Release 15*).

Objective 1 – Numerator

The "gold standard" used for comparison will be direct observation of selected interventions by research observer, except for neonatal infection and ACS, where in-patient note verification will be used. Data extracted from facility routine register records and data collected during maternal recall survey will be compared with this "gold standard" separately (**Figure 4**, panel A). Accuracy of each individual coverage indicator will be assessed by constructing two-by-two tables to analyse the sensitivity and positive-predictive value of routine data (**Figure 4**, panel B). Specificity of routine data will be assessed for those indicators with true negatives and confidence intervals will be computed. "Area Under the Curve" previously used for coverage indicators validation will be used for indicators with true negatives [25,46-48,64].

Objective 2 – Denominators

Various denominator options (**Table 1**) will be compared using descriptive statistics to assess variation in estimated coverage and undertake analyses to guide benchmarking. Information on denominators will come from the EN-BIRTH data set, facility total birth data collected from facility reports, and population birth data from estimates based on census or survey and fertility rates, as used in DHIS2. For indicators with a whole population denominator (ie, uterotonics, breastfeeding) or a clearly measurable "true" denominator regarding clinical need (eg, KMC – birth weight <2000g), the inflation factor will be used. Inflation factor is the ratio of estimated routine recording-based prevalence to true (observed) population-based prevalence. It represents the magnitude of over- or under-estimation in the study setting relative to true population-based prevalence.

Objective 3 – Content/quality of care

Multiple recorded parameters will be analysed to assess measurement related to content/quality of care, particularly regarding timing of interventions and in relation to WHO Guideline recommendations (**Table 3**).

Objective 4 – Barriers and enablers

To assess barriers and enablers to indicator data recording and use, mixed methods will be used based on a framework adapted from PRISM [62] and considering other tools [65]. Quantitative analysis of routine register data collected prior to and during the study will address two research questions: (1) Are some indicators recorded more completely than others? (2) Has routine recording changed during the study time? Qualitative data from FGDs, in-depth and key informant interviews will be analysed using QSR International's NVivo 12 qualitative software (NVivo qualitative data analysis Software; QSR International Pty Ltd Version 12.1, 2018). Predetermined codes will be applied by two independent researchers, data managed into units of information covering broad categories with grouping of relevant emerging themes of importance.

DISCUSSION

EN-BIRTH is the first large study to assess validity of newborn and maternal care indicators in routine data systems, doing so at very large scale (>20 000 observed births) across three countries with a high-burden of mortality. Previous maternal and newborn indicator validation studies have focused on testing the validity of women's self-report method, used in population-based household surveys [25,46-48,64]. Validation of facility registers have focussed on outcome measures [29]. The EN-BIRTH study seeks to validate both routine registers and maternal recall at discharge for coverage indicators of high impact interventions. The novel software developed for this research allows detailed and precise recording of events around the time of birth, and particularly the timing of interventions. There are many studies examining quality of care at birth [66,67], and this research is not repeating that, but is focused on accuracy of routine reporting of care.

This research responds to calls from country and programme leaders for guidance on indicators for maternal and newborn services, tracking progress towards meeting national targets and Universal Health Coverage [9,17,68]. The high reporting load for many countries with multiple programmes, donors, and indicators, may result in the so-called data rich, information poor (DRIP) syndrome [69]. In addition to high reporting burden on the system, the individual midwives and doctors are responsible for recording data in multiple registers and patient records, sometimes at the expense of providing respectful quality care for women and babies. Hence a shorter list of evidence-based, indicators is required for national tracking, taking in to account validity and utility in low-resource, high-burden settings. The results of this study will inform recommendations for indicators appropriate for uptake within HMIS, and may also identify some that are not appropriate for use at higher levels of the health system (**Figure 3**). This research will also help inform improved capture and quality of data in HMIS, and especially DHIS2.

During the MDG-era, population-level surveys were the most common data source in high-burden countries, but studies consistently demonstrate challenges with maternal recall data, especially regarding details of clinical interventions [24,25,46,47,64]. For data that require medical knowledge and especially events that women may not have closely witnessed (eg, neonatal resuscitation), we expect poor maternal recall, which may reflect the lack of information given to families experiencing complications. Given continued reliance on household surveys for demographic and health data in many remote or unstable settings, we anticipate the main value of our maternal recall survey validation findings will be to contribute to the understanding of which indicators are not suitable for use in household surveys. We anticipate that if the woman does not know about the intervention at discharge from hospital, then recall later will not be useful.

A strength of this study design is the rigorous assessment of validity at scale, of facility routine data by comparison with direct observation, defined here as the "gold standard". Another strength is a specific focus on the denominator challenge. In an era of Universal Health Coverage, with discussions surrounding scale-up of more complex care for targeted populations, the science of denominator measurement, use of proxies, and selection of benchmarks will be increasingly important. This challenge applies to denominator measurement for maternal and newborn complications (as well as other large burden conditions, notably non-communicable diseases). This study, however, is not designed to validate the denominator based on subjective assessment of clinical need (eg, requiring neonatal resuscitation). Hence, we will only be able to measure true negatives, calculate specificity, and undertake analysis of "area under the curve" for interventions with a total population or clearly defined denominator [25,46-48,64].

This research also offers a unique opportunity to examine quality of care data from >20 000 births and assess to what extent we can accurately capture specific components including content and timing of selected interventions. Although multiple specific aspects of care may be measured locally to drive quality of care improvement at facility level, here we will focus on quality of care indicators that may be useful at district or national levels of the health system. Timing of interventions is a critical marker of quality of care, since delays are a matter of life or death: a woman may die in hours, a baby in minutes. Moreover, the sequence of interventions is complex and even concurrent (eg, how often is the correct dose of uterotonic given <1 minute after birth to prevent a woman bleeding from postpartum haemorrhage; How soon is bag-and-mask ventilation initiated for a baby who is not breathing; How many hours each day is a baby kept in KMC position). The time-stamped design of EN-BIRTH data collection software will permit analysis of such sequences.

Whilst direct observation is considered the "gold standard", data collectors might miss interventions, with concurrent actions at birth, especially in an emergency. We will limit potential recording bias by using observers with health backgrounds who are familiar with the procedures under observation [70-72]. EN-BIRTH data will also be directly on the tablet software to allow fast data capture. The study also presents several ethical challenges including the dilemma of observing a life-threatening situation without appropriate response from facility staff, and gaining informed consent during labour [58]. The clinically trained observers will have underlying familiarity of hospital environments, experience to uphold study protocols correctly [70] and experience in maintaining participant confidentiality. Training and processes will be put in place to take account of professional and legal duty of care.

The "Hawthorne effect" describes the phenomenon when a research participant's behavior is altered as a consequence of being studied or observed, and can be a source of bias in observational research [73]. Within this study, it is possible that clinical observers' presence will influence health workers to change their approach to care and routine register data. However, there is some evidence to suggest that sustained contact with participants (as with this study) may mitigate altered behaviors in health care settings [74]. To assess this bias, we will analyze changes in register data completeness and quality before and during the study.

Although the EN-BIRTH study is not powered to validate an ACS administration indicator, this will be included. Current WHO guidelines provide strong recommendation for the provision of a single course of ACS for any woman at risk of imminent preterm birth (24-34 weeks of gestation) provided the following criteria are met: 1) accurate assessment of gestational age; 2) no evidence of maternal infection; 3) preterm birth is considered imminent; 4) available adequate childbirth and newborn care services [75]. EN-BIRTH study sites were assessed in accordance with these WHO guidelines. The Antenatal Corticosteroid Trial (ACT) evaluated use of ACS at lower levels of the health system, with half of study births in home settings and care often provided by traditional birth attendants [76]. ACT reported an adverse outcome risk particularly in cases where ACS administration was after 34 weeks and outlines important challenges for measurement of gestational age, and assessment of maternal infection. This demonstrated need for robust data and further evidence in such settings, along with the imperative of ensuring safety and effectiveness, make measurement of ACS coverage and outcomes essential. Therefore, the EN-BIRTH study ACS analysis will focus on assessing relevant documentation to report the current ACS administration practice, compared with WHO safety criteria [75].

Given the importance of the neonatal period in terms of risk and prevention of long-term adverse child development outcomes, we plan a five-year follow-up for EN-BIRTH study recruited children who received basic neonatal interventions [77]. The *Every Newborn* – **S**implified **M**easurement **I**ntegrating **L**ongitudinal **N**eurodevelopment & **G**rowth (EN-SMILING) aims to detect child development outcomes as early as possible for referral to services, and to improve routine measurement of child development outcomes in programme settings.

The EN-BIRTH study is richer through active involvement of experts and policymakers from the EN-BIRTH Expert Advisory Group, *Every Newborn* implementation community, EPMM, UN Agencies including WHO, UNICEF and UNFPA as well as many partners and donors. In further support of this goal, each of the three countries have National Advisory Committees who will actively participate in the research process and support uptake of findings. Results will also be published in peer reviewed journals and disseminated with all relevant audiences. Following EN-BIRTH study validity testing, an important next step will be to evaluate feasibility of a short-list of indicators at different levels of the health system.

Most of the 5.5 million deaths around the time of birth [3] still occur in settings with the least data. Household surveys remain a key data source in the poorest countries, and *Every Newborn* is also involved in a multi-site study, EN-INDEPTH, to assess and improve these data [78]. Data improvement is fundamental for monitoring more rapid progress towards meeting global and national mortality targets, and in achieving Universal Health Coverage for all women and newborns [15]. With ongoing investment in electronic data platforms (including DHIS2) and increasing country demand for evidence-based indicators, we anticipate that these results will advance availability and use of data to change coverage, quality and equity, to help end preventable maternal and newborn mortality, as well as stillbirths.

Acknowledgements: We credit the inspiration of the late Godfrey Mbaruku. Many thanks to Claudia DaSilva, Fion Hay, Alegria Perez, Sadie Sareen, Adeline Herman, Veronica Ulay, Mohammad Raisul Islam and Ziaul Haque Shaikh, Susheel Karki and Bhula Rai for their administrative support. We thank Sabrina Jabeen, Tamatun Islam Tanha, Goutom Banik and Md Moshiur Rahman for their support in providing training to data collectors in the Bangladesh sites. We would also like to thank Ann Blanc, Liliana Carvajal, Doris Chou, Kim Dickson, Tanya Marchant, Claire-Helene Mershon, Natalie Roos, Anna Seale, Theresa Shaver, Deborah Sitrin, Kate Somers, and Cindy Stanton for sharing relevant technical inputs and expertise.

We acknowledge the National Advisory Groups: **Tanzania**: Muhammad Bakari Kambi, Georgina Msemo, Asia Husein, Talhiya Yahya, Claud Kumalija, Eliakim Eliud, Mary Azayo, Onest Kimaro. **Bangladesh**: Mohammad Shahidullah, Khaleda Islam, Md Jahurul Islam (joining the EN-BIRTH Expert Advisory Group in 2018). **Nepal**: Tara Pokharel, Uwe Ewald.

Finally, and most importantly we thank the women, their families, the health workers and the hundreds of data collectors involved in the EN-BIRTH study.

Ethics and consent to participate: This study was granted ethical approval by institutional review boards in all operating counties including the London School of Hygiene & Tropical Medicine (Appendix S3 of **Online Supplementary Document**).

Availability of data and material: All collaborating partners have signed data sharing and transfer agreements.

Funding: The Children's Investment Fund Foundation (CIFF) are the main funder of this research which is administered via The London School of Hygiene & Tropical Medicine. The Swedish Research Council specifically funded the Nepal site through UNICEF and Lifeline Nepal. The main funding for the Windsor research design workshop was provided by CIFF and in addition, the United States Agency for International Development, Saving Newborn Lives/Save the Children, WHO and Bill & Melinda Gates Foundation through the United States Fund for UNICEF funded many participants' travel and accommodation for attendance.

Authorship contributions: The study was conceptualized by JEL in 2014, and the initial protocol was coordinated by HR with JEL during 2016, with inputs from SC, SM, HB, the EN-BIRTH advisory group (names listed above), and during a multi-stakeholder Windsor research design workshop (names listed above). From mid-2016 VSG, GGL, DB and AB, with LTD joining in 2017. GGL and HR led inputs to observation checklists and led development of training materials with country coordinators (Bangladesh, TT and AER, Nepal, NR, Tanzania, NS). DB led development of the maternal recall survey and coordinated the health facility assessments. Each of the three country research teams input to all the data collection tools and review processes. The iccdr,b team from Bangladesh (notably AER, TT, TH, QSR, SA and SBZ) led the development of the software application, data dashboards and database development with VSG and the LSHTM team. iccdr,b (AER) also led the development of the verification form for infection case management and the data variable dictionary. The IHI and MUHAS in Tanzania hosted the study implementation workshop, input to data collection tools and training materials and are leading work on objective 4 (barriers and enablers for data collection and use). The Nepal team input to data collection tools and training materials, and led on use of video filming. The manuscript was drafted by HR, JEL and LTD with further review of the analysis sections by VSG and SC, in addition to major inputs particularly from GM, AER and TT. All authors reviewed and helped to revise the manuscript.

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

Additional material

Online Supplementary Document

Authors' affiliations:

[1] Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene & Tropical Medicine (LSHTM), London, UK
[2] Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh
[3] Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden
[4] Pokhara Academy of Health Science, Pokhara Ranipauwa, Nepal
[5] Department of Health Services, Ministry of Health, Kathmandu, Nepal
[6] Nepal Health Research Council, Kathmandu, Nepal
[7] Golden Community, Kathmandu, Nepal
[8] LifeLine Nepal, Kathmandu, Nepal
[9] Matri Shishu Miteri Hospital, Pokhara, Nepal
[10] Kanti Children's Hospital, Kathmandu, Nepal
[11] Department of Paediatrics and Child Health, Muhimbili University of Health and Allied Sciences, Dar Es Salaam, Tanzania
[12] Department of Health Systems, Impact Evaluation and Policy, Ifakara Health Institute, Dar es Salaam, Tanzania
[13] Public Health Sciences – Global Health – Health Systems and Policy, Karolinska Institutet, Stockholm, Sweden
[14] Institute for International Programs, Department of International Health, Johns Hopkins University, Baltimore, Maryland, USA
[15] MEAUSRE Evaluation, University of North Carolina, North Carolina, USA
[16] Knowledge Management & Implementation Research Unit, Health Section, UNICEF, New York, USA
[17] Family and Reproductive Health WHO Tanzania
[18] Centre for Maternal and Newborn Health, Liverpool School of Tropical Medicine, Liverpool, UK
[19] Department for Sexual and Reproductive Health, UNFPA, New York, USA
[20] Department of Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland
[21] Office of Health, Infectious Disease and Nutrition, Bureau for Global Health, United States Agency for International Development, Washington, DC, USA
[22] Jhpiego Baltimore, Baltimore, MD, USA
[23] Department for Informatics, University of Oslo, Oslo, Norway
[24] Division of Reproductive Health, Centres for Disease Control and Prevention (CDC), Atlanta, Georgia, USA
[25] Save the Children, Washington, DC, USA

REFERENCES

1 United Nations. Sustainable Development Goals. 2016. Available: http://www.un.org/sustainabledevelopment/health/. Accessed: 5 April 2018.
2 United Nations. Global Strategy for Women's, Children's and Adolescents' Health, 2016-2030. New York: United Nations; 2015.
3 UN IGME. Levels and Trends in Child Mortality Report 2018. Estimates developed by United Nations inter-agency group for child mortality estimation (UN IGME). New York: United Nations Children's Fund: 2018.
4 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
5 World Health Organization, UNFPA, World Bank Group, United Nations Population Division. Trends in maternal mortality: 1990 to 2015: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. Available: http://www.who.int/reproductivehealth/publications/monitoring/maternal-mortality-2015/en/. Accessed: 4 December 2018.
6 Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, et al. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals. Lancet. 2016;388:3027-35. Medline:27839855 doi:10.1016/S0140-6736(16)31593-8
7 Darmstadt GL, Kinney MV, Chopra M, Cousens S, Kak L, Paul VK, et al. Who has been caring for the baby? Lancet. 2014;384:174-88. Medline:24853603 doi:10.1016/S0140-6736(14)60458-X
8 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
9 World Health Organization. *Every Newborn*: An action plan to end preventable deaths (ENAP). 2014 ISBN 9789241507448.
10 World Health Organization. Strategies towards ending preventable maternal mortality (EPMM). 2015 ISBN 9241508485.
11 World Health Organization. What is the Quality of Care Network? 2017. Available: http://www.who.int/maternal_child_adolescent/topics/quality-of-care/network/en/. Accessed: 19 December 2017.
12 Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? Lancet. 2005;365:891-900. Medline:15752534 doi:10.1016/S0140-6736(05)71048-5

13 Dickson KE, Simen-Kapeu A, Kinney MV, Huicho L, Vesel L, Lackritz E, et al. *Every Newborn*: health-systems bottlenecks and strategies to accelerate scale-up in countries. Lancet. 2014;384:438-54. Medline:24853600 doi:10.1016/S0140-6736(14)60582-1

14 World Health Organization. UNICEF, LSHTM. *Every Newborn* Metrics Report Cards. 2016. Available: https://www.healthynewbornnetwork.org/resource/enap-metrics-cards. Accessed: 21 July 2017.

15 World Health Organization. WHO technical consultation on newborn health indicators: *Every Newborn* Action Plan metrics, Ferney Voltaire, France, 3-5 December 2014. 2015 Contract No.: ISBN: 9789241509381.

16 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:S8. Medline:26391444 doi:10.1186/1471-2393-15-S2-S8

17 Mason E, McDougall L, Lawn JE, Gupta A, Claeson M, Pillay Y, et al. From evidence to action to deliver a healthy start for the next generation. Lancet. 2014;384:455-67. Medline:24853599 doi:10.1016/S0140-6736(14)60750-9

18 Demographic and Health Surveys. The DHS program. 2017. Available: https://dhsprogram.com/. Accessed: 19 December 2017.

19 UNICEF. Multiple Indicator Cluster Surveys, 2017. Available: http://mics.unicef.org/. Accessed: 19 December 2017.

20 Munos MK, Stanton CK, Bryce J. Improving coverage measurement for reproductive, maternal, neonatal and child health: gaps and opportunities. J Glob Health. 2017;7:010801. Medline:28607675 doi:10.7189/jogh.07.010801

21 Marchant T, Tilley-Gyado RD, Tessema T, Singh K, Gautham M, Umar N, et al. Adding content to contacts: measurement of high quality contacts for maternal and newborn health in Ethiopia, North East Nigeria, and Uttar Pradesh, India. PLoS One. 2015;10:e0126840. Medline:26000829 doi:10.1371/journal.pone.0126840

22 World Health Organization. Standards for improving quality of maternal and newborn care in health facilities. Geneva. WHO; 2016.

23 Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. Lancet Glob Health. 2018;6:e1196-252. Medline:30196093 doi:10.1016/S2214-109X(18)30386-3

24 Campbell H, el Arifeen S, Hazir T, O'Kelly J, Bryce J, Rudan I, et al. Measuring coverage in MNCH: challenges in monitoring the proportion of young children with pneumonia who receive antibiotic treatment. PLoS Med. 2013;10:e1001421. Medline:23667338 doi:10.1371/journal.pmed.1001421

25 Stanton CK, Rawlins B, Drake M, dos Anjos M, Cantor D, Chongo L, et al. Measuring coverage in MNCH: Testing the validity of women's self-report of key maternal and newborn health Interventions during the peripartum period in Mozambique. PLoS One. 2013;8:e60694. Medline:23667427 doi:10.1371/journal.pone.0060694

26 Bryce J, Arnold F, Blanc A, Hancioglu A, Newby H, Requejo J, et al. Measuring coverage in MNCH: new findings, new strategies, and recommendations for action. PLoS Med. 2013;10:e1001423. Medline:23667340 doi:10.1371/journal.pmed.1001423

27 Maternal and Child Survival Program. What Data on Maternal and Newborn Health do National Health Management Information Systems include? A review of data elements for 24 low- and lower middle income countries. 2018 May 2018. Report No.

28 Maina I, Wanjala P, Soti D, Kipruto H, Droti B, Boerma T. Using health-facility data to assess subnational coverage of maternal and child health indicators, Kenya. Bull World Health Organ. 2017;95:683-94. Medline:29147041 doi:10.2471/BLT.17.194399

29 Broughton EI, Ikram AN, Sahak I. How accurate are medical record data in Afghanistan's maternal health facilities? An observational validity study. BMJ Open. 2013;3:e002554. Medline:23619087 doi:10.1136/bmjopen-2013-002554

30 Duffy S, Crangle M. Delivery room logbook–fact or fiction? Trop Doct. 2009;39:145-9. Medline:19535748 doi:10.1258/td.2009.080433

31 Chiba Y, Oguttu MA, Nakayama T. Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. Midwifery. 2012;28:329-39. Medline:21684639 doi:10.1016/j.midw.2011.05.005

32 Melberg A, Diallo AH, Storeng KT, Tylleskar T, Moland KM. Policy, paperwork and 'postographs': Global indicators and maternity care documentation in rural Burkina Faso. Soc Sci Med. 2018;215:28-35. Medline:30205276 doi:10.1016/j.socscimed.2018.09.001

33 Ronsmans C, Achadi E, Cohen S, Zazri A. Women's recall of obstetric complications in South Kalimantan, Indonesia. Stud Fam Plann. 1997;28:203-14. Medline:9322336 doi:10.2307/2137888

34 Wall SN, Lee AC, Niermeyer S, English M, Keenan WJ, Carlo W, et al. Neonatal resuscitation in low-resource settings: what, who, and how to overcome challenges to scale up? Int J Gynaecol Obstet. 2009;107 Suppl 1:S47-S64. Medline:19815203 doi:10.1016/j.ijgo.2009.07.013

35 Souza JP, Betran AP, Dumont A, de Mucio B, Gibbs Pickens CM, Deneux-Tharaux C, et al. A global reference for caesarean section rates (C-Model): a multicountry cross-sectional study. BJOG. 2016;123:427-36. Medline:26259689 doi:10.1111/1471-0528.13509

36 Ye J, Betrán AP, Guerrero Vela M, Souza JP, Zhang J. Searching for the optimal rate of medically necessary cesarean delivery. Birth. 2014;41:237-44. Medline:24720614 doi:10.1111/birt.12104

37 Vogel JP, Betrán AP, Vindevoghel N, Souza JP, Torloni MR, Zhang J, et al. Use of the Robson classification to assess caesarean section trends in 21 countries: a secondary analysis of two WHO multicountry surveys. Lancet Glob Health. 2015;3:e260-70. Medline:25866355 doi:10.1016/S2214-109X(15)70094-X

38 Betran AP, Torloni MR, Zhang JJ, Gulmezoglu AM. WHO Statement on Caesarean Section Rates. BJOG. 2016;123:667-70. Medline:26681211 doi:10.1111/1471-0528.13526

39 Ye J, Zhang J, Mikolajczyk R, Torloni M, Gülmezoglu A, Betran A. Association between rates of caesarean section and maternal and neonatal mortality in the 21st century: a worldwide population-based ecological study with longitudinal data. BJOG. 2016;123:745-53. Medline:26331389 doi:10.1111/1471-0528.13592

40 Boatin AA, Schlotheuber A, Betran AP, Moller AB, Barros AJD, Boerma T, et al. Within country inequalities in caesarean section rates: observational study of 72 low and middle income countries. BMJ. 2018;360:k55. Medline:29367432 doi:10.1136/bmj.k55

41 Heywood A, Rohde J. Using information for action - a manual for health workers at facility level. University of Western Cape/HISP/MSH/EQUITY Project.

42 World Health Organization. 2018 Global Reference List of 100 Core Health Indicators - plus health-related SDGs. 2018 World Health Organization 2018. Licence: CC BY-NC-SA 3.0 IGO.

43 Maternal & Child Survival Program. Health Management Information Systems Review - Survey on Data Availability in Electronic Systems for Maternal and Newborn Health Indicators in 24 USAID Priority Countries. 2016.

44 DHIS2. Oslo: Health Information Systems Programme. 2016. Available: https://www.dhis2.org/. Accessed: 25 January 2018.

45 Munos MK, Blanc AK, Carter ED, Eisele TP, Gesuale S, Katz J, et al. Validation studies for population-based intervention coverage indicators: design, analysis, and interpretation. J Glob Health. 2018;8:020804. Medline:30202519 doi:10.7189/jogh.08.020804

46 Blanc AK, Warren C, McCarthy KJ, Kimani J, Ndwiga C. RamaRao S. Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya. J Glob Health. 2016;6:010405. Medline:27231541 doi:10.7189/jogh.06.010405

47 McCarthy KJ, Blanc AK, Warren CE, Kimani J, Mdawida B, Ndwidga C. Can surveys of women accurately track indicators of maternal and newborn care? A validity and reliability study in Kenya. J Glob Health. 2016;6:020502. Medline:27606061 doi:10.7189/jogh.06.020502

48 Blanc AK, Diaz C, McCarthy KJ, Berdichevsky K. Measuring progress in maternal and newborn health care in Mexico: validating indicators of health system contact and quality of care. BMC Pregnancy Childbirth. 2016;16:255. Medline:27577266 doi:10.1186/s12884-016-1047-0

49 The London School of Hygiene & Tropical Medicine. *Every Newborn* Action Plan Metrics Design Workshop for Facility-based Testing of Coverage Metrics, Windsor. 2016. Available: https://www.healthynewbornnetwork.org/hnn-content/uploads/ENAP-Metrics-Facility-based-Workshop-Report_April-2016_FINAL.pdf Accessed: 5 April 2018.

50 World Health Organization. Accountability for Women's and Children's Health Countries Oversight Platform. 2017. Available: http://www.who.int/woman_child_accountability/countries/en/. Accessed: 23 October 2017.

51 Lee AC, Cousens S, Wall SN, Niermeyer S, Darmstadt GL, Carlo WA, et al. Neonatal resuscitation and immediate newborn assessment and stimulation for the prevention of neonatal deaths: a systematic review, meta-analysis and Delphi estimation of mortality effect. BMC Public Health. 2011;11:S12. Medline:21501429 doi:10.1186/1471-2458-11-S3-S12

52 Oza S, Lawn JE, Hogan DR, Mathers C, Cousens SN. Neonatal cause-of-death estimates for the early and late neonatal periods for 194 countries: 2000–2013. Bull World Health Organ. 2015;93:19-28. Medline:25558104 doi:10.2471/BLT.14.139790

53 Vesel L, Bergh A-M, Kerber KJ, Valsangkar B, Mazia G, Moxon SG, et al. Kangaroo mother care: a multi-country analysis of health system bottlenecks and potential solutions. BMC Pregnancy Childbirth. 2015;15:S5. Medline:26391115 doi:10.1186/1471-2393-15-S2-S5

54 Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller A-B, Narwal R, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet. 2012;379:2162-72. Medline:22682464 doi:10.1016/S0140-6736(12)60820-4

55 Vogel JP, Souza JP, Gülmezoglu AM, Mori R, Lumbiganon P, Qureshi Z, et al. Use of antenatal corticosteroids and tocolytic drugs in preterm births in 29 countries: an analysis of the WHO Multicountry Survey on Maternal and Newborn Health. Lancet. 2014;384:1869-77. Medline:25128271 doi:10.1016/S0140-6736(14)60580-8

56 World Health Organisation. World Health Statistics data visualization dashboard. 2015. Available: http://apps.who.int/gho/data/view.sdg.3-2-data-ctry. Accessed.

57 World Health Organisation. Monitoring emergency obstetric care: a handbook, 2009. Available: http://www.who.int/reproductivehealth/publications/monitoring/9789241547734/en/ Accessed: 31 August 2017.

58 EN-BIRTH Study at London School Hygiene and Tropical Medicine Data Compass. 2018. Available: https://doi.org/10.17037/DATA.00000955. Accessed: 4 December 2018.

59 Lindbäck C, Ashish K, Wrammert J, Vitrakoti R, Ewald U, Mílqvist M. Poor adherence to neonatal resuscitation guidelines exposed; an observational study using camera surveillance at a tertiary hospital in Nepal. BMC Pediatr. 2014;14:233. Medline:25227941 doi:10.1186/1471-2431-14-233

60 Ashish KC, Mílqvist M, Wrammert J, Verma S, Aryal DR, Clark R, et al. Implementing a simplified neonatal resuscitation protocol-helping babies breathe at birth (HBB)-at a tertiary level hospital in Nepal for an increased perinatal survival. BMC Pediatr. 2012;12:159. Medline:23039709

61 Ashish KC, Wrammert J, Clark RB, Ewald U, Vitrakoti R, Chaudhary P, et al. Reducing perinatal mortality in Nepal using helping babies breathe. Pediatrics. 2016;137:e20150117. Medline:27225317 doi:10.1542/peds.2015-0117

62 Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009;24:217-28. Medline:19304786 doi:10.1093/heapol/czp010

63 Gwet KL. Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters: Advanced Analytics, LLC, Gaithersburg, Maryland, USA; 2014.

REFERENCES

64 Liu L, Li M, Yang L, Ju L, Tan B, Walker N, et al. Measuring coverage in MNCH: A validation study linking population survey derived coverage to maternal, newborn, and child health care records in rural China. PLoS One. 2013;8:e60762. Medline:23667429 doi:10.1371/journal.pone.0060762

65 World Health Organization. Data quality review: a toolkit for facility data quality assessment. Module 2: Desk review of data quality. Geneva: WHO; 2017.

66 Tripathi V. A literature review of quantitative indicators to measure the quality of labor and delivery care. Int J Gynaecol Obstet. 2016;132:139-45. Medline:26686027 doi:10.1016/j.ijgo.2015.07.014

67 The Lancet Global Health Commission on High Quality Health Systems in the SDG Era. (in press). 2018. Available: https://www.hqsscommission.org/. Accessed: 11 April 2018.

68 Tunçalp Ö, Were W, MacLennan C, Oladapo O, Gülmezoglu A, Bahl R, et al. Quality of care for pregnant women and newborns—the WHO vision. BJOG. 2015;122:1045-9. Medline:25929823 doi:10.1111/1471-0528.13451

69 Goodwin S. Data rich, information poor (DRIP) syndrome: is there a treatment? Radiol Manage. 1996;18:45-9. Medline:10158370

70 Jackson D, McDonald G, Luck L, Waine M, Wilkes L. Some strategies to address the challenges of collecting observational data in a busy clinical environment. Collegian. 2016;23:47-52. Medline:27188039 doi:10.1016/j.colegn.2014.10.001

71 Rawlins B, Christenesen A, Bluestone J. Clinical Observer Learning Resource Package. 2013. Available: http://reprolineplus.org/resources/clinical-observer-learning-resource-package. Accessed: 9 December 2017.

72 Fry M, Curtis K, Considine J, Shaban RZ. Using observation to collect data in emergency research. Australas Emerg Nurs J. 2017;20:25-30. Medline:28169134 doi:10.1016/j.aenj.2017.01.001

73 McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol. 2014;67:267-77. Medline:24275499 doi:10.1016/j.jclinepi.2013.08.015

74 Paradis E, Sutkin G. Beyond a good story: from Hawthorne Effect to reactivity in health professions education research. Med Educ. 2017;51:31-9. Medline:27580703 doi:10.1111/medu.13122

75 World Health Organization. WHO Recommendations on Interventions to Improve Preterm Birth Outcomes. 2015. Available: http://apps.who.int/iris/bitstream/10665/183037/1/9789241508988_eng.pdf. Accessed: 4 Dec 2018.

76 Althabe F, Belizán JM, McClure EM, Hemingway-Foday J, Berrueta M, Mazzoni A, et al. A population-based, multifaceted strategy to implement antenatal corticosteroid treatment versus standard care for the reduction of neonatal mortality due to preterm birth in low-income and middle-income countries: the ACT cluster-randomised trial. Lancet. 2015;385:629-39. Medline:25458726 doi:10.1016/S0140-6736(14)61651-2

77 Lawn JE, Blencowe H, Darmstadt GL, Bhutta ZA. Beyond newborn survival: the world you are born into determines your risk of disability-free survival. Pediatr Res. 2013;74 Suppl 1:1-3. Medline:24240732 doi:10.1038/pr.2013.202

78 Baschieri A, Gordeev VS, Akuze J, Kwesiga D, Blencowe H, Cousens S, 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. doi:10.7189/jogh.09.010901

Chapter 4 - Objective 2: Availability of routine data for tracking use

This chapter assesses labour and delivery ward baseline routine register data availability, quality, and utility in five EN-BIRTH study hospitals.

The chapter was published in August 2020 in BMC Health Services Research. The manuscript was published under a creative commons license (CC BY 4.0) and no further permission are needed.

The published manuscript is included in full below and supplementary material referenced in the paper is available at https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-020-5028-7#Sec24

4.1 List of Figures

4.2 List of Tables

4.3 Citation

Day LT, Gore-Langton GR, et al Labour and delivery ward register data availability, quality, and utility - Every Newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries.
BMC Health Serv Res 2020, **20**(1):737.
https://doi.org/10.1186/s12913-020-5028-7 [24]

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 <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	034282	**Title**	Dr
First Name(s)	Louise Tina		
Surname/Family Name	Day		
Thesis Title	Quality of care and quality of data for hospital births – tension or traction?		
Primary Supervisor	Associate Professor Cally Tann		

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?	Labour and delivery ward register data availability, quality, and utility - Every Newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries. BMC Health Serv Res 2020, 20(1):737. https://doi.org/10.1186/s12913-020-5028-7 Day, L. T., Gore-Langton, G. R., Rahman, A. E., Basnet, O., Shabani, J., Tahsina, T., Poudel, A., Shirima, K., Ameen, S., KC Ashish, Salim, N., Zaman, S. B., Shamba, D., Blencowe, H., Ruysen, H., El Arifeen, S., Boggs, D., Gordeev, V. S., Rahman, Q. S., Hossain, T., Joshi, E., Thapa, S., Poudel, R. P., Poudel, D., Chaudhary, P., Karki, R., Chitrakar, B., Mkopi, N., Wisiko, A., Kitende, A. P., Shirati, M. R., Chingalo, C., Semhando, A. O., Mtei, C., Mwenisongole, V., Bakuza, J. M., Kombo, J., Mbaruku, G. and Lawn, J. E.		
When was the work published?	August 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 jointly designed the objectives with the PI and co-first author. I designed and refined matrices for routine labour ward registers. I designed data cleaning and analysis plan, my co-first author colleague ran the statistical analysis under my leadership. I jointly drafted the manuscript with my co-first author. In particular, I led on results, discussion and conclusion. I am the corresponding author and jointly revised the manuscript in response to peer-review.

SECTION E

Student Signature	[redacted]
Date	

Supervisor Signature	
Date	

Day et al. BMC Health Services Research (2020) 20:737
https://doi.org/10.1186/s12913-020-5028-7

BMC Health Services Research

RESEARCH ARTICLE

Open Access

Labour and delivery ward register data availability, quality, and utility - Every Newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries

Louise Tina Day[1*†], Georgia R. Gore-Langton[1†], Ahmed Ehsanur Rahman[2], Omkar Basnet[3], Josephine Shabani[4], Tazeen Tahsina[2], Asmita Poudel[3], Kizito Shirima[4], Shafiqul Ameen[2], Ashish K.C.[5], Nahya Salim[4,6], Sojib Bin Zaman[2], Donat Shamba[4], Hannah Blencowe[1], Harriet Ruysen[1], Shams El Arifeen[2], Dorothy Boggs[1], Vladimir S. Gordeev[1,7], Qazi Sadeq-ur Rahman[2], Tanvir Hossain[2], Elisha Joshi[8], Sabu Thapa[3], Rajendra Prasad Poudel[9], Durga Poudel[9], Priyanka Chaudhary[3], Rabina Karki[3], Bibek Chitrakar[3], Namala Mkopi[6,10], Anna Wisiko[4], Alodear Patrick Kitende[4], Marystella Revocatus Shirati[4], Christostomus Chingalo[4], Amina Omari Semhando[10], Cleopatra Mtei[10], Victoria Mwenisongole[11], John Mathias Bakuza[4], Japhet Kombo[4], Godfrey Mbaruku[4^] and Joy E. Lawn[1]

Abstract

Background: Countries with the highest burden of maternal and newborn deaths and stillbirths often have little information on these deaths. Since over 81% of births worldwide now occur in facilities, using routine facility data could reduce this data gap. We assessed the availability, quality, and utility of routine labour and delivery ward register data in five hospitals in Bangladesh, Nepal, and Tanzania. This paper forms the baseline register assessment for the *Every Newborn*-Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study.

Methods: We extracted 21 data elements from routine hospital labour ward registers, useful to calculate selected maternal and newborn health (MNH) indicators. The study sites were five public hospitals during a one-year period (2016–17). We measured 1) availability: completeness of data elements by register design, 2) data quality: implausibility, internal consistency, and heaping of birthweight and explored 3) utility by calculating selected MNH indicators using the available data.

(Continued on next page)

* Correspondence: louise-tina.day@lshtm.ac.uk
†Louise Tina Day and Georgia R. Gore-Langton are joint co-authors.
^Godfrey Mbaruku is deceased.
[1]Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene and Tropical Medicine, London, UK
Full list of author information is available at the end of the article

(Continued from previous page)

Results: Data were extracted for 20,075 births. Register design was different between the five hospitals with 10–17 of the 21 selected MNH data elements available. More data were available for health outcomes than interventions. Nearly all available data elements were > 95% complete in four of the five hospitals and implausible values were rare. Data elements captured in specific columns were 85.2% highly complete compared to 25.0% captured in non-specific columns. Birthweight data were less complete for stillbirths than live births at two hospitals, and significant heaping was found in all sites, especially at 2500g and 3000g. All five hospitals recorded count data required to calculate impact indicators including; stillbirth rate, low birthweight rate, Caesarean section rate, and mortality rates.

Conclusions: Data needed to calculate MNH indicators are mostly available and highly complete in EN-BIRTH study hospital routine labour ward registers in Bangladesh, Nepal and Tanzania. Register designs need to include interventions for coverage measurement. There is potential to improve data quality if Health Management Information Systems utilization with feedback loops can be strengthened. Routine health facility data could contribute to reduce the coverage and impact data gap around the time of birth.

Keywords: Maternal, Newborn, Stillbirth, Registers, Birth, Hospital, Routine Health Management Information Systems, Measurement, Indicators

Background

Improving quality of care at birth could save an estimated 3 million lives per year [1, 2]. To drive progress, accurate data are essential, however, the majority of deaths around the time of birth occur in settings with the least information on these deaths, the "inverse data law" [3]. Improving impact and coverage data for action is central to the Sustainable Development Goal (SDG) aspiration of "no-one left behind" [4], the United Nation's Global Strategy for Women's Children's and Adolescents' Health [5], and The *Every Newborn* Action Plan (ENAP). One of five ENAP strategic objectives is to transform metrics and use of data to improve outcomes and track progress towards ending preventable maternal and newborn deaths, including stillbirths [6].

Labour and Delivery (L&D) ward registers are routinely completed by facility health workers and used to track ward admissions and discharges in a parallel system to patient case notes. Birth outcomes, care and interventions for women and babies are also often documented in these registers. However, concerns of poor register data quality in low- and middle- income counties (LMIC), have reduced confidence in full utilization of this data source in Health Management Information Systems (HMIS). As global facility births increase, currently >81%, [7], it is important to reassess the availability and quality of this routine data to help address the current data gap around the time of birth.

Research assessing labour ward register data in LMICs provides some explanation for the scepticism surrounding programmatic use of this source. Maternal and newborn health (MNH) data elements were not consistently available in facility registers in 24 high burden countries [8]. In a rural primary health care context in north eastern Nigeria health workers documented in labour ward registers most completely for birthweight (99%) and woman's age at delivery (97%); documentation was less complete for the composite indicator essential newborn care (82%) and preterm birth (77%) [9]. In two rural Kenyan hospitals, entire labour ward registers were missing for months, and when present many data elements were less than 80% complete; the proportion of data legible/correctly coded/appropriate/recognized ranged from 29 to 100% [10]. In one Ethiopian hospital, among the 20% of births missing from the labour ward register, 91% had received a clinical intervention, thus the register both underestimated total births and interventions [11]. However, the picture is not wholly negative and routine data can be improved. Data quality improvement efforts across 20 L&D wards in South Africa, including data collection training and monthly data reviews, demonstrated increased completeness from 26 to 64%, and accuracy from 37 to 65% [12]. In Rwanda, health system strengthening measures including performance review feedback activities, mentoring, and enhanced supervision led to increased value and ownership of data among health workers [13]. In Zanzibar, quarterly data use workshops with active engagement of data users, grew and improved the HMIS, enhancing staff capacity for information use, presentation and analysis for decision making [14].

The *Every Newborn*-Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study aimed to assess the validity of selected newborn and maternal health care interventions indicators (coverage, content/quality, and/or safety) in hospitals [6] (Fig. 1). Our assessment of existing routine registers in EN-BIRTH study hospitals formed the baseline against which to evaluate any changes in documentation resulting from the presence of researchers in the L&D ward [18].

Aim

This study aimed to assess the availability, quality, and utility of routine data in labour ward registers in five hospitals for 1 year before EN-BIRTH data collection.

Every Newborn-Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study [6]

- The *Every Newborn* Measurement Improvement Roadmap aims to increase the evidence base to inform selection and use of Maternal and Newborn Health indicators in national Health Management Information Systems e.g. District Health Information System (DHIS2).
- Ten core ENAP indicators were prioritised for national and global tracking, including those to measure impact, coverage, and input. Whilst the clarity of these indicator definitions vary, the availability and quality of data is limited for all ENAP core indicators [15-17].
- The EN-BIRTH study aims to test validity of selected newborn and maternal health intervention indicators (coverage/ quality aspects and/or safety) in hospitals.
- The study includes the direct clinical observation of >20,000 births in five Comprehensive Emergency Obstetric and Neonatal Care (CEmONC) public hospitals in Bangladesh, Nepal and Tanzania (Annex, Table 1).
- Observation/ verification data as gold standards will be compared to women's report at exit survey and to routine hospital register data [6].

Fig. 1 Summary of the EN-BIRTH study [6, 15–17]

Objectives

To evaluate routine hospital labour ward registers for 21 selected maternal and newborn data elements (Annex, Table 2):

1. Data availability: measure completeness in relation to register design.
2. Data quality: assess implausibility, internal consistency, and birthweight heaping.
3. Data utility: cross-tabulate and transform available count data to coverage/impact indicators.

Methods

Study settings

The five EN-BIRTH study sites are public hospitals in high burden Sub-Saharan African and South Asian LMIC settings and implementing the selected MNH interventions. Two hospitals in Bangladesh (BD) - Maternal and Child Health Training Institute (MCHTI) Azimpur, and Kushtia District Hospital; one in Nepal (NP) - Pokhara Academy of Health Sciences; and two in Tanzania (TZ) - Temeke Regional Hospital and Muhimbili National Hospital (Annex, Table 1) [6].

Data collection

Data elements/ count data required to calculate selected priority global MNH indicators were identified ($n = 21$) (Annex, Table 2). The data elements were extracted from routine hospital labour ward registers by trained researchers. In Bangladesh, for Caesarean section births, additional data from routine "Operation Registers" were extracted and included in the dataset. All data were extracted at the end of the 12 month study period, prior to EN-BIRTH observational data collection; in Tanzania and Bangladesh, 1st January 2016 – 31st December 2016, and in Nepal 1st April 2016 - 31st March 2017. Data were extracted for all births in Bangladesh and Nepal, and a 20% simple random monthly sample in Tanzania, due to the high case volume. Data were directly entered into customized databases in Tanzania and Nepal and into Microsoft Excel (Version 2007) from register photographs in Bangladesh.

Data analysis

Data were analyzed in Stata 15 (StataCorp, 2017, College Station, TX).

The following data analysis methods were applied for each study objective (Table 1):

Objective 1: availability of labour ward register data elements

Availability of data elements: mapped across the five hospital registers by classifying the register design into one of three categories:

Specific column allotted for data element e.g. Column title: "Uterotonic for third stage of labour", documentation requires "Yes" or "No".

Non-specific column allotted for data element e.g. Column title: "Drugs given". Uterotonic drugs are documented alongside other drugs e.g. analgesics, antibiotics etc.

No column allotted for the data element in the routine register (but may be recorded elsewhere e.g. patient case notes).

Completeness of data element recording: the percentage of total births recorded in the register with data recorded for the data element (Table 1). Whilst data completeness is often considered a data quality dimension, for the purpose of this study, we consider it separately [20].

Table 1 Terms and definitions of data availability, quality and utility assessed by study objectives. EN-BIRTH Baseline Register Analysis

Study Objective	Term	Definition
Objective 1: Data Availability	**Availability**	A measure of whether the specific data element is recorded in the register in relation to register design [8, 19]
	Completeness	A measure of the proportion of entries in the register that had any data recorded for the specified data element for: Numerator – women or babies for whom intervention received/not received or health outcome of interest recorded Denominator – mothers delivered or babies born [20].
Objective 2: Data Quality	**Implausibility**	A measure of whether individual data are outside pre-defined ranges of biological credibility.
	Heaping	A measure of the proportion of values falling on specific values (e.g. for birthweight on 2000g or 2500g) or rounded (i.e. ending in "00" or "50").
	Internal consistency	A measure of whether the observed relationship between related data elements is as expected [20, 21].
Objective 3: Data Utilization	**Utility**	The transformation of count data into indicators by using them as numerators and denominators or cross-tabulation.
	Coverage	Number of individuals receiving an intervention or service (numerator), from among the hospital population in need of the intervention or service (denominator) [6].
	Impact	A measure of the extent to which health status of the facility target population is being achieved (e.g. maternal and newborn mortality); used for global tracking [22] .

Objective 2: quality of labour ward register data

Three facets of data quality were assessed for a subset of data elements:

Implausibility: The proportion of extreme or unlikely values were calculated for three data elements: birthweight (< 350g or > 6000g), gestational age (< 20 weeks or > 44 weeks), and women's age (< 10 years or > 49 years).

Birthweight heaping and rounding were assessed in three ways. First, the proportion of birthweights rounded to 100g (ending "00") or 50g (ending "50") was calculated.

Second, rounded weight values (e.g. 2500g) were calculated as a proportion of all weights within the adjacent 250g brackets (e.g. 2250-2750g). Third, the heaping ratio of the rounded weight value (e.g. 2500g) relative to the number of weights within the adjacent 250g brackets, excluding the rounded value (e.g. 2250–2499 plus 2501-2749g) was calculated.

Internal consistency of data elements with expected associations were examined by cross tabulation [23]: birth outcome and breastfeeding and [1] baby outcome at discharge [20, 24].

Objective 3: utility

To explore potential use of available MNH data elements, indicators (coverage, impact, and others of programmatic relevance) (Annex Table 3) were calculated with 95% confidence intervals (95% CI) using the register count data as numerators and denominators (Annex Table 4). For indicators using live births as the denominator, our calculations include only recorded live births in both numerator and denominator. Birth outcomes were further disaggregated by birthweight [6]. The effect of birthweight heaping on the Low Birth Weight (LBW) rate was explored by reallocating 50% of the birthweights recorded as exactly 2500g to the LBW (<2500g) category.

Ethical approval

Institutional review boards in all sites, and at the London School of Hygiene and Tropical Medicine granted ethical approval and administrative data sharing agreements were in place.

Results

Objective 1: availability of data

Data were extracted for 20,075 babies in total, 8544 in Nepal, 7111 in Bangladesh, and 4420 in Tanzania (Table 2). Across the five hospitals, 396 babies were either twins or triplets.

The labour ward registers were named: "Delivery Register" in Azimpur BD which differed from "Delivery Register" in Kushtia, BD. Both Bangladesh hospitals used "Operation Registers" for Caesarean births (Table 2). "Obstetric Register" is the national standardized register in Pokhara NP. Both Tanzanian hospitals use the national standardized HMIS labour ward register and additional data elements are captured in Muhimbili within a informal perinatal register known locally as "Midwifery Book".

The labour ward register designs are summarized in Fig. 2, shaded in black if the data element is not captured. Labour ward registers contained ten of 21 data elements in Azimpur BD, 11 in Kushtia BD, 15 in Temeke TZ, 17 in Muhimbili TZ, and 12 in Pokhara NP (Table 2).

Table 2 Availability of data in labour ward/ operation theatre registers in five EN-BIRTH study hospitals at baseline, total births recorded *n*=20,075

	Bangladesh				Nepal	Tanzania		Total
	Azimpur Tertiary		Kushtia District		Pokhara Regional	Temeke Regional	Muhimbili National	
Register Name	Labour Ward Register	Operation Register	Labour Ward Register	Operation Register	Obstetric Register	HMIS Labour Ward Register	HMIS Labour Ward Register & Midwifery Book	
Total number of babies extracted in register	1415	3253	1742	701	8544	2560	1860	20,075
Babies of multiple births (twins, triplets)	26	60	93	6	76	121	14	396
Total data elements in register	18	21	19	21	31	43	45	
Total data elements of 21 requested	10		11		12	15	17	

Across the five hospital registers, these 21 data elements were recorded in 65 separate columns, of which 61 columns were "specific" for the data element and four columns were "non-specific". High completeness (>80%) was found for 85.2% of the 61 specific columns compared to 25.0% of the four non-specific colums.

Data Element	Numerator	Denominator	Data Element Completeness* n (%)				
			Bangladesh		Nepal	Tanzania	
			Azimpur Tertiary	Kushtia District	Pokhara Regional	Temeke Regional	Muhimbili National
			Total births n= 4,668	Total births n= 2,443	Total births n= 8,544	Total births n= 2,560	Total births n= 1,860
Intervention Count Data (necessary for coverage)							
Uterotonics for prevention of postpartum haemorrhage (PPH)	✓		0 (0)**	1,731 (70.9)**		2,560 (100)	1,860 (100)
Breastfeeding early initiation in first hour after birth	✓					2,545 (99.4)	0 (0)
Baby resuscitation - stimulation	✓					2,560 (100)	1,860 (100)
Baby resuscitation - bag-mask-ventilation	✓					2,560 (100)	1,860 (100)
Antenatal Corticosteroid	✓						
Chlorhexidine applied to cord	✓					**Not applicable**	**Not applicable**
Health Outcome Count Data (necessary for impact)							
Birth Outcome (liveborn/ stillborn)	✓	✓	4,657 (99.8)	1,920 (78.6)	8,544 (100)	789 (30.8)	1,838 (98.8)
Stillbirth type (macerated/ fresh)	✓			0 (0)	60 (45.5)	61 (100)	104 (100)
Birthweight	✓	✓	4,638 (99.4)	1,614 (66.1)	8,509 (99.6)	2,560 (100)	1,856 (99.8)
Gestational age (completed weeks)	✓	✓			8,148 (95.4)		1,793 (96.4)
Woman's outcome at discharge from L&D	✓		1,406 (30.1)	1,727 (70.7)	8,544 (100)	2,548 (99.5)	1,860 (100)
Baby's outcome at discharge from L&D	✓		1,398 (29.9)	1,725 (70.6)	8,543 (99.9)	2,508 (98.0)	1,860 (100)
Other Count Data							
Woman's age	✓	✓	4,650 (99.6)	2,380 (97.3)	8,544 (100)	2,559 (99.9)	1,849 (99.4)
Sex of Baby	✓	✓	4,659 (99.8)	1,925 (78.8)	8,544 (100)	2,560 (100)	1,860 (100)
Date of Delivery	✓		4,668 (100)	2,443 (100)	8,544 (100)	2,560 (100)	1,860 (100)
Time of Delivery	✓		4,668 (100)	2,443 (100)	8,533 (99.8)	2,560 (100)	1,860 (100)
Mode of delivery	✓	✓	4,668 (100)	2,435 (99.7)	8,536 (99.9)	2,560 (100)	1,860 (100)
Retained placenta	✓					2,560 (100)	1,860 (100)
Woman's Estimated blood loss in ml	✓						
Date of discharge from L&D	✓				8,544 (100)		1,860 (100)
Time of discharge from L&D	✓						

* number of babies for whom data element recorded

**Register designed for "intervention not given" to be left intentionally blank, so not a measure of true completeness

KEY	
	Register design - allotted specific column for data element
	Register design - allotted non-specific column for data element
	Register design - no column allotted

Fig. 2 Availability and completeness of data elements in labour ward registers, by intervention, health outcome and other count data coded by register design. EN-BIRTH Baseline Register Analysis n=20,075

Availability of intervention count data

Uterotonics for prevention of postpartum haemorrhage (PPH) was not captured in the register in Pokhara NP. In both the Bangladesh hospital registers uterotonics were recorded in a non-specific column “medicine given”, left blank for “not given”, so true completeness could not be calculated. The Tanzanian register had a specific column headed “Oxytocin, Ergometrine, or Misoprostol” and completeness was 100% in both Muhimbili and Temeke TZ (Fig. 2).

Immediate breastfeeding was not captured in either Bangladesh or Nepal registers. The Temeke TZ register had 99.4% completeness, but the same data element in Muhimbili TZ was not completed.

Newborn resuscitation data [25] were also only recorded in the Tanzanian registers, within a specific column “Helping Babies Breathe” coded: “1” suction, “2” stimulation, “3” bag-mask-ventilation and "no" for no resuscitation. Completeness in both hospitals was 100%.

Availability of health outcome data

The baby’s outcome at birth, live birth or stillborn, was documented in a non-specific column in Pokhara and a specific column in the other four hospitals. Completeness of recording was 30.8% in Temeke TZ, 78.6% in Kushtia BD and above 98.0%, for the remaining hospitals (Fig. 2).

Data elements for stillbirth (SB) timing (antepartum/intrapartum) were not available in any register and proxy measures (fresh/macerated) were allotted a specific column in Kushtia BD (completeness 0%) and both Tanzanian hospitals (completeness 100%) and a non-specific column in Pokhara NP (completeness 45.5%) (Fig. 2).

Birthweight was documented in a specific column in all five registers, completeness was >99% in four hospitals and 66.1% in Kushtia BD. Stratifying birthweight completeness by outcome showed that in Bangladesh stillbirths were much less complete, 50.0% compared to 100% for live births in Azimpur BD and 26.3% compared to 87.8% for live births in Kushtia BD (Fig. 3). Gestational age was allotted a specific column only in Pokhara NP and in the additional perinatal register in Muhimbili TZ, completeness was >95% in both. Women’s and baby’s condition at discharge from the L&D ward had specific columns in all registers, completeness was <80% in Azimpur BD and Kushtia BD and >99.5% in Muhimbili and Temeke TZ and Pokhara NP (Fig. 2).

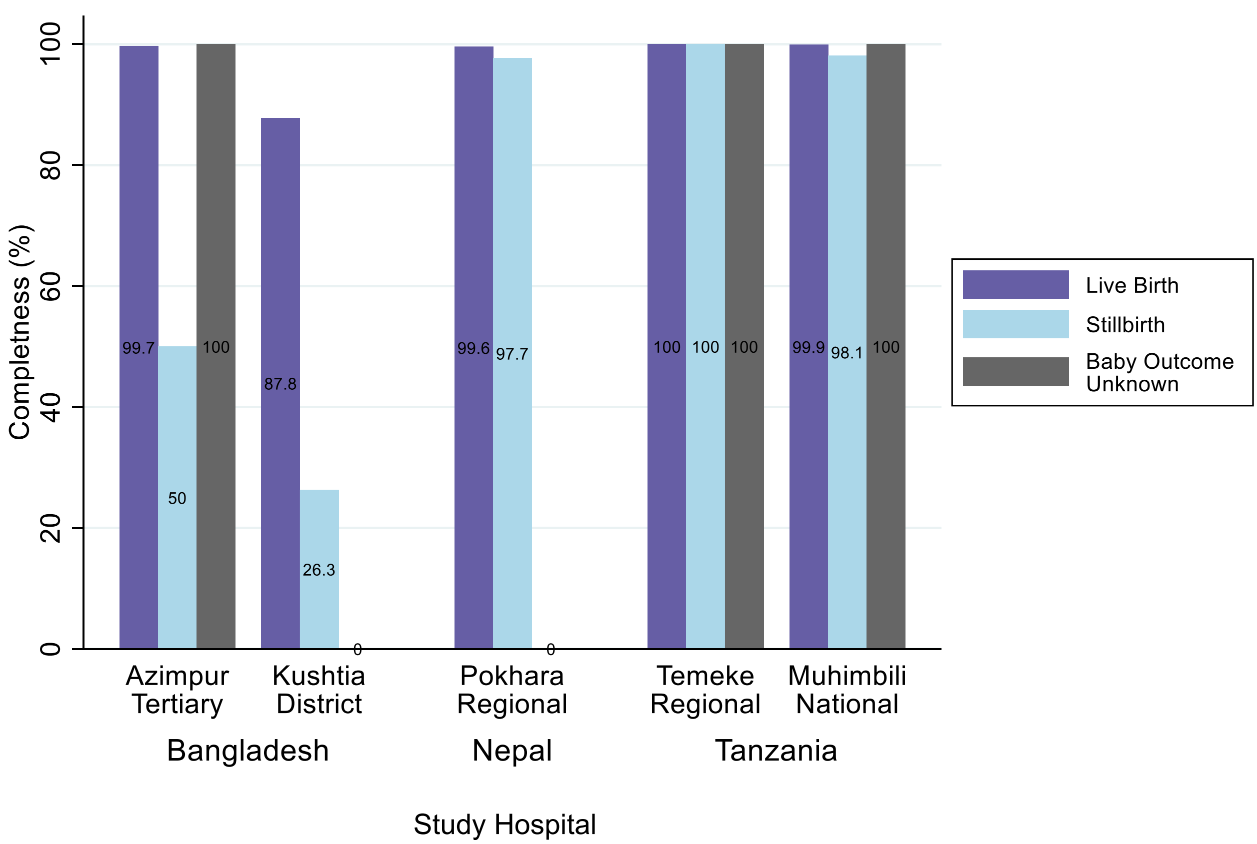


Fig. 3 Completeness (%) of recording of birthweight data stratified by birth outcome (live birth/stillbirth/birth outcome unknown). EN-BIRTH Baseline Register Analysis n=19,177

Availability of other count data

All five labour ward registers, had specific columns and >95% completeness for: woman's age, date and time of birth, and mode of birth (Fig. 2). Baby's sex was >95% complete in all registers except Kushtia BD, 78.8%.

Antenatal corticosteroids, chlorhexidine application to cord (implemented in BD and NP), and time of discharge from L&D ward were not allotted columns in any register. Date of discharge from L&D ward was only allotted a specific column in Pokhara NP and Muhimbili TZ, 100% complete (Fig. 2).

Objective 2: quality of data

Implausibility

The proportion of implausible values was low across hospitals – for birthweight 0–1.2%, for gestational age 0–0.2%, and woman's age 0–0.2%.

Heaping

Birthweight data were heaped in all five hospitals, in four registers more than 74% of weights were rounded to the nearest 100g (Fig. 4, Annex, Table 5). The heaping ratio was highest at 2.00 in Kushtia for 3000g, i.e. twice as many babies were recorded as exactly 3000g than at any other weight within the two adjacent 250g brackets (2750–2999 and 3000-3249g) (Annex, Table 5). For the critical 2500g LBW cut-off weight, among all babies with a birthweight within range 2250-2749g, the babies with birthweight recorded as exactly 2500g was very high; 60.7% in Kushtia BD, 43.5% in Pokhara NP, 42.0% in Temeke TZ, 19.5% in Azimpur BD and 18.9% in Muhimbili TZ (Annex, Table 5).

Internal consistency

Babies with birth outcome "stillbirth" should also be recorded "died" for baby outcome at L&D ward discharge. In Bangladesh, the non-specific discharge term "unwell" was recorded for 96.2% stillbirths ($n = 25$) in Azimpur and 94.6% ($n = 106$) in Kushtia. The discharge term "alive" or "well" was used for 5.4% ($n = 6$) in Kushtia BD, 16.3% ($n = 17$) in Muhimbili TZ and 6.6% ($n = 4$) in Temeke TZ. Stillbirths recorded as having been breastfed were 11.5% ($n = 7$) in Temeke TZ.

Objective 3: utilization of data

Intervention coverage indicators

Coverage indicators calculated from the available register count data are shown in Table 3. Uterotonics coverage to prevent PPH ranged from 19.5% of live births in Temeke TZ to 89.1% of live births in Kushtia BD.

The neonatal resuscitation coverage true denominator is "babies in need of resuscitation", and as this was not available in these routine registers, a surrogate of total births (live births plus stillbirths) was used. Bag-mask ventilation (BMV) was received by 4.1% ($n = 105$) of total births in Temeke TZ (Table 3), among these 25.7% ($n = 27$) were live births, 1.9% (n = 2) were fresh stillbirths, and 72.4% ($n = 74$) had birth outcome missing. Among babies receiving BMV in Temeke TZ only 24.8% ($n = 26$) were recorded to have also received stimulation.

Impact indicators

The facility stillbirth rate (SBR) was lowest at 7.4 in Azimpur BD and highest at 55.9 in Muhimbili TZ per 1000 total births, Table 3. The fresh SBR ranged from 2.2 in Pokhara NP to 18.9 in Muhimbili TZ per 1000 total births.

Low Birth Weight (LBW) prevalence, ranged from 10.3% in Temeke TZ to 22.6% in Muhimbili TZ, Table 3. The adjusted LBW rate (after re-allocating 50% of babies with a recorded birthweight of exactly 2500g to the LBW category) increased the LBW prevalence by 1.7% in Muhimbili TZ and by 7.2% in Kushtia BD (Table 4).

Cross-tabulating categorical birthweight with outcome (live birth/ fresh stillbirth/ macerated stillbirth) showed 62.4% ($n = 212$) of total stillbirths and 49.3% ($n = 41$) of fresh stillbirths were categorised LBW compared to 13.1% ($n = 2225$) of live births (Table 5).

The preterm birth rate (number of babies < 37 weeks per 100 live births) was 4.5% in Pokhara NP and 32.5% in Muhimbili TZ.

Maternal deaths were recorded in Pokhara NP ($n = 3$), Muhimbili TZ ($n = 1$), and Temeke TZ ($n = 5$), with none in Azimpur BD or Kushtia BD. Thus facility Maternal mortality ratio (MMR) before discharge from L&D ward ranged from zero in both Bangladesh hospitals to 137.4 per 100,000 live births in Temeke TZ (Table 3). The neonatal mortality rate (NMR) before discharge from L&D ward ranged from zero in Azimpur BD and Kushtia BD to 7.5 per 1000 live births in Muhimbili TZ.

Other indicators of programmatic relevance

The proportion of hospital births to adolescents (11–19 years) ranged from 4.8% in Muhimbili TZ to 18.6% in Azimpur BD. Ratio of male:female babies was highest in Pokhara NP at 118:100, Table 3.

Caesarean section rate, using a live birth denominator [26, 27], was 43.4%, ranging from 10.4% in Kushtia to 70.2% in Azimpur BD, Table 3. As 69 stillbirths (20.2% of total stillbirths) were also delivered by Caesarean, if these were included in the denominator [27], the Caesarean rate would decrease overall to 37.3%.

Discussion

This is the largest multi-country study we are aware of in LMICs to assess labour ward register data availability,

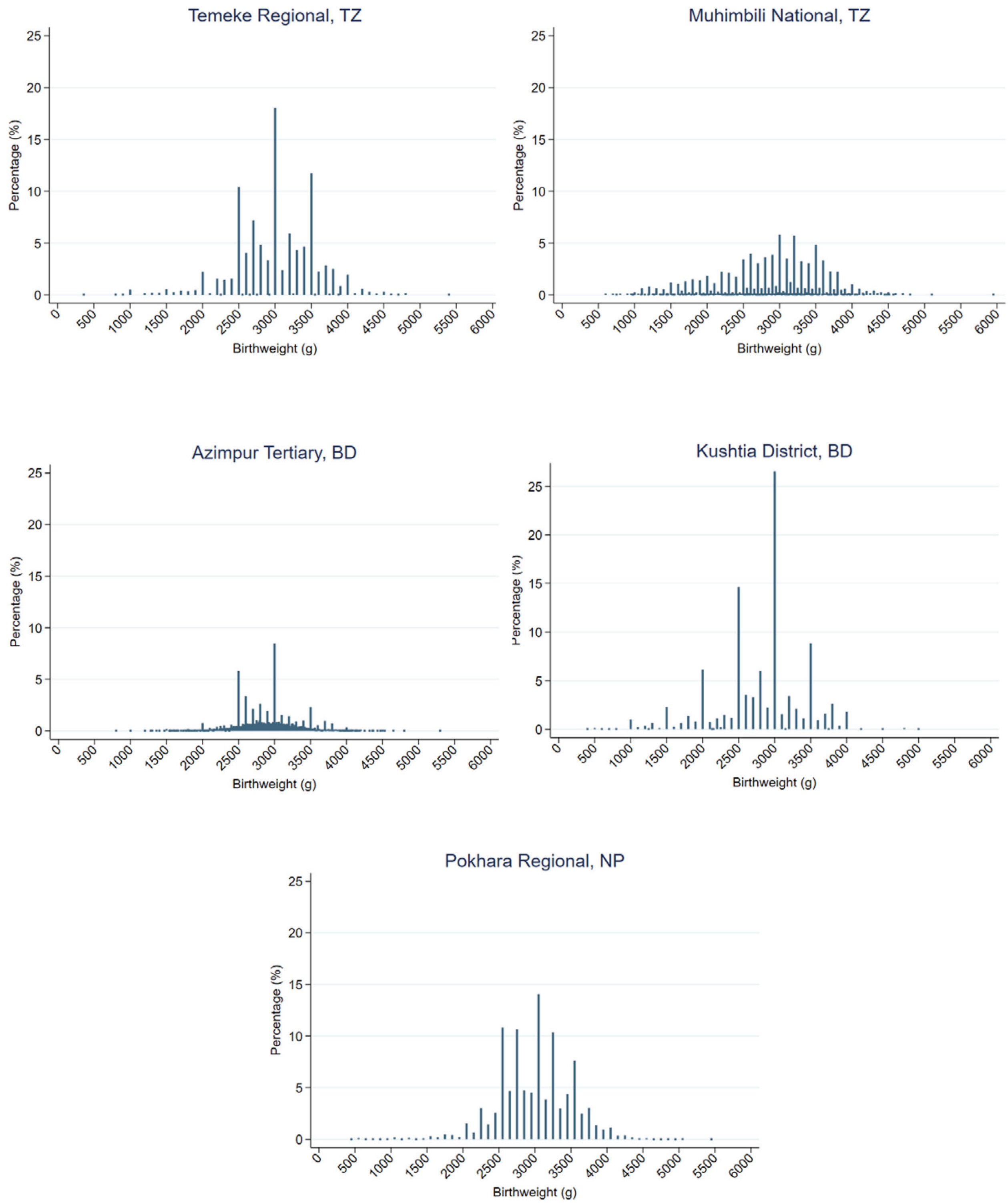


Fig. 4 Distribution of plausible birthweights recorded in each of the five EN-BIRTH study hospital labour ward registers. EN-BIRTH Baseline Register Analysis, n = 19,140

Table 3 Examples of data utilization - transformation of count data into indicators - EN-BIRTH registers baseline analysis n = 20,075

		Hospitals				
		Bangladesh		Nepal	Tanzania	
		Azimpur Tertiary	Kushtia District	Pokhara Regional	Temeke Regional	Muhimbili National
Indicator Type	*Denominator: Recorded Total Births*	4,668	2,443	8,544	2,560	1,860
	Denominator: Recorded Live births	4,623	1,802	8,412	728	1,734
	Recorded Stillbirths	34	11	132	61	104
	Recorded birth outcome missing (babies excluded from live birth denominator)	45	641	132	1832	126
	Indicator					
Coverage	Women receiving uterotonics for the prevention of PPH during the third stage of labour/ live births % (95% CI)	0	89.1 (87.5, 90.4)*		19.5 (16.8, 22.6)*	34.3 (32.1, 36.6)*
	Babies breastfeeding initiated/live births % (95% CI)				84.6 (81.7, 87.2)*	0
	Babies receiving stimulation for resuscitation/total births % (95% CI)				3.4 (2.3, 5.0)	0.4 (0.2, 0.8)
	Babies receiving BMV for resuscitation/total births % (95% CI)				4.1 (2.6, 5.4)	0
Impact	Stillbirth rate / 1,000 total births (95% CI)	7.4 (5.3, 10.3)	47.6 (39.8, 56.9)	15.4 (13.0, 18.3)	23.8 (18.6, 30.5)	55.9 (46.3, 67.3)
	Fresh stillbirth rate / 1,000 total births (95% CI)		0	2.2 (1.4, 3.5)	10.9 (7.6, 15.8)	18.9 (14.4, 27.3)
	Low Birth Weight rate (<2500g)/ 100 live births (95% CI)	12.6 (11.6, 13.6)*	18.5 (16.7, 20.5)*	10.5 (9.8, 11.2)*	10.3 (8.2, 12.7)*	22.6 (20.7, 24.6)*
	Preterm birth rate (<37 weeks)/100 live births (95% CI)			4.5 (4.1, 5.0)*		32.5 (30.3, 34.8)*
	Term birth rate (37-41 weeks)/ 100 live births (95% CI)			87.4 (86.7, 88.1)*		61.6 (59.3, 63.9)*
	Post-term birth rate (42 weeks or more)/ 100 live births (95% CI)			3.6 (3.2, 4.0)*		2.4 (1.7, 3.6)*
	Maternal Mortality Ratio at discharge from L&D ward/ 100,000 live births (95% CI)	0	0	23.8 (2.8, 85.86)*	137.4 (3.48, 762.94)*	57.7 (1.46, 320.9)*
	Neonatal Mortality Rate at discharge from L&D ward / 1,000 live births (95% CI)	0	0	0.8 (0.3, 1.7)*	5.49 (1.50, 14.01)*	7.5 (4.0, 12.7)*
Other	Adolescent birth rate-facility (11-19 years)/100 total births (95% CI)	18.6 (17.5, 19.7)	7.5 (6.6, 8.7)	13.3 (12.6, 14.1)	12 (14.8, 17.7)	4.8 (4.1, 6.2)
	Neonatal Sex Ratio (# Males per 100 Females)	107	102	118	100	106
	Caesarean section Rate/ 100 live births (95% CI)	70.2 (68.8, 71.5)*	10.4 (9.0, 11.9)*	23.9 (23.0, 24.8)*	14.4 (12.0, 17.2)*	64.0 (61.6, 66.2)*
	Mothers with retained placenta occurred/100 total births (95% CI)				0.3 (0.1, 0.6)	0.2 (0.05, 0.5)

* For indicators which use live births as the denominator: calculations include only live births in the numerator given the incomplete recording of birth outcome data (denominator) in all facilities
Grey cells indicate data element required to calculate indicator not present in the Labour Ward Register

quality, and utility. Hospital registers are key tools used to collect individual data for aggregation and transmission up the HMIS data pyramid [6]. Data extracted from five CEmONC hospitals show that a large amount of data are being collected in labour ward registers. The calculation of MNH coverage and impact indicators require the availability of specific data elements for use as numerators and/or denominators, yet none of the labour ward registers contained all 21 selected data elements. Data for outputs, outcomes, and impact measurement were more widely available, than for intervention coverage. Only the Tanzanian registers captured most of the selected interventions and gestational age was only captured by the Nepal register and the additional register in Muhimbili TZ.

The Performance of Routine Information System Management (PRISM) framework identifies complexity and design as technical factors in routine health information systems performance [28]. The register designs were different between countries, and within country in Bangladesh. Data were captured from the additional perinatal register in Muhimbili TZ and from operation registers for babies born by Caesarean in Bangladesh, highlighting further complexity in multiple recording systems.

Whether a specific column was allotted for the data element related to completeness of recording. Across all five hospitals, a much lower proportion of non-specific columns had high levels of completeness than did specific columns (25.0% versus 85.2%). However, there were other examples of low completeness within specific columns (e.g. 0% SB type Kushtia BD) or high completeness within non-specific columns (e.g. 100% birth outcome, Pokhara NP) highlighting that technical factors alone are necessary but not sufficient for data availability.

Table 4 Adjusted and unadjusted Low Birth Weight rate - EN-BIRTH register baseline analysis n=17,033

	Bangladesh		Nepal	Tanzania	
	Azimpur Tertiary	Kushtia District	Pokhara Regional	Temeke Regional	Muhimbili National
Unadjusted Low Birth Weight rate (< 2500 g) / 100 live births (95% CI)	12.6 (11.6, 13.6)	18.5 (16.7, 20.5)	10.5 (9.8, 11.2)	10.3 (8.2, 12.7)	22.6 (20.7, 24.6)
Adjusted Low Birth Weight rate (< 2500 g) / 100 live births (95% CI)	15.6 (14.5, 16.6)	25.7 (23.6, 27.9)	15.9 (15.2, 16.7)	15.0 (12.6, 17.8)	24.4 (22.4, 26.4)
Increase in Low Birth Weight rate (< 2500 g)	2.93	7.20	5.41	4.67	1.73

Table 5 Birth outcomes cross-tabulated by categorical birthweight, pooled data all EN-BIRTH hospitals baseline register analysis n=17,595

Birth Outcome	Total Births[a]	Categorical Birthweight (g) n (%)				
		≤999g	1000–1999g	2000–2499g	2500–3999g	≥4000g
Total Live births (%)	17,033	27 (0.2)	508 (3.0)	1690 (9.9)	14,402 (84.6)	361 (2.4)
Total Stillbirths (%)[b]	340	29 (8.5)	125 (36.8)	58 (17.1)	119 (35.0)	9 (2.6)
Fresh Stillbirths (%)[c]	83	4 (4.8)	27 (32.5)	10 (12.0)	39 (47.0)	3 (3.6)
Macerated Stillbirths (%)[2]	139	11 (7.9)	66 (47.5)	24 (17.3)	36 (25.9)	2 (1.4)

[a]babies with a recorded birth outcome and a plausible birthweight recorded (*n* = 17,595)
[b]includes all stillbirths from all five hospitals, [c]Information on fresh or macerated stillbirths presented where available (i.e. for 100% of SB in Tanzania, 45.5% of SB in Nepal and no SB in Bangladesh)

Other factors associated with data completeness included mode of birth, e.g. in Kushtia BD 97.1% of babies for whom birth outcome was missing were delivered by Caesarean. This finding is similar to previous research in Ethiopia, where a high proportion of babies not recorded in the register had required a clinical intervention [11]. Previous studies have highlighted the low value placed on stillbirths and the resultant data gaps [29–31] and similarly we found birthweight data completeness was lower for these babies in Bangladesh.

Incomplete count data affect indicator calculation results. When intervention coverage numerators are missing, rates will appear low, e.g. when only the outcome "unwell" was recorded and no maternal/ newborn deaths, the MMR, NMR at discharge from L&D ward and SBR may be inappropriately low - zero in Azimpur BD and Kushtia BD during this study. When denominators such as birth outcome are missing (e.g. for 69.2% of babies in Temeke TZ and 21.4% in Kushtia BD) many indicators which use live birth as the denominator will be adversely affected. For example, coverage of breastfeeding in Temeke TZ would be 292.5% had the numerator not been restricted to include only babies in the live births denominator (i.e. exluding babies with a birth outcome unknown). Calculating coverage using a total birth denominator instead of a clinical need denominator, as we have done for neonatal resuscitation, requires benchmarked rates for meaningful tracking across hospitals.

Alternatively, data completeness may be high, but if inaccurate, coverage will be falsely low or high. For example, uterotonic coverage was apparently low in the Tanzanian registers. These data are handwritten using a Swahili abbreviation "N" for "Ndiyo" (Yes) or "H" for "Hapana" (No) which can be hard to distinguish and possibly incorrectly extracted. Numbered coding systems may be helpful when the design is simple e.g. the "Helping Babies Breathe" column in Tanzania. Blank data elements in the register can mean either "incomplete" or a true zero, as in the Bangladesh register design, which if not differentiated can introduce another source of data inaccuracy [32].

Beyond data completeness, our data quality evaluation showed variable results. Birthweight rounding and heaping were substantial across all hospitals. If a baby whose true birthweight of 2470g is rounded to 2500g the LBW rate will be underestimated – in our model by up to 7.2%. Both analogue and digital scales were used for birthweight across the five hospitals which may contribute to rounding. Additional EN-BIRTH analyses are exploring accuracy and processes of birthweight measurement [33, 34]. In these high mortality burden countries, very large variation in hospital mortality rates may suggest data quality issues; Muhimbili TZ had a stillbirth rate almost eight times higher that Azimpur BD. The EN-BIRTH mixed-methods study aims to test validity of these indicator measurements against the gold standard of observation data.

Barriers and enablers to recording in routine hospital registers are being explored in the wider EN-BIRTH study [6]. Quality of register data is affected by HMIS input determinants described by the PRISM framework [28] including technical, organisational and behavioural factors. Factors known to negatively impact routine data quality include poor use of data, lack of feedback, low management support, lack of health worker confidence, low motivation, lack of competence and low perceived utility of routine recording tasks [14, 28, 35]. Health worker training and supportive supervision regarding the importance of routine recording around the time of birth could improve data quality for all babies, especially stillbirths. Innovations to increase health worker data utilization skills could also help sustain improvements in data recording as the purpose of these activities is recognised. Previous studies have demonstrated large gains from such efforts [12–14, 35, 36].

Further research is needed to understand the effect of labour ward register design on data quality, the impact of increased reporting burden on frontline health workers, and ways to optimize the utility of register data whilst reducing duplication. Standardized and harmonized registers with inclusion of an appropriate number of selected key data elements need evaluating against

What is already known:

- >81% births are in hospitals and labour ward registers have potential to close the gap for coverage data around the time of birth.
- Research regarding availability and quality of register data has been sparse, especially in LMIC where the majority of deaths occur.
- 5.1 million deaths of babies around the time of birth (2.5 million neonatal deaths and 2.6 million stillbirths), majority are preventable
- Global goals to reduce these deaths require quality data to track progress and drive action.

What this research adds and what next:

- First large (>20,000 babies) multi-country, multi-site detailed assessment of labour ward registers for data completeness and quality.
- Large quantities of numerator and denominator count data are currently available in the labour ward registers, which could be used now.
- Impact indicators to inform programme decision making can be calculated with currently available data, notably intrapartum stillbirth rates, with cross tabulation by weight.
- Coverage indicator data currently available in some hospitals include uterotonics, breast-feeding and neonatal bag-mask-ventilation.
- Improvements in data quality is required for certain data elements e.g. birthweight heaping
- Health workers invest time in documentation in labour ward registers, yet these data are often under-used and currently not reaching their potential to address the data gap around the time of birth.
- Large scale validation studies are needed in order to have confidence in these data, the EN-BIRTH study is a response to this need.

Fig. 5 Summary figure: Labour and delivery ward register data, what is already known, what the EN-BIRTH baseline register study adds and what next

registers that contain large numbers of data elements. At the five study hospitals, all documentation at the health worker/ mother and baby interface was in paper-based routine registers. As electronic platforms increase, the effect of digitization on data quality around the time of birth requires attention from the source data to the top of the data pyramid [6].

Utilizing the EN-BIRTH multi-country study hospitals, a strength of this research is the large amount of data extracted (20,075 births), providing the first in-depth and multi-country analysis of routine labour ward register data. However, EN-BIRTH study hospitals may not be entirely representative of routine recording practices in facilities at different levels of the health system nor in other LMIC settings. Some EN-BIRTH hospitals have been involved in previous research, thus routine recording may be better than typical. Conversely, staff workload in these high-volume CEmONC hospitals could reduce data quality. We were unable to assess whether all babies born on the labour ward were recorded in the register, nor the relationship between staff levels and data quality. Research in facilities at different levels of the health system is required before wider conclusions can be drawn. Furthermore, evidence of completeness and quality of register data do not necessarily correlate with accurate aggregation and reporting from the facility up the HMIS data pyramid, therefore research is required to review quality of facility-reported data used for district/national/global tracking of MNH indicators.

The *Every Newborn* strategic objective to transform measurement aims to increase availability and quality of data to use for action. Unless all births occur in hospitals, facility data will overestimate population coverage. However, as hospital births increase (globally now 81% [7]) this data source is increasingly valuable. Improving facility data quality would also benefit wider health system indicators (e.g. immunization coverage) which currently use census projection data. Whilst household surveys are useful to provide information on contact with MNH services at a population level in LMICs, they have been shown to be less valid for the capture of content or coverage of interventions around the time of birth, hence new strategies incorporating multiple data sources, including register data, are required [32, 37–39]. Clarity is needed on the calculation of Caesarean section rates; the current denominator recommendation is live births, but more than 10% of stillbirths in this dataset were delivered by Caesarean [26, 27]. In our study, inclusion of stillbirths in the denominator as well as the numerator increased Caesarean rate by nearly 6%. We propose Caesarean rates be calculated using hospital total births and stratified for live births and stillbirths.

Findings presented here could be used now by decision makers at various levels of the health system. In the hospital for quality improvement e.g. if no fresh stillbirths are being resuscitated this could lead to review of guidelines, practice and/or documentation. Birthweight data were readily available in all five hospitals, so LBW rate reporting, one of WHO 100 core health indicators could be improved [40]. If the LBW rate is implausibly low, hospitals might use the same data to improve quality of birthweight measurement. Using birthweight categories and birth outcome data, we found differences between live births and stillbirths, e.g. the differential growth of stillbirths where 36.8% weighed 1000-1999 g compared to 3.0% of live births. Our study showed that 50.6% of fresh stillbirths had a normal birthweight yet died, this metric could also be tracked to improve quality of care.

Changes in register recording practices during the EN-BIRTH study will be explored [6]. Importantly, the EN-BIRTH observational study will further validate indicators from labour ward register data to inform use in HMIS and areas of focus to further improve data availability and quality.

Data used for action is foundational for tracking progress towards global goals for every woman and every child to survive and thrive [4, 5]. As data is used, data quality and overall HMIS performance improves [14, 28]. As data quality improves, coverage and outcome indicators can more confidently be used for action to track progress and drive change.

Conclusions

This study shows that large amounts of specific MNH data elements are currently available in routine labour ward registers in five hospitals in Bangladesh, Nepal and Tanzania. Data quality varied when assessed for completeness and implausibility. There is potential to improve the quality of available data if HMIS utilization with feedback loops can be strengthened. By advancing routine health facility data for use, labour ward registers can contribute to much needed regular coverage and impact measurements around the time of birth (Fig. 5). To optimize care around the critical time of birth, labour ward register data offer huge potential to be improved and used.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12913-020-5028-7.

Additional file 1: Table S1. Summary information on country context and five EN-BIRTH study hospitals. **Table S2.** MNH data elements ($n = 21$) extracted from the labour ward and OT ward register of each of five EN-BIRTH hospitals. **Table S3.** Definitions and examples of coverage and impact indicators, including numerator and denominator. **Table S4.** Indicators calculated for potential data utilization from routine register numerator and denominatorIndicators calculated for potential data utilization from routine register numerator and denominator. **Table S5.** Heaping of birthweight data recorded labour ward/operation registers in EN-BIRTH study hospitals.

Abbreviations

ACS: Antenatal corticosteroids; BD: Bangladesh; BMV: bag-mask ventilation; CIFF: Children's Investment Fund Foundation; DHIS2: District Health Information Systems 2; EmONC: Emergency Obstetric and Newborn Care; ENAP: *Every Newborn* Action Plan now branded as *Every Newborn*; EN-BIRTH: *Every Newborn*-Birth Indicators Research Tracking in Hospitals study; EPMM: Ending Preventable Maternal Mortality; HMIS: Health Management Information Systems; icddr, b: International Centre for Diarrheal Disease Research, Bangladesh; IHI: Ifakara Health Institute; L&D: Labour and Delivery; LBW: Low Birth Weight; LMIC: Low and Middle Income Countries; LSHTM: London School of Hygiene and Tropical Medicine; MMR: Maternal Mortality Ratio; MNH: Maternal and Newborn Health; MUHAS: Muhimbili University of Health and Allied Sciences; NMR: Neonatal Mortality Rate; NP: Nepal; PPH: Post-partum haemorrhage; PRISM: Performance of Routine Information System Management; RHIS: Routine Health Information Systems; SB: Stillbirth; SBR: Stillbirth Rate; TZ: Tanzania; UNICEF: United Nations International Children's Emergency Fund; WHO: World Health Organization

Acknowledgements

We credit the inspiration of the late Godfrey Mbaruku. We thank the health workers who completed the registers. We thank the National Advisory Groups in Bangladesh, Nepal and Tanzania and the EN-BIRTH Expert Advisory group for their collaboration in the EN-BIRTH study. We thank collaborating colleagues in Nepal: Nisha Rana, Jagat Jeevan Ghimire and Sushil Karki and in LSHTM: Simon Cousens, Sarah G Moxon, Angela Baschieri for their input to the EN-BIRTH study. Many thanks to Claudia DaSilva, Fion Hay, Alegria Perez, Sadie Sareen, Adeline Herman, Veronica Ulaya, Mohammad Raisul Islam and Ziaul Haque Shaikh, and Bhula Rai for their administrative support.

Authors' contributions

The full EN-BIRTH study was conceptualized by JEL, who acquired the funding and led the overall design with support from HR including inputs from collaborating colleagues and HB, the EN-BIRTH Expert Advisory Group, and during a multi-stakeholder research design workshop. The IHI and MUHAS in Tanzania hosted the study implementation workshop. HR, GRGL and DB led development of data collection tools with input from each of the three country research teams (AER, TT, EJ, OB, AKC, RP, JJ, NS, DS, JS, KS, GM). GRGL, HR and DB led development of training materials with country coordinators (Bangladesh: TT and AER, Nepal: NR, Tanzania: NS). Data was collected by BD: SA, SBZ; NP: AP, PC, BC, RK; TZ: AW, AK, MS, CC, AS, CM, VM, JB, JK. VSG led with development of the register matrices with major inputs from LTD OB, SBZ and JS. The cleaning and analysis of data was carried out by GRGL with LTD. The manuscript was drafted with equal contributions by LTD and GRGL: with major inputs from JEL, HR, HB, DB, VSG, AER, EJ, JS, TT, OB, NS, AKC, SBZ, SA, QSR, TH, SEA, AP, ST, RP, DP, DS, JS, KS, NM, AW, APK, MRS, CC, AOS, CM, VM, JMB, JK. All authors reviewed and finalized the manuscript.

Funding

The Children's Investment Fund Foundation (CIFF) are the main funder of the EN-BIRTH Study and funding is administered via The London School of Hygiene and Tropical Medicine. CIFF were involved, with many other collaborative partners, in the protocol development and implementation workshops. CIFF have not been involved in data collection or analysis. The Swedish Research Council specifically funded the Nepal site through Golden Community but was not involved in study design, data collection or analysis.

Availability of data and materials

The datasets used and/or analysed during the current study are available on LSHTM Data Compass repository, https://datacompass.lshtm.ac.uk/955/.

Ethics approval and consent to participate

This study was granted ethical approval by Institutional Review Board of London School of Hygiene And Tropical Medicine (ref: 11780), Institutional Review Board of Ifakara Health Institute, Tanzania (ref: 032–2016), Institutional Review Board of Muhimbili University of Health and Allied Sciences, Tanzania (ref: 2016-10-21/AEC/Vol.XI/310), Institutional Review Board of National Institute for Medical Research, Tanzania (ref: NIMR/HQ/R.8a/Vol.IX/2394), Institutional Review Board of International Centre for Diarrhoeal Disease Research, Bangladesh (ref: PR-16055), Ethical Review Committee of International Centre for Diarrhoeal Disease Research, Bangladesh (ref: PR-16055) and Nepal Health Research Council (ref: 187 /2O16).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene and Tropical Medicine, London, UK. [2]Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh. [3]Golden Community, Lalitpur, Nepal. [4]Department of Health Systems, Impact Evaluation and Policy, Ifakara Health Institute, Dar es Salaam, Tanzania. [5]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden. [6]Muhimbili University of Health and Allied Sciences (MUHAS), Dar es Salaam, Tanzania. [7]Institute of Population Health Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK. [8]LifeLine Nepal, Baneshwor, Kathmandu, Nepal. [9]Yagiten Pvt Ltd, Kathmandu, Nepal. [10]Muhimbili National Hospital (MNH), Dar es Salaam, Tanzania. [11]Temeke Regional Referral Hospital, Dar es Salaam, Tanzania.

Received: 12 July 2019 Accepted: 24 February 2020
Published online: 12 August 2020

References
1. Bhutta ZA, Das JK, Bahl R, Lawn JE, Salam RA, Paul VK, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? Lancet. 2014;384(9940):347–70.
2. 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(9938): 189–205.
3. Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? Lancet. 2005;365(9462):891–900.
4. United Nations. Sustainable Development Goals 2016 [28/11/2018]. Available from: https://www.un.org/sustainabledevelopment/health/.
5. United Nations. Global strategy for Women's, Children's and Adolescents' health, 2016–2030. New York: United Nations; 2015 Available from: https://www.who.int/life-course/partners/global-strategy/en/.
6. Day LT, Ruysen H, Gordeev VS, Gore-Langton GR, Boggs D, Cousens S, et al. "Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. J Glob Health. 2019;9(1):010902.
7. World Health Organization. WHO GHO and Data Portal for Global Strategy. 2019. http://apps.who.int/gho/data/node.gswcah. Accessed 7 Jan 2020.
8. Maternal Child and Survival Program. What data on maternal and newborn health do National Health Management Information Systems Include? 2018. Available from: https://www.mcsprogram.org/resource/what-data-on-maternal-and-newborn-health-do-national-health-management-information-systems-include/.
9. Bhattacharya AA, Allen E, Umar N, Usman AU, Felix H, Audu A, et al. Monitoring childbirth care in primary health facilities: a validity study in Gombe state, northeastern Nigeria. J Glob Health. 2019;9(2):020411.
10. Chiba Y, Oguttu MA, Nakayama T. Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. Midwifery. 2012;28(3):329–39.
11. Duffy S, Crangle M. Delivery room logbook - fact or fiction? Trop Dr. 2009; 39(3):145–9.
12. Mphatswe W, Mate KS, Bennett B, Ngidi H, Reddy J, Barker PM, et al. Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. Bull World Health Organ. 2012;90(3):176–82.
13. Wagenaar BH, Hirschhorn LR, Henley C, Gremu A, Sindano N, Chilengi R, et al. Data-driven quality improvement in low-and middle-income country health systems: lessons from seven years of implementation experience across Mozambique, Rwanda, and Zambia. BMC Health Serv Res. 2017;17(3):830.
14. Braa J, Heywood A, Sahay S. Improving quality and use of data through data-use workshops: Zanzibar, United Republic of Tanzania. Bull World Health Organ. 2012;90(5):379–84.
15. World Health Organization. Every Newborn: An action plan to end preventable deaths (ENAP) 2014.
16. Mason E, McDougall L, Lawn JE, Gupta A, Claeson M, Pillay Y, et al. From evidence to action to deliver a healthy start for the next generation. Lancet. 2014;384(9941):455–67.
17. 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(2):S8.
18. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol. 2014;67(3):267–77.
19. MEASURE Evaluation. Routine data quality assessment tool - user manual. 2017. Available from: https://www.measureevaluation.org/resources/publications/ms-17-117.
20. World Health Organization. Data quality review: a toolkit for facility data quality assessment. Module 2. Desk review of data quality. 2017.
21. Munos MK, Stanton CK, Bryce J. Improving coverage measurement for reproductive, maternal, neonatal and child health: gaps and opportunities. J Glob Health. 2017;7(1):010801.
22. World Health Organization. Child Health and Development 2014 [18/12/2018]. Available from: http://www.emro.who.int/child-health/research-and-evaluation/indicators/All-Pages.html.
23. World Health Organization. Global Strategy for Women's, Children's and Adolescents' Health (2016–2030). 2016.
24. Munos MK, Blanc AK, Carter ED, Eisele TP, Gesuale S, Katz J, et al. Validation studies for population-based intervention coverage indicators: design, analysis, and interpretation. J Glob Health. 2018;8(2):020804.
25. American Academy of Pediatrics. Guide for Implementation of Helping Babies Breathe (HBB): Strengthening neonatal resuscitation in sustainable programs of essential newborn care, vol. 2011; 2011. Report No.
26. MEASURE Evaluation. Cesarean sections as a percent of all births [cited 2018 29/11]. Available from: https://www.measureevaluation.org/prh/rh_indicators/womens-health/sm/cesarean-sections-as-a-percent-of-all-births.
27. World Health Organisation. Indicator Metadata Registry: Births by caesarean section (%) [05/04/2019]. Available from: http://apps.who.int/gho/data/node.wrapper.imr?x-id=68.
28. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009;24(3):217–28.
29. Lawn JE, Yakoob MY, Haws RA, Soomro T, Darmstadt GL, Bhutta ZA. 3.2 million stillbirths: epidemiology and overview of the evidence review. BMC Pregnancy Childbirth. 2009;9(Suppl 1):S2.
30. Lumbiganon P, Panamonta M, Laopaiboon M, Pothinam S, Patithat N. Why are Thai official perinatal and infant mortality rates so low? Int J Epidemiol. 1990;19(4):997–1000.
31. Froen JF, Gordijn SJ, Abdel-Aleem H, Bergsjo P, Betran A, Duke CW, et al. Making stillbirths count, making numbers talk - issues in data collection for stillbirths. BMC Pregnancy Childbirth. 2009;9:58.
32. Bhattacharya AA, Umar N, Audu A, Felix H, Allen E, Schellenberg JRM, et al. Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: a case study from Gombe state, Nigeria. PLoS One. 2019;14(1):e0211265.
33. Gladstone ME, Ogillo K, Shamba D, Gore-Langton GR, Day LT et al. Opportunities to improve quality of routine Birthweight measurement – a mixed methods study in Temeke hospital, Tanzania. BMC Pregnancy & Childbirth [in press].
34. Kong S, Day LT, Zaman SB, Peven K, Salim N, Sunny AK, et al. Birthweight: EN-BIRTH multi-country study informing measurement quality in surveys and routine systems. BMC Pregnancy & Childbirth [in press].
35. Sæbø JI, Mesheck Moyo C, Nielsen P. Promoting transparency and accountability with district league tables in Sierra Leone and Malawi. Health Policy Technol. 2018;7(1):35–43.
36. Wagenaar BH, Gimbel S, Hoek R, Pfeiffer J, Michel C, Manuel JL, et al. Effects of a health information system data quality intervention on concordance in Mozambique: time-series analyses from 2009–2012. Popul Health Metrics. 2015;13(1):9.
37. Seoane G, Castrillo M, O'Rourke K. A validation study of maternal self reports of obstetrical complications: implications for health surveys. Int J Gynecol Obstet. 1998;62(3):229–36.
38. Stanton CK, Rawlins B, Drake M, dos Anjos M, Cantor D, Chongo L, et al. Measuring coverage in MNCH: testing the validity of women's self-report of key maternal and newborn health interventions during the peripartum period in Mozambique. PLoS One. 2013;8(5):e60694.

39. Liu L, Li M, Yang L, Ju L, Tan B, Walker N, et al. Measuring coverage in MNCH: a validation study linking population survey derived coverage to maternal, newborn, and child health care records in rural China. PLoS One. 2013;8(5):e60762.
40. Geneva: World Health Organization. 2018 Global Reference List of 100 Core Health Indicators (plus health-related SDGs). 2018.

Publisher's Note

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

Chapter 5 - Objective 3: Validity of indicator data elements

This chapter assesses validity of measurement of newborn and maternal health-care coverage in EN-BIRTH study hospitals

The chapter was published in December 2020 in The Lancet Global Health. The manuscript was published under a creative commons license (CC BY-NC-ND 4.0 license) and no further permissions are needed.

The published manuscript is included in full below and supplementary material referenced in the paper is available at https://www.thelancet.com/cms/10.1016/S2214-109X(20)30504-0/attachment/6a5bf618-706c-4fff-be9b-ae03aad93973/mmc1.pdf

5.1 List of Figures

5.2 List of Tables

5.3 Citation

Day LT, Rahman QS, et al.
Assessment of the validity of the measurement of newborn and maternal health-care coverage in hospitals (EN-BIRTH): an observational study.
The Lancet Global Health 2021; **9**(3): E267-79.
https://doi.org/10.1016/S2214-109X(20)30504-0.[25]

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	034282	**Title**	Dr
First Name(s)	Louise Tina		
Surname/Family Name	Day		
Thesis Title	Quality of care and quality of data for hospital births – tension or traction?		
Primary Supervisor	Associate Professor Cally Tann		

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?	Every Newborn-BIRTH observational study to assess validity of newborn and maternal coverage measurement in hospitals. The Lancet Global Health 2021; 9(3): E267-79. https://doi.org/10.1016/S2214-109X(20)30504-0.16 Day, Louise Tina, Rahman, Qazi Sadequr, Rahman, Ahmed Ehsanur, Salim, Nahya, Kc, Ashish, Ruysen, Harriet, Tahsina, Tazeen, Masanja, Honorati, Basnet, Omkar, Gore-Langton, Georgia R., Zaman, Sojib Bin, Shabani, Josephine, Jha, Anjani Kumar, Gordeev, Vladimir Sergeevich, Ameen, Shafiqul, Shamba, Donat, Jha, Bijay, Boggs, Dorothy, Hossain, Tanvir, Shirima, Kizito, Bastola, Ram Chandra, Peven, Kimberly, Siddique, Abu Bakkar, Mbaruku, Godfrey, Paudel, Rajendra, Baschieri, Angela, Hossain, Aniqa Tasnim, Kong, Stefanie, Paudel, Asmita, Ahmed, Anisuddin, Cousens, Simon, El Arifeen, Shams, Lawn, Joy E.,		
When was the work published?	December 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	**Yes**

		review?	

*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 managed the collaborative multi-country/site study from since mid-way through data collection. Specific contributions included designing the duplicate observation/ extraction process and the variable renaming and collated the code book. I jointly designed the analysis plan with the PIs and co-authors. I co-led the data cleaning process with my co-first author. I designed the variable matrix for population-level coverage comparison and individual-level validity testing. The statistical analysis was run by my joint first author. I conceptualised the additional research questions to stratify by mode of birth (all sites) and the natural experiment of original and revised registers for the Bangladesh study hospitals. I drafted the manuscript, conceptualised the validity ratios and designed all figures. I incorporated collaborative inputs from PIs, country teams, national/global expert advisory groups.I am the corresponding author, lead on the two rounds of manuscript revision in response to peer-review, including co-ordinating collaborative inputs from co-authors. I managed proofs and presentation of results at dissemination activities. I presented results at dissemination activities including preliminary results at MoNITOR, study results launch.

SECTION E

Student Signature	[redacted]
Date	

Supervisor Signature	[redacted]
Date	

Assessment of the validity of the measurement of newborn and maternal health-care coverage in hospitals (EN-BIRTH): an observational study

Louise Tina Day, Qazi Sadeq-ur Rahman*, Ahmed Ehsanur Rahman, Nahya Salim, Ashish KC, Harriet Ruysen, Tazeen Tahsina, Honorati Masanja, Omkar Basnet, Georgia R Gore-Langton, Sojib Bin Zaman, Josephine Shabani, Anjani Kumar Jha, Vladimir Sergeevich Gordeev, Shafiqul Ameen, Donat Shamba, Bijay Jha, Dorothy Boggs, Tanvir Hossain, Kizito Shirima, Ram Chandra Bastola, Kimberly Peven, Abu Bakkar Siddique, Godfrey Mbaruku†, Rajendra Paudel, Angela Baschieri , Aniqa Tasnim Hossain, Stefanie Kong, Asmita Paudel, Anisuddin Ahmed, Simon Cousens, Shams El Arifeen‡, Joy E Lawn‡, EN-BIRTH validation collaborative group§, EN-BIRTH Expert Advisory Group§*

Summary

Background Progress in reducing maternal and neonatal deaths and stillbirths is impeded by data gaps, especially regarding coverage and quality of care in hospitals. We aimed to assess the validity of indicators of maternal and newborn health-care coverage around the time of birth in survey data and routine facility register data.

Methods Every Newborn-BIRTH Indicators Research Tracking in Hospitals was an observational study in five hospitals in Bangladesh, Nepal, and Tanzania. We included women and their newborn babies who consented on admission to hospital. Exclusion critiera at admission were no fetal heartbeat heard or imminent birth. For coverage of uterotonics to prevent post-partum haemorrhage, early initiation of breastfeeding (within 1 h), neonatal bag-mask ventilation, kangaroo mother care (KMC), and antibiotics for clinically defined neonatal infection (sepsis, pneumonia, or meningitis), we collected time-stamped, direct observation or case note verification data as gold standard. We compared data reported via hospital exit surveys and via hospital registers to the gold standard, pooled using random effects meta-analysis. We calculated population-level validity ratios (measured coverage to observed coverage) plus individual-level validity metrics.

Findings We observed 23 471 births and 840 mother–baby KMC pairs, and verified the case notes of 1015 admitted newborn babies regarding antibiotic treatment. Exit-survey-reported coverage for KMC was 99·9% (95% CI 98·3–100) compared with observed coverage of 100% (99·9–100), but exit surveys underestimated coverage for uterotonics (84·7% [79·1–89·5]) *vs* 99·4% [98·7–99·8] observed), bag-mask ventilation (0·8% [0·4–1·4]) *vs* 4·4% [1·9–8·1]), and antibiotics for neonatal infection (74·7% [55·3–90·1] *vs* 96·4% [94·0–98·6] observed). Early breastfeeding coverage was overestimated in exit surveys (53·2% [39·4–66·8) *vs* 10·9% [3·8–21·0] observed). "Don't know" responses concerning clinical interventions were more common in the exit survey after caesarean birth. Register data underestimated coverage of uterotonics (77·9% [37·8–99·5] *vs* 99·2% [98·6–99·7] observed), bag-mask ventilation (4·3% [2·1–7·3] *vs* 5·1% [2·0–9·6] observed), KMC (92·9% [84·2–98·5] *vs* 100% [99·9–100] observed), and overestimated early breastfeeding (85·9% (58·1–99·6) *vs* 12·5% [4·6–23·6] observed). Inter-hospital heterogeneity was higher for register-recorded coverage than for exit survey report. Even with the same register design, accuracy varied between hospitals.

Interpretation Coverage indicators for newborn and maternal health care in exit surveys had low accuracy for specific clinical interventions, except for self-report of KMC, which had high sensitivity after admission to a KMC ward or corner and could be considered for further assessment. Hospital register design and completion are less standardised than surveys, resulting in variable data quality, with good validity for the best performing sites. Because approximately 80% of births worldwide take place in facilities, standardising register design and information systems has the potential to sustainably improve the quality of data on care at birth.

Funding Children's Investment Fund Foundation and Swedish Research Council.

Introduction

Investment in global health measurement has particularly focused on outcomes, notably deaths. Accurate data are urgently needed to track the progress towards the Sustainable Development Goals (SDGs) to end the annual 5·1 million preventable stillbirths and newborn deaths, plus 0·3 million maternal deaths by 2030.[1–4] Despite nearly 80% of births worldwide now being in health-care facilities,[5] many avoidable deaths occur, notably intrapartum stillbirths and in preterm newborn babies.[6,7] Many evidence-based, high-impact interventions for maternal and newborn health are

Lancet Glob Health 2021; 9: e267–79

Published **Online** December 14, 2020
https://doi.org/10.1016/S2214-109X(20)30504-0

This online publication has been corrected. The corrected version first appeared at thelancet.com/lancetgh on February 16, 2021

See **Comment** page e221

*Joint first authors

†Deceased

‡Joint senior authors

§Group members listed at the end of the Article

Maternal, Adolescent, Reproductive, & Child Health Centre, London School of Hygiene & Tropical Medicine, London, UK (L T Day MRCPCH, H Ruysen MSc, G R Gore-Langton MSc, V S Gordeev PhD, D Boggs MSc, K Peven MPH, A Baschieri PhD, S Kong MSc, S Cousens MA, J E Lawn FMedSci)**; Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh** (Q Sadeq-ur Rahman MSc, A E Rahman MPH, T Tahsina MSc, S B Zaman MPH, S Ameen MBBS, T Hossain MSc, A B Siddique MSc, A T Hossain MSc, A Ahmed MSc, S E Arifeen DPH)**; Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania** (N Salim PhD)**; Department of Health Systems, Impact Evaluation and Policy, Ifakara Health Institute, Dar es Salaam, Tanzania** (N Salim, H Masanja PhD, J Shabani MSc, D Shamba MSc, K Shirima HDiploma, G Mbaruku PhD)**; International Maternal and Child Health,**

Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden (A KC PhD); Research Division, Golden Community, Lalitpur, Kathmandu, Nepal (O Basnet BBS, R Paudel MSc, A Paudel BN); Nepal Health Research Council, Kathmandu, Nepal (Anjani Kumar Jha MD, Bijay Jha BCom); The Institute of Population Health Sciences, Queen Mary University of London, London, UK (V S Gordeev); Pokhara Academy of Health Science, Pokhara, Nepal (R C Bastola); Ministry of Health and Population, Kathmandu, Nepal (R C Bastola); and Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, King's College London, London, UK (K Peven)

Correspondence to: Dr Louise T Day, Maternal, Adolescent, Reproductive, & Child Health Centre, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK **louise-tina.day@lshtm.ac.uk**

Research in context

Evidence before this study

Increasing coverage and quality of care around the time of birth is fundamental to reducing 5·3 million maternal, fetal, and newborn deaths and disability every year and achieving the Sustainable Development Goals. National and global tracking of maternal and newborn health has mostly used measures of contact coverage (eg, skilled birth attendance), with little attention paid to indicators of content or quality of care. We searched Ovid MEDLINE for articles published in English after Jan 1, 2010, with the search terms "valid" AND "maternal or newborn or neonate or labour or childbirth or delivery or intrapartum" AND "coverage or indicator or measure or track or numerator or denominator" and restricted to low-income and middle-income settings. Of the 598 papers identified, 17 met our inclusion criteria, among which most observational studies focused on indicator measurement validation in survey with two for routine register data. We found no additional relevant documents in the grey literature. Previous studies have reported low validity for measures around the time of birth. High-priority newborn interventions, such as neonatal bag-mask ventilation and kangaroo mother care (KMC), have not undergone validity testing in surveys or register data. Early initiation of breastfeeding has only been validated in surveys. WHO and UNICEF's *Every Newborn* Action Plan prioritised validation of data on indicators of maternal and newborn coverage of care to improve measurement of content of care and especially through routine data systems.

Added value of this study

The *Every Newborn*-Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study was done in five hospitals in Bangladesh, Nepal, and Tanzania, and aimed to assess the validity of data for five indicators of maternal and newborn health-care coverage in two data sources: routine facility registers and women's report at exit survey. These data were compared against a gold standard of direct observation or case note verification. EN-BIRTH was about 10 times larger than previous facility-based validation studies and included more than 25 000 cases—either observed births or KMC or newborn admission case notes. This study is the first to assess the validity of routine register data for most of these indicators, and of survey data for the hospital newborn care indicators. We found that exit survey reports had low accuracy for uterotonic coverage and for early initiation of breastfeeding, consistent with previous smaller studies. Population-based household surveys already capture early initiation of breastfeeding coverage but research is needed to improve accuracy. We also found survey-reported data on bag-mask ventilation had low accuracy and neonatal infection treatment with antibiotics had low sensitivity among the target group. KMC coverage was accurately reported at exit survey by women who had practised KMC. Thus, further work is required to assess whether KMC report remains reliable after the typical household survey recall period of 2–5 years, and also to ascertain the extent to which mothers who did not practise KMC misreport having done so. Routine registers in some hospitals were found to be highly complete, but accuracy varied between hospitals, even with the same register design. Register accuracy for uterotonics was excellent in two hospitals, and KMC sensitivity was excellent in two hospitals and good in two hospitals. One hospital had good register accuracy for bag-mask ventilation. For early initiation of breastfeeding, register accuracy was poor in all four study hospitals with a register column.

Implications of all the available evidence

For care around the time of birth, surveys are important for estimating population-based contact coverage and family-led behaviours such as breastfeeding, but our findings do not support adding questions about clinical interventions to surveys, with the possible exception of admission for KMC. Given that approximately 80% of the world's births are now in facilities, routine registers can provide data on intervention delivery more rapidly and at lower cost than surveys. Optimising register design, filling, and flow into national routine information systems requires investment in implementation research. Caesarean birth negatively affected the accuracy of survey-reported and register-recorded data. Further research is required regarding the measurement implications of increasing caesarean section rates. Reliable data are necessary, but not sufficient to improve care around the time of birth—health workers, policy makers, and politicians must also value and use these data to drive change.

delivered in health-care facilities, yet gaps in quality of care and gaps in data result in missed opportunities.[7,8] The SDGs aspire to achieve universal health coverage, which will be difficult without addressing the crucial measurement gaps regarding effective coverage and quality around the time of birth.[9]

For the **Demographic and Health Surveys programme** see https://dhsprogram.com/

For **UNICEF's Multiple Indicator Cluster Surveys** see http://mics.unicef.org/

Coverage is defined as the proportion of individuals receiving an intervention or practice (numerator) among the population in need of that intervention or practice (denominator). National and global tracking of coverage has primarily focused on survey measurement of contacts with the health system (eg, institutional birth), with few indicators capturing content (eg, interventions) or quality of care.[10,11] Maternal and newborn coverage of care in low-income and middle-income countries (LMICs) is mainly tracked through large-scale population-based household surveys, notably the Demographic and Health Surveys (DHS) programme and UNICEF's Multiple Indicator Cluster Surveys (MICS). Although these surveys importantly measure population-based contact coverage, previous research has found mixed validity for the content of care around the time of birth (eg, breastfeeding).[12–16] DHS has done more than 400 surveys in 90 countries. With over 400 questions in the DHS core questionnaire, focus has been on the need to validate additional questions before adding. Overall, few survey indicators have been validated, and those relating to clinical care for small and ill newborn babies. have not previously been validated.[17]

The shift to most births worldwide being in facilities, paired with rapid improvements for routine national health management information systems (HMIS) including digitalisation, have potential to transform measurement of coverage and quality of care for women and newborn babies, including in high-burden settings.[17] Health workers document details of admitted women and newborn babies in routine facility registers, usually in parallel to individual patient case notes. These registers are the primary source for aggregate data that flow into routine HMIS. The quality of HMIS data, or mistrust of quality, impedes full use by policy makers.[18–20] Previous studies of routine paper-based registers in facilities in LMICs have reported on data availability,[21,22] but only two small observational studies have examined some aspects of register measurement validity for care around the time of birth.[23,24]

Transforming measurement and use of data to track coverage and quality of care is one of five strategic objectives in the *Every Newborn* Action Plan, led by WHO and UNICEF, implemented in more than 92 countries.[25] Validation of coverage measurement for interventions and practices (content of care) was prioritised.[26] Core indicators regarding high-impact maternal and newborn care recommended by WHO were selected as outlined in the measurement improvement roadmap for *Every Newborn*[17,26] and Ending Preventable Maternal Mortality monitoring framework.[27] The *Every Newborn*-Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study was designed to address these evidence gaps by assessing measurement validity for high-priority indicators of newborn and maternal health coverage to inform their use in routine HMIS and population-based surveys for national and global tracking. None of these indicators had previously been validated in routine register data and few in survey data.[28]

Criterion validity testing compares measurement against an objective gold standard to assess whether indicators measure what they intend to, and can provide accurate evidence to inform programmes.[29,30] The EN-BIRTH study protocol outlined four objectives: (1) to assess the validity of numerator measurement, (2) to compare denominators, (3) to evaluate content and quality of care, and (4) to assess barriers and enablers to routine register measurement.[28] We report results on the EN-BIRTH study's first objective for five prioritised core coverage indicators (appendix p 4). Validity testing of other core or additional indicators will be reported separately.

Methods

Study design and participants

EN-BIRTH was a mixed-methods observational study that compared directly observed or verified interventions and practices (considered to be the gold standard) with coverage measured by two different data sources: women's report at exit survey after discharge and hospital routine register records (appendix p 5). The contexts and methods are detailed in a previously published protocol.[28] Five public hospitals providing comprehensive emergency obstetric and neonatal care and including the interventions of interest in the contexts of high mortality burden in sub-Saharan Africa and south Asia were identified in Bangladesh, Nepal, and Tanzania (appendix p 6). Inclusion and exclusion critieria are given in the panel.

Participants gave voluntary informed written consent before recruitment for observation and again for exit surveys. This study was granted ethics approval by institutional review boards in all three countries and by the London School of Hygiene & Tropical Medicine, London, UK (appendix p 7). All collaborating partners have signed data sharing and transfer agreements.

Procedures

In each hospital, consenting participants were recruited in three service delivery contexts to collect data on five selected interventions (panel). Trained researchers in each hospital obtained informed consent and collected participant data prospectively. Data collection varied by site from 7 months to 12 months to achieve the required precision-based sample size (appendix p 8). Clinical observers worked in shifts, covering 24 h each day. Separate groups of data collectors were responsible for observation and verification, exit survey, and routine register data extraction. We used an Android-tablet-based electronic data capture system that was custom built and designed to restrict access by data collector group but linked records at an individual level.[28,31] All data were stored locally on the tablet and synchronised to a country database server managed in Microsoft SQL. A centralised

See Online for appendix

***Panel*: Service delivery contexts and selected interventions**

Labour and delivery ward

(1) Uterotonics to prevent post-partum haemorrhage, (2) early initiation of breastfeeding, and (3) neonatal resuscitation by bag-mask ventilation. Eligible women and their newborn babies were observed while admitted on the labour ward. Exclusion criteria at admission no fetal heartbeat heard or imminent birth.

Kangaroo mother care ward or corner

(4) Skin-to-skin contact or kangaroo mother care position between mother and baby. All mother and baby pairs who were admitted to the kangaroo mother care ward or corner were eligible for observation and exit survey.

Newborn care ward or paediatrics ward

(5) Antibiotic treatment for neonatal infection. Infection diagnosis and name of injectable antibiotic treatment was verified from case notes because observation was not feasible.[30] Eligible neonates were those admitted with clinically defined infection (sepsis, pneumonia, or meningitis). Antibiotic treatment was not documented in routine registers in these hospitals.

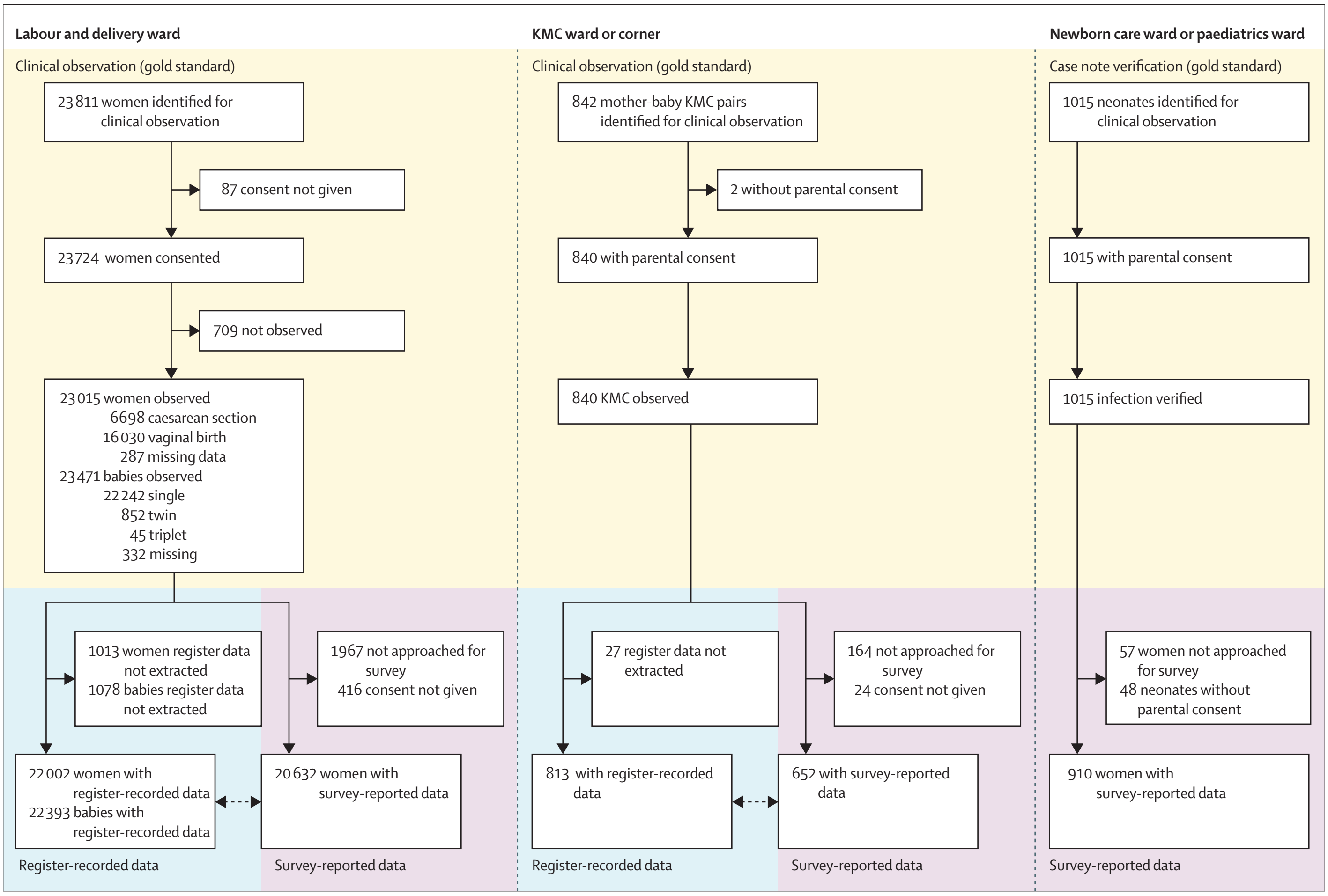


Figure 1: **Flow diagram for EN-BIRTH datasets**
Dashed line indicates methods were compared. KMC=kangaroo mother care.

web-based dashboard was developed to monitor the progress of data capture for selected interventions.

Observation data regarding the interventions and practices were time-stamped, captured in real time by touching a specific data element: once for "observed done" and twice for "observed not done" to override the default "not observed". Women's responses to close-ended questions in the exit surveys regarding the interventions and practices were captured in real time and recorded as "yes", "no", or "don't know, don't remember".

Health workers record each admitted individual woman and baby on one row in the routine registers and record data elements in columns that are either specific columns for that data element or non-specific columns (eg, other details). Research data collectors extracted intervention and practice data from these existing register records after hospital discharge, captured as "yes", "no", "not recorded", or "not readable". Registers were assessed according to two dimensions of data quality, completeness and external consistency with observation. If the register column was blank for an intervention, the column was extracted as "not recorded" in Tanzania and Nepal, but in Bangladesh as "no" (intervention not given) to align with register filling instructions. Hence calculating data element completeness in the Bangladesh registers was not possible. Bangladesh, Nepal, and Tanzania each had different routine register designs, all of which were paper based. The labour ward register design in Bangladesh changed during the study because of national standardisation. In the tables and figures we present data from the revised registers. Original and revised register design and data are shown in the appendix (pp 9–10). In Muhimbili, Tanzania, additional data elements were captured in a long-standing informal perinatal register.

Statistical analysis

Data were anonymised before pooling for analysis using Stata (version 14.2) To assess coverage for the selected indicators, we calculated observed, survey-reported, and register-recorded coverage (appendix p 11). We excluded participants with missing data from their relevant sample. For numerator validation, we used the

	Observed coverage	Survey-recorded coverage	"Don't know" responses	Sensitivity	Specificity	Percent agreement*
Labour and delivery ward† (n=23 471)						
Uterotonics	99·4% (98·7–99·8)	84·7% (79·1–89·5)	8·7% (4·5–14·1)	84·9% (79·6–89·6)	32·5% (21·2–44·6)	84·7% (79·4–89·4)
Early breastfeeding	10·9% (3·8–21·0)	53·2% (39·4–66·8)	0·6% (0·1–1·3)	76·9% (70·7–82·7)	50·0% (32·3–67·7)	53·8% (40·2–67·2)
Neonatal resuscitation (bag-mask ventilation)	4·4% (1·9–8·1)	0·8% (0·4–1·4)	5·9% (2·4–10·7)	9·3% (4·7–15·0)	99·5% (99·2–99·8)	96·0% (93·1–98·1)
Kangaroo mother care ward or corner (n=840)						
Kangaroo mother care	100% (99·9–100)	99·9% (98·3–100)	0·0% (0·0–0·1)	100% (99·8–100)	‡	100% (99·8–100)
Newborn care ward or paediatrics ward (n=1015)						
Antibiotic injection for neonatal infection§	96·7% (94·0–98·6)	74·7% (55·3–90·1)	16·9% (7·4–29·2)	75·9% (55·6–91·6)	‡	75·3% (56·4–90·2)
Named antibiotic for neonatal infection§	96·7% (94·0–98·6)	12·3% (3·5–25·1)	16·9% (7·4–29·2)	12·7% (3·7–25·6)	‡	16·1% (8·0–26·2)

Further individual-level validity statistics and site-specific results by mode of birth are given in the appendix (p 17). *(true positives + true negatives) / n. †Data are for all modes of birth. ‡Specificity not reported because all true negatives not captured. §Verified from case notes.

Table 1: **Individual-level validity testing for survey-recorded coverage versus observed coverage**

simplest denominator (total women observed, total births [livebirths plus stillbirths] observed, kangaroo mother care (KMC) mother–baby pairs observed, or newborn babies treated for infection). A comparison of denominators, including true denominators of clinical need where relevant, will be analysed subsequently. Often, population-based surveys (eg, DHS or MICS) measure coverage from "yes" responses, therefore "don't know" and "no" responses might both be used to suggest no coverage. For registers, monthly data aggregation typically involves counting column ticks, so that "not recorded" is treated as "intervention not given". In our analyses we present these typical scenarios and compare the effect on validity of excluding "don't know" and "not recorded" responses (appendix p 12).

We calculated the absolute differences between observed coverage and exit-survey-reported and register-recorded coverage. Cut-off ranges were adapted from data quality review methods (overestimate or underestimate by 0–5%, 6–10%, 11–15%, 16–20%, and >20%) and used to generate heatmaps.[32] To assess population-level validity for all indicators across both platforms (survey and registers with specific columns) we calculated validity ratios, similar to verification factors in data quality review methods.[32] Validity ratios can be applied when interventions and practices are intended for all women or newborn babies or a small target group. A ratio higher than 1·0 implies overestimation of survey-reported or register-recorded coverage compared with observed, and a ratio lower than 1·0 implies an underestimate. Ratio measure cut-offs used 0·05 increments, defining "excellent" as 0·95–1·05. Measures of individual-level validity were calculated as follows: when two-way table column totals were 10 or more we calculated sensitivity, and if relevant, specificity, negative predictive value, positive predictive value, area under the curve (AUC), and inflation factor. Otherwise we present percent agreement.[28,30] We excluded participants with missing data on a pairwise basis.

All calculations were first done separately for each participating hospital and exact 95% CIs were calculated using the binomial distribution. We then combined the hospital-specific results using a random effects meta-analysis approach (Stata metan command). We calculated I^2 and τ^2 to assess heterogeneity between hospitals. In addition to the protocol planned analyses, because of the increasing proportion of caesarean sections globally, and the many caesarean births in this study, we did analyses stratified by mode of birth (vaginal birth [normal vaginal or vacuum extraction and forceps combined] and caesarean section) to assess any effect on measurement.

STARD guidelines were followed throughout (appendix p 13).

To assess the reliability of our gold standard observation, we calculated Cohen's κ coefficient for the 5% of the sample observed by both supervisors and data collectors.[28] We included all indicators in our analyses, discussing κ scores below high or substantial cut offs, less than 0·71 for observation and less than 0·91 for data extraction, in study limitations. To assess any change in recording practices in routine registers due to study observer presence, we compared absolute differences between completeness of extracted study data with register data from 1 year pre-study collected retrospectively.[22] We also calculated κ coefficients for a 5% sample of double-extracted study register data.

EN-BIRTH is registered with Research Registry, 4833.

Role of the funding source
The funder of the study attended the study design workshop but had no role in data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study and had final responsibility for publication submission decision.

Results
Between July 3, 2017, and May 30, 2018, among 23 015 women on the labour and delivery ward, 23 471 births were observed, 20 632 (89·6%) women had an exit survey, and 22 002 (95·6%) had register data

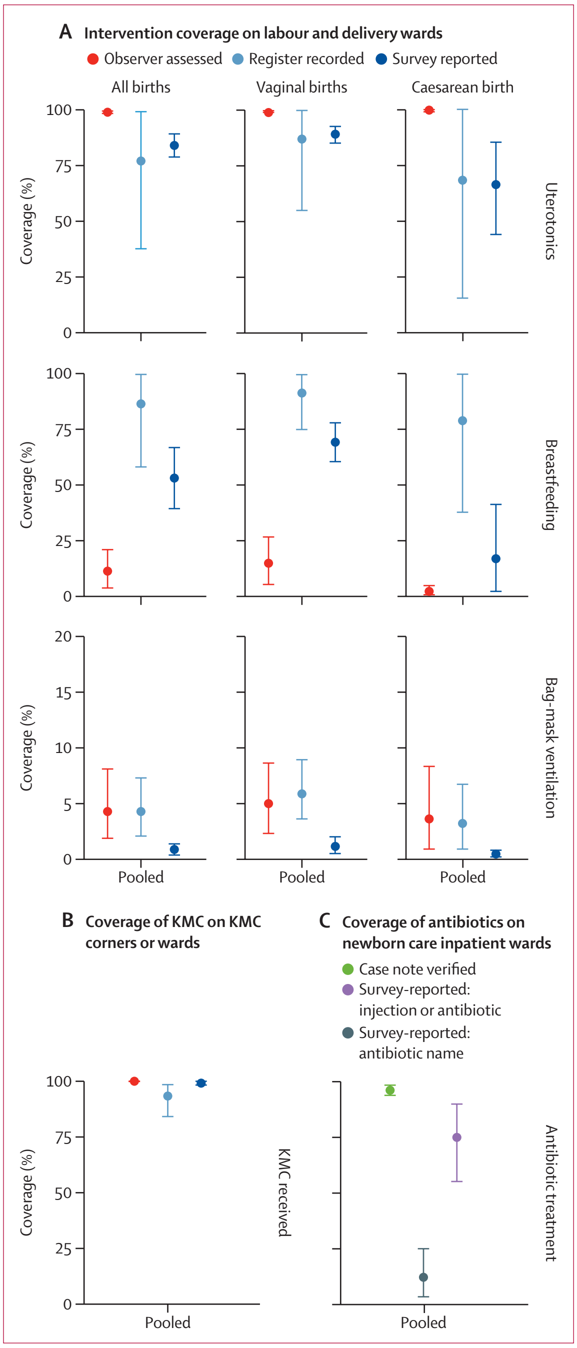


Figure 2: **Coverage for five selected indicators measured by observation, register, and survey, overall and by mode of birth**
Labour and delivery ward (n=23 015; A), KMC (n=840; B), verified neonatal infection (n=1015; C). Error bars show 95% CI. Pooled using random effects meta-analysis. Site-specific results by mode of birth are given in the appendix (pp 21–22). KMC=kangaroo mother care.

extracted (figure 1). Exit survey interviews were done at mean 1·4 days (SD 2·9) after birth (appendix p 8). Consent was not granted by 87 (0·4%) women for observation and 416 (1·8%) for exit survey. 1967 (8·5%) women left the hospital before the survey could be done.

Birth outcomes and background characteristics are shown in the appendix (pp 15–16). Women younger than 19 years comprised 4·6–17·2% of the sample and secondary education completion varied between sites (34·8–61·2%). Caesarean sections were done for 6698 participants, ranging from 7·0% in Temeke, Tanzania, to 72·8% in Azimpur, Bangladesh (appendix p 15–16). The proportion of in-facility stillbirths ranged from three per 1000 to 49 per 1000 total births and the proportion of newborn babies with low birthweight (<2500 g) ranged from 7·4% to 26·5% (appendix p 16).

Survey-reported coverage of uterotonics for prevention of post-partum haemorrhage was 84·7% (95% CI 79·1–89·5), compared with observed coverage of 99·4% (98·7–99·8; table 1). Exit survey heterogeneity was low (τ^2 0·027), with 8·7% (4·5–14·1) of responses given as "don't know" (appendix p 24). At the individual validity level, sensitivity of survey-reported coverage of uterotonics was 84·9% (79·6–89·6) and specificity 32·5% (21·2–44·6; table 1). Exclusion of "don't know" responses resulted in sensitivity increasing to 93·5% (91·1–95·5) and specificity decreasing to 17·3% (7·4–29·4; appendix p 24). Survey-reported coverage of uterotonics after caesarean birth was 66·3% (44·0–85·3) compared with observed coverage of 99·5% (98·9–99·9), and after vaginal birth was 89·3% (85·3–92·8) compared with observed coverage of 99·6% (99·1–99·9; figure 2, appendix p 21).

Register-recorded coverage of uterotonics for prevention of post-partum haemorrhage was 77·9% (95% CI 37·8–99·5) with high heterogeneity (τ^2 0·729) between hospitals, compared with observed coverage of 99·2% (98·6–99·7; table 2; appendix pp 17, 24). In the hospital in Pokhara, Nepal, the register had no column to capture uterotonics, all other hospitals used specific columns. Temeke, Tanzania underestimated coverage by 1·7% and Azimpur overestimated by 0·5% (appendix pp 17, 24, 29). However, different hospitals in these countries using the same registers underestimated coverage, by 78·2% in Kushtia and by 33·9% in Muhimbili. At the individual validity level, sensitivity was 78·0% (37·8–99·5) and specificity 22·8% (1·7–53·6). Exclusion of "not recorded" records resulted in sensitivity increasing to 86·1% (48·5–100·0) and specificity decreasing to 3·5% (0·0–17·2). Register-recorded coverage of uterotonics after caesarean birth was 68·5% (15·5–100·0) compared with observed coverage of 99·4% (98·7–99·9), and after vaginal birth was 86·6% (55·0–100·0) compared with observed coverage of 99·4% (98·7–99·9; figure 2; appendix p 24).

Validity ratios for coverage of uterotonics by exit survey were "very good" in Temeke, Tanzania, "good" in Kushtia, Bangladesh, and "moderate" in the remaining three sites. Register validity ratios for coverage of uterotonics were "excellent" in Azimpur and Temeke, but "poor" in Kushtia, and Muhimbili (figure 3).

Among newborn babies who were observed for more than 1 h after birth (n=6304), exit surveys substantially

	Observed coverage	Register-recorded coverage	Not recorded	Not readable	Sensitivity	Specificity	Percent agreement*
Labour and delivery ward (n=23 471)†							
Uterotonics	99·2% (98·6–99·7)	77·9% (37·8–99·5)	3·1% (0·0–19·1)	0·1% (0·0–0·1)	78·0% (37·8–99·5)	22·8% (1·7–53·6)	77·2% (37·7–99·3)
Early breastfeeding	12·5% (4·6–23·6)	85·9% (58·1–99·6)	7·6% (1·1–19·2)	0·1% (0·0–0·2)	97·6% (83·9–100)	6·4% (0·0–29·2)	24·6% (8·5–45·7)
Neonatal resuscitation (bag-mask ventilation)	5·1% (2·0–9·6)	4·3% (2·1–7·3)	65·4% (15·3–99·2)	0·3% (0·2–0·6)	23·6% (7·3–45·2)	96·8% (94·2–98·7)	93·2% (88·0–97·0)
Kangaroo mother care ward or corner (n=840)							
Kangaroo mother care	100% (99·9–100)	92·9% (84·2–98·5)	0·9% (0·2–2·0)	0·0% (0·0–0·3)	93·0% (84·5–98·5)	‡	93·0% (84·5–98·5)

Further individual-level validity statistics and site-specific results by mode of birth are given in the appendix (p 17). *(true positives + true negatives) / n. †Data are for all modes of birth. ‡Specificity not reported because all true negatives not captured. Antibiotic treatment for neonatal infections was not documented in routine registers in these hospitals.

Table 2: **Individual-level validity testing for register-recorded coverage versus observed coverage**

Ratio	Hospital ward	Selected indicator	Azimpur, Bangladesh Tertiary	Kushtia, Bangladesh District	Pokhara, Nepal Regional	Temeke, Tanzania Regional	Muhimbili, Tanzania National	All sites pooled (95% CI)
Survey to observed	A. Labour ward	1 Uterotonics to prevent post-partum haemorrhage	0·81	0·88	0·82	0·92	0·81	0·85 (0·80–0·91)
		2 Early initiation of breastfeeding	27·40	6·70	10·13	2·68	1·43	2·92 (1·58–5·38)
		3 Neonatal Resuscitation - Bag-mask ventilation	0·60	0·30	0·14	0·10	0·14	0·20 (0·12–0·34)
	B. KMC ward	4 KMC	0·89	1·00	1·00	1·00	1·00	1·00 (0·99–1·01)
	C. Neonatal ward	5.1 Neonatal infection antibiotics - injection	0·83	0·60	0·49	0·92	0·96	0·74 (0·55–1·00)
		5.2 Neonatal infection antibiotics - name	0·05	0·26	0·03	0·22	0·15	0·11 (0·05–0·25)
Register to observed	A. Labour ward	1 Uterotonics to prevent post-partum haemorrhage	1·01	0·22		0·98	0·66	0·61 (0·45–0·84)
		2 Early initiation of breastfeeding	50·40	9·84		3·67	2·30	4·29 (7·22–7·25)
		3 Neonatal resuscitation - Bag-mask ventilation	1·14	1·27		0·75	0·55	0·85 (0·59–1·23)
	B. KMC ward	4 KMC	1·00	0·98		0·85	0·85	0·92 (0·82–1·03)
	C. Neonatal ward	5 Neonatal infection antibiotics/injection						

<0·80 or >1·20 poor; 0·80 to 0·84 or 1·16 to 1·20 moderate; 0·85 to 0·89 or 1·11 to 1·15 good; 0·90 to 0·94 or 1·06 to 1·10 very good; 0·95 to 0·99 or 1·00 to 1·05 excellent; Data not captured in specific column in register

Figure 3: **Heatmap for selected indicator validity ratios**
Validity ratios study by sites and pooled (heatmap cut offs 5%, 10%, 15%, and 20%). Pooled using random effects meta-analysis. KMC=kangaroo mother care. Cut-off ranges adapted from WHO Data Quality Review, Module 2.[32]

overestimated the coverage of breastfeeding initiated within 1 hour after birth compared with observed coverage (table 1). Exit survey heterogeneity was low (τ^2 0·101), with 0·6% (95% CI 0·1–1·3%) of responses "don't know" (appendix pp 17, 31, 36). Survey-reported coverage of early breastfeeding after caesarean birth was 17·1% (2·3–41·3), compared with observed coverage of 2·4% (1·2–3·9) and 69·5% (60·5–77·9) after vaginal birth, compared with observed coverage of 14·4% (5·4–26·7; figure 2, appendix pp 31, 36).

Register-recorded coverage also substantially over-estimated early breastfeeding compared with observed coverage (table 2), with high heterogeneity between hospitals (τ^2 0·423; appendix pp 31, 36). Register-recorded coverage of early breastfeeding after caesarean birth was 78·3% (95% CI 37·8–99·7), compared with observed coverage of 2·2% (0·9–4·0) and after vaginal birth was 91·4% (74·9–99·5) compared to observed 17·3% (8·0–29·1; figure 2, appendix pp 31, 36). Validity ratios for both survey and registers were categorised as "poor" in all sites (figure 2).

Exit survey-reported coverage underestimated bag-mask ventilation compared with observed coverage, using a total birth denominator (table 1). Exit survey heterogeneity was low (τ^2 0·003) with 5·9% (95% CI 2·4–10·7) of responses "don't know" (appendix pp 18, 38, 43). Sensitivity was 9·3% (4·7–15·0) and specificity was 99·5% (99·2–99·8; table 1). Exclusion of "don't know" responses resulted in sensitivity increasing slightly to 12·5% (6·5–19·9) with no decrease in specificity (appendix pp 38, 43). Survey-reported coverage of bag-mask ventilation after caesarean birth was 0·4% (0·2–0·8), compared with observed coverage of 3·7% (0·9–8·3), and after vaginal birth was 1·1% (0·5–2·0), compared with observed coverage of 5·0% (2·3–8·6) (figure 2, appendix pp 38, 43).

Register-recorded coverage underestimated bag-mask ventilation, with low heterogeneity (τ^2 0·017), compared

A Labour and delivery ward		Azimpur, Bangladesh Tertiary	Kushtia, Bangladesh District	Pokhara, Nepal Regional	Temeke, Tanzania Regional	Muhimbili, Tanzania National
1. Uterotonics to prevent post-partum haemorrhage						
Register design: column allotted data element		Specific column	Specific column	No column	Specific 2 columns	Specific 2 columns
Completeness	Data element recorded in register	Not possible	Not possible		99·2%	68·6%
External Consistency	Indicator: observed coverage %	98·9%	99·8%		99·3%	98·4%
	Indicator: measured coverage, register recorded %	99·4%	21·6%		97·6%	64·5%
	Measurement gap: register recorded and observed	0·6% underestimate	78·2% underestimate		1·7% underestimate	34·0% underestimate
2. Early Initiation of breastfeeding (observed > 1 h)						
Register design: column allotted data element		specific column	specific column	No column	specific 2 columns	specific 2 columns
Completeness	Data element recorded in register	Not possible	Not possible		97·7%	76·6%
External Consistency	Indicator: observed coverage %	1·8%	9·8%		26·0%	19·1%
	Indicator: measured coverage, register recorded %	91·7%	96·8%		95·3%	43·8%
	Measurement gap: register recorded and observed	89·9% overestimate	87·0% overestimate		69·3% overestimate	24·7% overestimate
3. Neonatal resuscitation (bag-mask ventilation)						
Register design: column allotted data element		Specific column	Specific column	No column	Specific 2 columns	Specific 2 columns
Completeness	Data element recorded in register	Not possible	Not possible		91·1%	55·7%
External Consistency	Indicator: observed coverage %	0·8%	6·1%		7·1%	9·0%
	Indicator: measured coverage, register recorded %	0·9%	7·7%		5·3%	5·0%
	Measurement gap: register recorded and observed	0·1% overestimate	1·6% overestimate		1·8% underestimate	4·0% underestimate
B Kangaroo mother care ward or corner						
4. Kangaroo mother care						
Register design: column allotted data element		Specific column	Specific column	Non-specific	Specific 2 columns	Specific 2 columns
Completeness	Data element recorded in register	100·0%	98·5%	93·0%	99·1%	98·8%
External Consistency	Indicator: observed coverage %	100·0%	99·9%	99·9%	99·8%	99·5%
	Indicator: measured coverage, register recorded %	100·0%	97·8%	21·2%	84·8%	85·2%
	Measurement gap: register recorded and observed	0·0% accurate	2·1% overestimate	78·7% underestimate	15·0% underestimate	14·3% underestimate
C Newborn care ward or paediatrics ward						
5. Neonatal Infection Antibiotic Treatment						
Register design		No column	No column	No column	No column	No column

No column for data element | Non-specific column for data element | Specific column
>20% poor | 16–20% moderate | 11–15% good | 6–10% very Good | 0–5% excellent

Figure 4: **Routine register design for EN-BIRTH study sites and accuracy of data quality dimensions**
Cut-off ranges adapted from WHO Data Quality Review, Module 2.[32] Completeness calculations were denoted "not possible" for Bangladesh registers because the register was designed to be left blank if the intervention or practice was not done. An expanded version of this figure is shown in the appendix (p 9).

with observed coverage (table 2; appendix pp 18, 38). Register column design varied (figure 4). In Bangladesh, the column was ticked when bag-mask ventilation was done and was otherwise left blank, and coverage was slightly overestimated by 0·1–1·6% (figure 4). In Tanzania, a numerical code ("3" for bag-mask ventilation) or "no" was written in the column, and completeness was 55·7% for Muhimbili and 91·1% for Temeke (appendix pp 38, 43). The Pokhara register did not capture this data element. Sensitivity was 23·6% (95% CI 7·3–45·2) and specificity was 96·8% (94·2–98·7; table 2). Exclusion of "not recorded" records resulted in sensitivity increasing to 53·6% (28·1–78·1) and specificity decreasing to 77·7% (57·9–92·5; appendix pp 38, 43). Register-recorded coverage of bag-mask ventilation after caesarean birth was 3·2% (0·9–6·7), compared with observed coverage of 4·1% (0·6–10·4), and after vaginal birth register-recorded coverage of bag-mask ventilation was 5·9% (3·6–8·9), compared with observed coverage of 5·9% (3·3–9·2; figure 2, appendix pp 38, 43). Survey validity ratios were all categorised as "poor" (figure 3). For register validity ratios, Azimpur was "good" but all other sites were "poor" (figure 3).

In KMC wards or corners, 840 mother–baby pairs, including 91 babies who were born outside of the study hospital, were observed, with 652 exit surveys done and 813 routine register records extracted (figure 1). 815 (97%) of KMC babies had a birthweight of 2000 g or less (WHO recommendation or KMC). Background characteristics are shown in the appendix (p 45). Pre-discharge mortality was low (0·0–1·8%; appendix p 45).

Exit survey-reported coverage of KMC was 99·9% (95% CI 98·3–100), compared with observed coverage of 100% (99·9–100), and sensitivity was 100% (99·8–100), with zero "don't know" responses (table 1). KMC coverage was captured in specific national standardised KMC ward registers, except in Nepal, which used the general child register (figure 4). Register-recorded coverage from standardised specific KMC registers was 92·9% (84·2–98·5), with low heterogeneity (τ^2 0·065), compared with observed coverage of 100% (99·9–100) and sensitivity of 93·0% (84·2–98·5; table 2; appendix pp 19, 46, 48). Validity ratios for KMC by exit surveys were categorised as "excellent" for all sites, except Azimpur, which was "good", and for registers validity ratios were "excellent" in the Bangladesh sites and good in the Tanzania sites (figure 3).

In newborn care inpatient wards, among 1532 newborn babies, a diagnosis of clinically defined infection (sepsis, pneumonia, or meningitis) was documented in individual case notes for 1015 (66·6%) neonates and 910 exit surveys were done (figure 1). Background characteristics of these neonates are shown in the appendix (p 49). Coverage of antibiotic treatment verified by case notes (gold standard) was 96·7% (95% CI 94·0–98·6; table 1). Exit-survey reported coverage was 74·7% (55·3–90·1) when measured using a general question regarding injection or antibiotic use, with sensitivity of 75·9% (55·6–91·6) and high heterogeneity (τ^2 0·204; appendix pp 19, 50, 52). The proportion of "don't know" responses was high (table 1). When adding the question regarding the name of the antibiotic, survey-reported coverage dropped to 12·3% (3·5–25·1; table 1; appendix pp 20, 50, 52). Validity ratios for the general question (injection or antibiotic) ranged from excellent in Muhimbili to poor in Pokhara; for the more specific question (antibiotic name), ratios were poor in all sites (figure 3).

Analysis of inter-rater reliability for gold standard data showed high or substantial κ scores for most data elements but moderate scores for observed uterotonic coverage in Temeke, and early breastfeeding in Nepal and Temeke (appendix pp 53–54). Lower κ scores were found for both KMC and verification of antibiotics in Pokhara (appendix p 51). Inter-rater reliability for routine register data was lower than the high or substantial cut offs for all labour ward indicators in Kushtia, Temeke, and Muhimbili; and for KMC in Pokhara (appendix pp 55–56). Register completeness comparison, before and during the study, revealed decreases of more than 5% for bag-mask ventilation coverage in both Tanzanian sites and for uterotonic coverage in Muhimbili. Breastfeeding completeness increased in Muhimbili from 0 to 99·4% (appendix p 54).

Discussion

We examined validity of coverage indicators for selected maternal and newborn care indicators in two data systems: exit surveys and routine registers. Surveys are highly standardised in question design and interview technique. Registers are variable in design and filling techniques. We found much higher heterogeneity for register-recorded coverage compared with exit-survey reported coverage. Even with the same register design, accuracy varied between hospitals, with good validity for the highest performing sites. We stratified data by mode of birth and found that caesarean birth affected measurement in surveys and registers. With rising caesarean rates, especially in LMICs,[33] this finding needs further consideration. Register data that are aggregated for use are typically located in the labour and delivery ward, but with high caesarean section rates, specific registers in operating theatres might be necessary. For survey reports after caesarean birth, low accuracy might relate to not seeing an intervention happening (eg, whether or not the baby had bag-mask ventilation) but might also be a reflection on gaps in respectful care; women have a right to communication and informed consent regarding care for themselves or their newborn babies.[34,35]

Survey report had low accuracy for indicators of clinical intervention coverage led by health workers around the time of birth, notably neonatal care, and to a lesser extent for uterotonics. This study is the first to test the validity of survey reporting of indicators of care for small and ill newborn babies. We found that survey reporting of bag-mask ventilation coverage had low accuracy and reporting of neonatal infection treatment with antibiotics had low sensitivity among the target group. By contrast, KMC, which is led by the woman had high sensitivity in exit surveys, with potential for further testing, including report from women did not practice KMC. Use in population-based surveys would require sufficient sample size for the target group of small babies requiring KMC. For neonatal infection, even interviewing only women whose babies had been admitted, we found high proportions of "don't know" responses and underestimates of observed coverage. A high proportion of "don't know" responses suggests that the survey question is a poor way of measuring intervention coverage. We reported "don't know" as "no" in line with the practice of the DHS and

MICS for yes or no questions, but we note that sensitivity increased when "don't know" responses were excluded for uterotonic and bag-mask ventilation coverage. Our exit survey had a recall period of a few days, but most household surveys cover the previous 2–5 years; other studies have found recall decay even by 1 year.[36] Hence, these results are likely to be worse in a routine household survey.

Early breastfeeding rates were observed to be very low across study hospitals, particularly after caesarean. The early initiation of breastfeeding indicator is already measured in the DHS and MICS household surveys, and national and global tracking rely on these data.[37,38] Our results indicate that both women's report and routine register data substantially overestimate coverage. This finding is in agreement with previous survey validation studies.[12–15] One register validation study for early breastfeeding in a composite indicator of essential newborn care also showed overestimation.[24] We postulate three explanations for these overestimates: first, this finding might be due to the timing component (ie, breastfeeding early but not within 1 h) being misreported by the woman or health worker. Second, successful initiation of breastfeeding is a process that involves the baby being put to the breast, then attachment, and then sucking. Putting the baby to the breast is one important step and is the focus of the survey-reported question, but might not have been considered as initiation by the observer. The observers were trained to click the stamp on the tablet at the point of initiation and recording is likely to vary given the challenges of observing this complex and dynamic process. Third, the overestimate might be due to social desirability bias among women to over-report in surveys or professional desirability pressure on the health worker to over-record in registers. More work is needed to improve measurement of this crucial indicator, including exploring whether changing the timing component could increase accuracy in surveys and registers. In Muhimbili, before and during the study, register data completeness for breastfeeding increased from 0 to 99%. This finding was probably due to the data being extracted from the informal perinatal register before the study, rather than the formal labour ward register, and highlights how complex documentation systems affect measurement.

Registers have unrealised potential as a useful data source, shown by the high accuracy and sensitivity for indicators in some EN-BIRTH study hospitals. Although hospital registers can only capture a limited number of data elements, we found that register designs in Tanzania and Bangladesh already have the relevant numerator or count data for selected indicators of maternal and newborn health coverage.[22] Many data are already being collected by frontline health workers. Register data were highly complete, and although data collectors rarely indicated data were not readable, we found low inter-rater κ results across register recorded data. Extracting data for aggregation is a crucial step for data flowing to higher levels in the health system, and more research is needed to inform data extraction quality. The accuracy of register-recorded coverage varied between hospitals even with identical register design, reflecting variation in implementation and data culture.[39] In both Bangladesh hospitals, register-recorded coverage increased in the revised registers when specific columns for data elements replaced absent or non-specific columns in original registers; however, change in accuracy varied by indicator. Sensitivity increased (and specificity decreased) when "not recorded" data were excluded from analyses (eg, uterotonics and bag-mask ventilation). Further work on register design for high quality monthly data aggregation, as well as feedback after use, is needed.[40] Qualitative research might help to understand the differences in these five hospitals by exploring barriers and enablers to routine documentation, specifically for register design, register filling, and register use.[28]

EN-BIRTH study strengths include multi-country sites, rigorous observational design, and large sample size (about 10-times higher than that in previous validation studies[12–15,24]) targeting previously missed clinical care groups, especially small and ill newborn babies—a priority in universal health coverage measurement.[16] Errors were minimised in observation data by the tablet-based app design, which was custom-built and user friendly, with major effort invested in being able to navigate between the recording of simultaneous events for the woman and baby with minimal delay.[31] The tablet app also importantly captured time-stamped data for time-sensitive interventions, around the time of birth when complications might mean that women can die within hours and babies within minutes. Detailed analyses, including quality of care with timing data, are published separately.[41–45] Dual observation by supervisors showed high or substantial agreement for most data elements, with breastfeeding coverage scoring the least well across all sites, probably due to challenges of capturing the process. Information bias during data collection was reduced by using different data collector groups for observation, exit surveys, and register data extraction. The effect on register recording completeness from the presence of researchers was assessed by comparison with register data extraction before the study.[22] As per protocol, we did not base our assessment of validity using AUC cut offs because our data were all binary (yes or no) for the coverage estimates; thus, the AUC is simply the average of the sensitivity and specificity.[28] I^2 estimates the percentage of variation that is attributable to study heterogeneity. For an intervention like uterotonics that was almost universally applied, there is little total variation so that even small differences between sites result in a large I^2. We therefore chose to place more emphasis in τ^2, which provides an estimate of the magnitude of the between-site variation. A small value of τ^2 indicates little absolute variation between sites even when I^2 might appear large.

However, our study also had limitations. Despite large samples, we did not achieve ten or more column counts in the two-way tables for all indicators, either because interventions had very high coverage (eg, correctly provided for all women and babies) or very low prevalence (eg, small clinical target group). This affected our ability to report on individual-level validity metrics. By contrast with other measurement validation studies, EN-BIRTH chose to use vigorous subset double observation inter-observer κ calculations to assess possible between-site variation in validation results. In the protocol, we pre-defined ranges for what would be considered high or substantial agreement. κ scores were lower than expected for some indicators in some sites. We postulate that this finding might be a real reflection of inter-observer variation, and we suspect this to be the case, for example, regarding breastfeeding in Temeke, with the lowest percent agreement and a low κ. We note the dependence of κ on prevalence with paradoxically low κ scores, due to the imbalance in marginal totals or with perfect symmetry in the imbalance, as our results showed.[46] The tablet app was not available for pre-study extraction of register data, and the different data collection methods might account for some of the differences in completeness before and during the study.

The five hospitals were high-volume public comprehensive emergency obstetric and neonatal care hospitals in cities and results might not be generalisable to lower-level or rural facilities with lower volumes. The study sample might have been healthier than is typical in such hospitals because recruitment excluded antepartum intrauterine fetal death and women who were too ill to consent after admission. This observational study was not designed to capture true denominators in terms of the need of intervention required for coverage indicators—eg, for accurate diagnosis of neonatal sepsis in terms of microbiological culture or molecular diagnosis. Although blood culture is still considered the definitive diagnostic method, even with excellent laboratory capacity, only about a third of neonatal sepsis cases have a positive culture result. The focus of this study was validity of routine health system data, especially of the clinical diagnosis, and was not addressing the need for better laboratory systems, which is also crucial.[45,47]

Time pressure on health workers in these busy hospitals might have affected their ability to deliver and communicate care to women, affecting exit survey reports and register documentation.[48] We consider this real-world time pressure to be present in many such contexts. Women's reports were collected at hospital exit, close in time to the event and without the typical 2–5 year gap in household surveys. If women cannot report accurately at exit from the facility, they are unlikely to report more accurately later in a household survey, so exit survey findings are relevant when considering adding questions to household surveys.[49] The EN-SMILING study is following up with the cohort, with the potential to repeat interviews in 5 years to investigate the change in women's report.[28]

Policymakers and programme managers require information to inform investments and programmatic course correction. Surveys are important in most LMICs for population-based outcome data and contact coverage of care, but given the high rate of "don't know" responses and low accuracy for the reporting of clinical interventions, adding these indicators to surveys is not justifiable. Because approximately 80% of births worldwide are now in facilities, standardising register design and linked information systems have the potential to sustainably improve data quality for care at birth. Routine register data can be accurate and health workers' time investment would not be wasted if these data were better used. Well-designed, standardised registers are important, and could reduce the burden on health workers of duplicative, or non-valid, data collection.[22] If interventions and practices are defined with a timing element (eg, early initiation of breastfeeding), this inclusion needs consideration in register design. The timing component of the uterotonics coverage indicator is not yet clearly defined, which will limit comparability if routinely measured. EN-BIRTH tested validity of measurement in paper-based registers at the interface with women and their babies. A further phase will explore the feasibility of these indicators flowing up through the national routine data systems, many of which are being rapidly digitised.[28] Another important research gap is how to best measure experience of care, including respectful care in all settings, and the provision and experience of care for women and newborn babies in fragile and humanitarian settings.[35,50,51]

Valid routine data alone will not save lives. Data need to be used by health-care professionals caring for women and their babies, and by policy makers and governments to invest and transform care, enabling universal health coverage as a reality that can be measured and improved.

Contributors

JEL conceived the EN-BIRTH study, acquired the funding and led the overall design. QS-uR was the main lead for analyses, working closely with LTD and the EN-BIRTH team. LTD and JEL drafted the manuscript with the principal investigators: SEA, AKC, HM, AER, NS, and country teams. All authors made substantial contributions to conception, design, data collection, analysis, or interpretation of data. GRG-L did the analysis for register recording practice comparison before and during the study. All authors revised the manuscript and gave final approval for publication and agree to be accountable for the work. Collaborative authors including the Expert Advisory Group made contributions to conception, design, data collection, analysis, or interpretation of data. The authors' views are their own, and not necessarily from any of the institutions they represent.

EN-BIRTH validation collaborative group

Bangladesh—Ayub Ali, Bilkish Biswas, Rajib Haider, Abu Hasanuzzaman, Amir Hossain, Ishrat Jahan, Rowshan Hosne Jahan, Jasmin Khan, M A Mannan, Tapas Mazumder, Hafizur Rahman, Ziaul Haque Shaikh,

Aysha Siddika, Taslima Akter Sumi, and Taqbir Us Samad Talha. *Tanzania*—Evelyne Assenga, Claudia Hanson, Edward Kija, Rodrick Kisenge, Karim Manji, Fatuma Manzi, Namala Mkopi, Mwifadhi Mrisho, and Andrea Pembe. *Nepal*—Jagat Jeevan Ghimire, Rejina Gurung, Elisha Joshi, Avinash K Sunny, Naresh P KC, Nisha Rana, Shree Krishna Shrestha, Dela Singh, Parashu Ram Shrestha, and Nishant Thakur. *London School of Hygiene & Tropical Medicine*—Hannah Blencowe and Sarah G Moxon.

EN-BIRTH Expert Advisory Group

Agbessi Amouzou, Tariq Azim, Debra Jackson, Theopista John Kabuteni, Matthews Mathai, Jean-Pierre Monet, Allisyn Moran, Pavani Ram, Barbara Rawlins, Jennifer Requejo, Johan Ivar Sæbø, Florina Serbanescu, and Lara Vaz.

Declaration of interests

We declare no competing interests.

Data sharing

For the **study data** see https://datacompass.lshtm.ac.uk/955/

The datasets generated during and analysed during the study are available on the London School of Hygiene & Tropical Medicine Data Compass repository.

Acknowledgments

This study was funded by The Children's Investment Fund Foundation through a grant awarded to JEL at the London School of Hygiene & Tropical Medicine. The Swedish Research Council funded the Nepal site through Lifeline Nepal and Golden Community. We acknowledge the core funders for all the partner institutions. We credit the inspiration of the late Godfrey Mbaruku. We thank Claudia DaSilva, Veronica Ulaya, Mohammad Raisul Islam, Sudip Karki, and Rabina Karki for their administrative support and Sabrina Jabeen, Goutom Banik, Shahidul Alam, Tamatun Islam Tanha, and Moshiur Rahman for support during data collector training. We also thank Ann Blanc, Emily Carter, and Melinda Munos for sharing relevant technical inputs and expertise. We acknowledge the National Advisory Groups: Mohammod Shahidullah, Muhiuddin Osmani, Sabina Ashrafee Lipi, and Jahurul Islam in Bangladesh; Naresh P KC in Nepal; and Muhammad Bakari Kambi, Georgina Msemo, Talhiya Yahya, Claud Kumalija, Eliudi Eliakimu, and Honest Kimaro in Tanzania. This paper is published with permission from the Directors of Ifakara Health Institute, Muhimbili University of Health and Allied Sciences, International Centre for Diarrhoeal Disease Research, Bangladesh, and Golden Community. We are also very grateful to fellow researchers who peer-reviewed this paper. Finally, and most importantly, we also thank the women, their families, the health workers and data collectors (appendix p 57).

References

1 WHO. Trends in maternal mortality 2000 to 2017 estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division 2019. Geneva: World Health Organization, 2019. https://www.who.int/reproductivehealth/publications/maternal-mortality-2000-2017/en/ (accessed May 20, 2020).

2 UNICEF. Levels and trends in child mortality report 2019. https://www.unicef.org/reports/levels-and-trends-child-mortality-report-2019 (accessed May 18, 2020).

3 Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *Lancet* 2014; **384:** 189–205.

4 WHO. Global strategy for women's, children's and adolescents health, 2016–2030. Geneva, World Health Organization, 2015. http://www.who.int/life-course/publications/global-strategy-2016-2030/en/ (accessed May 18, 2020).

5 UNICEF. The State of the World's Children 2019: statistical tables, https://www.unicef.org/media/63016/file/SOWC-2019.pdf (accessed June 11, 2020).

6 Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. *Lancet Glob Health* 2018; **6:** e1196–252.

7 Bhutta ZA, Das JK, Bahl R, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? *Lancet* 2014; **384:** 347–70.

8 Miller S, Abalos E, Chamillard M, et al. Beyond too little, too late and too much, too soon: a pathway towards evidence-based, respectful maternity care worldwide. *Lancet* 2016; **388:** 2176–92.

9 Marsh AD, Muzigaba M, Diaz T, et al. Effective coverage measurement in maternal, newborn, child, and adolescent health and nutrition: progress, future prospects, and implications for quality health systems. *Lancet Glob Health* 2020; **8:** e730–36.

10 Carvajal-Aguirre L, Vaz LM, Singh K, et al. Measuring coverage of essential maternal and newborn care interventions: an unfinished agenda. *J Glob Health* 2017; **7:** 020101.

11 Marchant T, Bryce J, Victora C, et al. Improved measurement for mothers, newborns and children in the era of the Sustainable Development Goals. *J Glob Health* 2016; **6:** 010506.

12 Stanton CK, Rawlins B, Drake M, et al. Measuring coverage in MNCH: testing the validity of women's self-report of key maternal and newborn health interventions during the peripartum period in Mozambique. *PLoS One* 2013; **8:** e60694.

13 Blanc AK, Warren C, McCarthy KJ, Kimani J, Ndwiga C, RamaRao S. Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya. *J Glob Health* 2016; **6:** 010405.

14 Blanc AK, Diaz C, McCarthy KJ, Berdichevsky K. Measuring progress in maternal and newborn health care in Mexico: validating indicators of health system contact and quality of care. *BMC Pregnancy Childbirth* 2016; **16:** 255.

15 McCarthy KJ, Blanc AK, Warren CE, Kimani J, Mdawida B, Ndwidga C. Can surveys of women accurately track indicators of maternal and newborn care? A validity and reliability study in Kenya. *J Glob Health* 2016; **6:** 020502.

16 Munos MK, Stanton CK, Bryce J. Improving coverage measurement for reproductive, maternal, neonatal and child health: gaps and opportunities. *J Glob Health* 2017; **7:** 010801.

17 Moxon SG, Ruysen H, Kerber KJ, et al. Count every newborn; a measurement improvement roadmap for coverage data. *BMC Pregnancy Childbirth* 2015; **15** (suppl 2): S8.

18 Mate KS, Bennett B, Mphatswe W, Barker P, Rollins N. Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa. *PLoS One* 2009; **4:** e5483.

19 Chiba Y, Oguttu MA, Nakayama T. Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. *Midwifery* 2012; **28:** 329–39.

20 Duffy S, Crangle M. Delivery room logbook - fact or fiction? *Trop Doct* 2009; **39:** 145–49.

21 Maternal and Child Survival Program. What data on maternal and newborn health do national health management information systems include? A review of data elements for 24 low- and lower middle income countries. 2018. https://www.mcsprogram.org/resource/what-data-on-maternal-and-newborn-health-do-national-health-management-information-systems-include (accessed May 18, 2020).

22 Day LT, Gore-Langton GR, Rahman AE, et al. Labour and delivery ward register data availability, quality, and utility - Every Newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries. *BMC Health Serv Res* 2020; **20:** 737.

23 Broughton EI, Ikram AN, Sahak I. How accurate are medical record data in Afghanistan's maternal health facilities? An observational validity study. *BMJ Open* 2013; **3:** e002554.

24 Bhattacharya AA, Allen E, Umar N, et al. Monitoring childbirth care in primary health facilities: a validity study in Gombe State, northeastern Nigeria. *J Glob Health* 2019; **9:** 020411.

25 WHO, UNICEF. Every Newborn: an action plan to end preventable deaths. Geneva: World Health Organization, 2014. https://apps.who.int/iris/handle/10665/127938 (accessed May 18, 2020).

26 WHO. WHO technical consultation on newborn health indicators: Every Newborn Action Plan metrics, Ferney Voltaire, France, 3–5 December 2014. Geneva: World Health Organization, 2015. https://apps.who.int/iris/handle/10665/184225 (accessed May 18, 2020).

27 Moran AC, Jolivet RR, Chou D, et al. A common monitoring framework for ending preventable maternal mortality, 2015–2030: phase I of a multi-step process. *BMC Pregnancy Childbirth* 2016; **16:** 250.

28 Shamba D, Day LT, Zaman SB, et al. Barriers and enablers to labour ward register data collection and use: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* 2020 (in press).

29 Benova L, Moller A-B, Moran AC. "What gets measured better gets done better": the landscape of validation of global maternal and newborn health indicators through key informant interviews. *PLoS One* 2019; **14:** e0224746.

30 Munos MK, Blanc AK, Carter ED, et al. Validation studies for population-based intervention coverage indicators: design, analysis, and interpretation. *J Glob Health* 2018; **8:** 020804.

31 Ruysen H, Rahman AE, Gordeev VS, et al. Electronic data collection for multi-country, hospital-based, clinical observation of maternal and newborn care: EN-BIRTH study experiences. *BMC Pregnancy Childbirth* 2020 (in press).

32 WHO. Data quality review: a toolkit for facility data quality assessment. Module 2: Desk review of data quality. Geneva: World Health Organization, 2017. https://apps.who.int/iris/handle/10665/259225 (accessed May 18, 2020).

33 Boerma T, Ronsmans C, Melesse DY, et al. Global epidemiology of use of and disparities in caesarean sections. *Lancet* 2018; **392:** 1341–48.

34 WHO. Standards for improving quality of maternal and newborn care in health facilities. Geneva: World Health Organization, 2016.

35 Bohren MA, Mehrtash H, Fawole B, et al. How women are treated during facility-based childbirth in four countries: a cross-sectional study with labour observations and community-based surveys. *Lancet* 2019; **394:** 1750–63.

36 McCarthy KJ, Blanc AK, Warren CE, Kimani J, Mdawida B, Ndwidga C. Can surveys of women accurately track indicators of maternal and newborn care? A validity and reliability study in Kenya. *J Glob Health* 2016; **6:** 020502.

37 Victora CG, Bahl R, Barros AJD, et al. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet* 2016; **387:** 475–90.

38 Fanzo J, Hawkes C, Udomkesmalee E, et al. 2018 global nutrition report: shining a light to spur action on nutrition. 2018. https://globalnutritionreport.org/reports/ (accessed May 18, 2020).

39 Lippeveld T, Sauerborn R, Bodart C. Design and implementation of health information systems. 2000. https://apps.who.int/iris/handle/10665/42289 (accessed May 18, 2020).

40 Braa J, Heywood A, Sahay S. Improving quality and use of data through data-use workshops: Zanzibar, United Republic of Tanzania. *Bull World Health Organ* 2012; **90:** 379–84.

41 Ruysen H, Shabani J, Hanson C, et al. Uterotonics for prevention of postpartum hemorrhage: EN-BIRTH multi-country study. *BMC Pregnancy Childbirth* (in press).

42 Tahsina T, Hossain AT, Ruysen H, et al. Immediate newborn care and breastfeeding: EN-BIRTH multi-country study. *BMC Pregnancy Childbirth* (in press).

43 Ashish KC, Peven K, Ameen S, et al. Neonatal resuscitation: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* (in press).

44 Salim N, Shabani J, Peven K, et al. Kangaroo mother care: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* 2020 (in press).

45 Rahman AE, Hossain AT, Ameen S, et al. Antibiotic use for inpatient newborn care with suspected infection: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* 2020 (in press).

46 Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; **43:** 543–49.

47 Fitchett EJA, Seale AC, Vergnano S, et al. Strengthening the Reporting of Observational Studies in Epidemiology for Newborn Infection (STROBE-NI): an extension of the STROBE statement for neonatal infection research. *Lancet Infect Dis* 2016; **16:** e202–13.

48 MEASURE Evaluation. Easing the Data collection burden on healthcare providers. 2017. https://www.measureevaluation.org/resources/publications/tr-17-211 (accessed May 30, 2020).

49 Ameen S, Siddique AB, Peven K, et al. Survey of women's report for 33 maternal and newborn indicators: EN-BIRTH multi-country validation study. *BMC Pregnancy and Childbirth* 2020 (in press).

50 Tunçalp Ö, Were WM, MacLennan C, et al. Quality of care for pregnant women and newborns—the WHO vision. *BJOG* 2015; **122:** 1045–49.

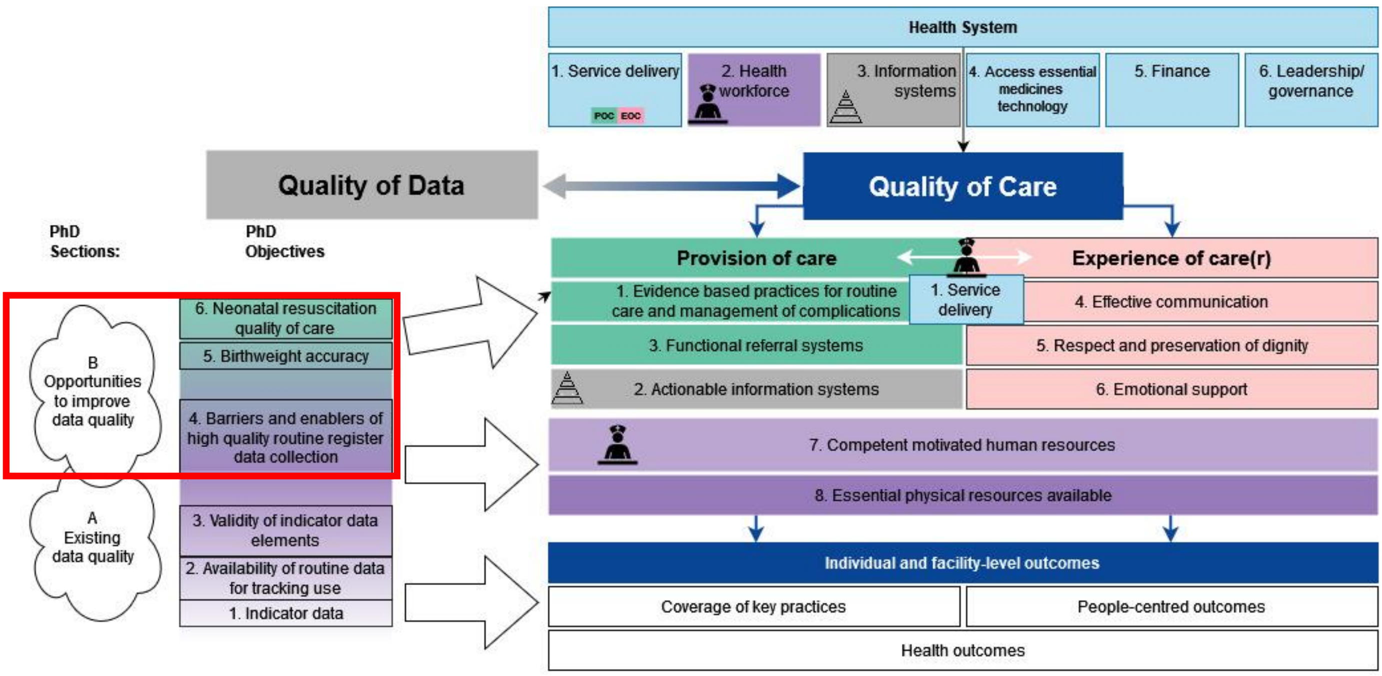
51 Afulani PA, Phillips B, Aborigo RA, Moyer CA. Person-centred maternity care in low-income and middle-income countries: analysis of data from Kenya, Ghana, and India. *Lancet Glob Health* 2019; **7:** e96–109.

SECTION B: Identifying opportunities to improve labour and delivery routine register data quality for hospital births

Section B – explores objectives 4, 5, and 6 as highlighted in the red box in Figure 6**:**

- Objective 4 – To explore barriers and enablers for health professionals to record high quality data for newborn and maternal health indicator measurement from labour ward routine registers in five EN-BIRTH study hospitals.
- Objective 5 – To assess routine birthweight in EN-BIRTH study hospitals: accuracy, gaps and opportunities to measure coverage and quality of care.
- Objective 6 – To assess measurement opportunities for neonatal resuscitation: indicator definitions and quality of care.

Figure 10: 'Quality of Care and Quality of Data conceptual framework' – highlighting PhD Thesis Section B - Identifying opportunities to improve labour and delivery routine register data quality for hospital births



Chapter 6 – Objective 4: Barriers and enablers to routine data

This chapter investigates the barriers and enablers for health professionals to hospital routine register documentation of coverage of care indicators for women and newborns in five EN-BIRTH study hospitals.

The chapter was published in March 2021 in BMC Pregnancy and Childbirth. The manuscript was published under a creative commons license (Creative Commons Attribution 4.0 International License) and no further permissions are needed.

The published manuscript is included in full below and supplementary material referenced in the paper is available at https://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/s12884-020-03517-3#Sec22

6.1 List of Figures

6.2 List of Tables

6.3 Citation

Shamba D, Day LT, Zaman SB, et al.
Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study.
BMC Pregnancy Childbirth, 2021; **21**(Suppl 1): 233. https://doi.org/10.1186/s12884-020-03517-3.

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	034282	**Title**	Dr
First Name(s)	Louise Tina		
Surname/Family Name	Day		
Thesis Title	Quality of care and quality of data for hospital births – tension or traction?		
Primary Supervisor	Associate Professor Cally Tann		

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?	Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth, 2021; 21(Suppl 1): 233. https://doi.org/10.1186/s12884-020-03517-3. Shamba, D., Day, L. T., Zaman, S. B., Sunny, A. K., Tarimo, M. N., Peven, K., Khan, J., Thakur, N., Talha, Mtus, K, C. A., Haider, R., Ruysen, H., Mazumder, T., Rahman, M. H., Shaikh, M. Z. H., Saebo, J. I., Hanson, C., Singh, N. S., Schellenberg, J., Vaz, L. M. E., Requejo, J., Lawn, J. E. and EN-BIRTH Study Group		
When was the work published?	March 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 jointly co-ordinated multi-country technical working group regular multi-site meetings. I jointly led collaborative design of data collection tools, thematic analyses and synthesis. I led the routine register design synthesis and the main results figure 1. I jointly drafted the manuscript with my co-first author. I am the corresponding author, led on the manuscript revision in response to peer-review, including co-ordinating collaborative inputs from co-authors. I managed proofs and co-presented results at dissemination activities.

SECTION E

Student Signature	[redacted]
Date	[redacted]

Supervisor Signature	[redacted]
Date	[redacted]

Shamba et al. BMC Pregnancy and Childbirth 2021, 21(Suppl 1):233
https://doi.org/10.1186/s12884-020-03517-3

BMC Pregnancy and Childbirth

From *Every Newborn* **BIRTH multi-country validation study: informing measurement of coverage and quality of maternal and newborn care**

RESEARCH **Open Access**

Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study

Donat Shamba[1†], Louise T. Day[2*†], Sojib Bin Zaman[3], Avinash K. Sunny[4], Menna Narcis Tarimo[1], Kimberly Peven[2,5], Jasmin Khan[3], Nishant Thakur[4], Md. Taqbir Us Samad Talha[3], Ashish K.C.[6], Rajib Haider[3], Harriet Ruysen[2], Tapas Mazumder[3], Md. Hafizur Rahman[3], Md. Ziaul Haque Shaikh[3], Johan Ivar Sæbø[7], Claudia Hanson[2,8], Neha S. Singh[2], Joanna Schellenberg[2], Lara M. E. Vaz[9], Jennifer Requejo[10], Joy E. Lawn[2†] and EN-BIRTH Study Group

Abstract

Background: Policymakers need regular high-quality coverage data on care around the time of birth to accelerate progress for ending preventable maternal and newborn deaths and stillbirths. With increasing facility births, routine Health Management Information System (HMIS) data have potential to track coverage. Identifying barriers and enablers faced by frontline health workers recording HMIS source data in registers is important to improve data for use.

Methods: The EN-BIRTH study was a mixed-methods observational study in five hospitals in Bangladesh, Nepal and Tanzania to assess measurement validity for selected *Every Newborn* coverage indicators. We described data elements required in labour ward registers to track these indicators. To evaluate barriers and enablers for correct recording of data in registers, we designed three interview tools: a) semi-structured in-depth interview (IDI) guide b) semi-structured focus group discussion (FGD) guide, and c) checklist assessing care-to-documentation. We interviewed two groups of respondents (January 2018–March 2019): hospital nurse-midwives and doctors who fill ward registers after birth ($n = 40$ IDI and $n = 5$ FGD); and data collectors ($n = 65$). Qualitative data were analysed thematically by categorising pre-identified codes. Common emerging themes of barriers or enablers across all five hospitals were identified relating to three conceptual framework categories.

Results: Similar themes emerged as both barriers and enablers. First, register design was recognised as crucial, yet perceived as complex, and not always standardised for necessary data elements. Second, register filling was performed by over-stretched nurse-midwives with variable training, limited supervision, and availability of logistical resources. Documentation complexity across parallel documents was time-consuming and delayed because of low staff numbers. Complete data were valued more than correct data. Third, use of register data included clinical handover and monthly reporting, but little feedback was given from data users.

(Continued on next page)

* Correspondence: Louise-Tina.Day@lshtm.ac.uk
†Donat Shamba and Louise T Day are joint first authors.
†Joy E. Lawn is the senior author.
[2]Centre for Maternal, Adolescent, Reproductive & Child Health (MARCH), London School of Hygiene & Tropical Medicine, Keppel St, London, UK
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusion: Health workers invest major time recording register data for maternal and newborn core health indicators. Improving data quality requires standardised register designs streamlined to capture only necessary data elements. Consistent implementation processes are also needed. Two-way feedback between HMIS levels is critical to improve performance and accurately track progress towards agreed health goals.

Keywords: Birth, Maternal, Newborn, Coverage, Facility registers, Indicators, Data quality, Health management information systems

Key findings

What is known and what is new about this study?

- Routine facility register data recorded in Health Management Information Systems (HMIS) in low- and middle-income countries (LMICs) provide an opportunity to close data gaps for tracking coverage of care at birth. Although around four-fifths of the world's births are now in facilities, labour ward register data are currently under-used and under-studied. Specifically, few studies have examined barriers and enablers for recording high quality routine maternal and newborn data, or on the use of labour and delivery ward registers.
- EN-BIRTH was the first multi-country, mixed-methods study to assess validity of register-recorded maternal and newborn coverage indicators. In the three study countries, we found register coverage measurement accuracy varied, even between hospitals in the same country using the same registers.
- Hence to assess barriers and enablers for health workers to record data in labour ward registers, we interviewed health workers ($n = 72$) and EN-BIRTH research data collectors ($n = 65$) across the five hospitals.

What did we find and what does it mean?

- **DESIGN** of national labour ward registers varied between the study countries, capturing between 35 and 58 data elements, duplicative with other recoding in other documents. Coverage indicators of interest (uterotonics, early initiation of breastfeeding and neonatal bag-mask-ventilation) are recorded in registers in Bangladesh and Tanzania but not in Nepal. Standardisation of registers and linkage of these registers to digital HMIS is urgently needed for global tracking. Registers also need local ownership to streamline with local facility documentation requirements, this is critical to reduce burden on frontline health workers.
- **FILLING** processes of routine registers are not systematically implemented within or between countries. Completeness was more highly valued than accuracy. Consistency and accuracy could be promoted by training and supportive supervision to realize the potential of this data source.
- **USE** of register data are impeded by lack of trust in its quality. Promotion of the importance of health facility data for clinical quality improvement, and monitoring is needed to improve data quality and use. Feedback from data users at supervisor/manager and district levels could increase the value frontline health workers attribute to these data and promote their use at the place of care.

What next and research gaps?

- Routine labour ward register data can be used now to contribute vital data around the time of birth. Implementation research is required on interventions to standardise labour ward register designs, and the processes for filling them with regular data quality review. Such research could test an improvement package to include a two-way data flow system up from labour ward registers into HMIS, and feedback returning to the facility.

Background

Data gaps to track care around the time of birth in low- and middle-income country (LMIC) settings impede action towards goals to end more than 5 million deaths annually of newborns, stillbirths and women [1–4]. Although > 80% of the world's births occur in facilities [5], routine records are under-utilised as a data source for maternal and newborn care. The *Every Newborn* Action Plan (ENAP), agreed by all United Nations member states and > 80 development partners, includes an ambitious measurement improvement roadmap with an urgent focus to improve measurement around the time of birth, especially routine Health Management Information System (HMIS) data [6]. Sustainable Development Goal 17 "Revitalise the global partnership for sustainable development" includes a specific target to increase the availability of high-quality, timely and reliable data [7]. Population-based surveys remain a major source of maternal and child health data in LMIC [8–10]. Such household surveys— e.g. the Demographic and Health Surveys (DHS) Program [11] and Multiple Indicator Cluster Surveys (MICS) [12, 13]— collect information regarding births over the preceding 2 to 5 years, thus are not designed to tracking progress on a month-to-month, or year-to-year basis [1, 14–16].

Routine HMIS data, in contrast, have potential to be available more regularly and used for more timely action by health workers, facility/district managers and policy makers [17]. The expansion of digital platforms e.g. District Health Information Software 2 (DHIS-2) in LMICs in recent years has increased awareness of the potential of HMIS to improve data availability at the subnational level and above [18]. Whilst household surveys are designed to be representative of populations, as institutional births rise, facility HMIS data is becoming increasingly useful. However, HMIS data quality has historically been considered poor, so increasing data quality and trust are essential [19, 20]. Studies in LMICs have shown how data use positively impacts quality of care and helps strengthen health systems [21, 22]. The performance of routine information system management (PRISM) framework illustrates the multiple factors (organisational, technical and behavioural) that influence data quality and information use (Table 1) [23, 24]. Routine register data are usually the source for HMIS facility data. Paper registers are books, typically located on a hospital ward; they contain a limited number of data elements as a parallel and usually duplicate system to individual patient case notes. Health workers record each admitted

Table 1 Performance of routine information system management (PRISM) conceptual framework components

Type	Category	Content
INPUTS RHIS Determinants	**Technical Factors**	Complexity of reporting forms, procedures
		HIS design
		Computer Software
		Information technology complexity
	Organisational Factors,	Governance
		Planning
		Training
		Supervision
		Quality
		Finance
		Promotion of culture of information
		Availability of resources
	Behavioural factors	Level of knowledge of content of HIS forms
		Data quality checking skills
		Problem solving for HIS tasks
		Competence in HIS tasks
		Confidence levels for HIS tasks
		Motivation
		Demand
PROCESS steps	**RHIS processes**	Data collection
		Data transmission
		Data processing
		Data analysis
		Data quality check
		Feedback
OUTPUT desired	**Improved RHIS performance**	Data quality/information use
OUTCOME desired	**Improved Health System performance**	
IMPACT desired	**Improved health status**	Improved health status

References: PRISM framework and monitoring framework for ending preventable maternal mortality [23, 24]

individual women/newborn on one row in the register with data regarding care practices and interventions in columns allotted either for "specific" data elements (e.g. bag-mask-ventilation) or "non-specific" data elements (e.g. other details). Previous studies have assessed availability and completeness of data elements for maternal and newborn coverage indicators in routine registers [25, 26]. Data for local and higher health system use in HMIS are typically aggregated from registers monthly, using paper tally sheets and/or summary forms. Data culture within the health facility influences register data collection, analysis and use [27].

The *Every Newborn* – Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study was a mixed methods observational study in three countries (Tanzania, Bangladesh and Nepal). EN-BIRTH aimed to assess measurement validity of newborn and maternal indicators for routine facility-based tracking of coverage, quality of care, and outcomes (21). Indicators were selected based on criteria outlined in global frameworks [6, 28, 29]. The EN-BIRTH validation assessment reported finding register-recorded coverage accuracy varied by indicator and by hospital [30].

Objectives

This paper is part of a supplement based on the EN-BIRTH multi-country validation study, *'Informing measurement of coverage and quality of maternal and newborn care'*. The purpose of this paper is to explore general barriers and enablers for health workers to record high-quality (complete and accurate) data in labour ward registers only. Data recorded in registers in neonatal and kangaroo mother care wards are explored in other papers in the supplement [31, 32]. This paper has three objectives:

1. **Describe the STRUCTURE OF ROUTINE LABOUR WARD REGISTERS** for measurement of coverage of key maternal and newborn health intervention indicators.
2. **Identify BARRIERS AND ENABLERS** for health workers to record and use labour ward register data for measurement of coverage of key maternal and newborn interventions.
3. **Explore the PROCESSES** of labour ward health care provision and register documentation including flow and sequence, by health workers for key maternal and newborn interventions at birth.

Methods

Study sites and overview

EN-BIRTH study was conducted in five public hospitals in three high-burden mortality countries: Maternal and Child Health Training Institute, Azimpur and Kushtia District Hospital in Bangladesh (BD), Pokhara Academy of Health Sciences in Nepal (NP), and Temeke Regional Hospital and Muhimbili National Referral Hospital in Tanzania (TZ) (Additional file 1). These comprehensive emergency obstetric and newborn care (CEmONC) hospitals were selected since they provided the interventions of interest across several different wards. Labour ward register findings for three indicators (uterotonics to prevent post-partum haemorrhage, early initiation of breastfeeding and neonatal bag-mask-ventilation) will be reported in this manuscript; other ward findings are reported in separate manuscripts [31–33]. The multi-partner research team co-designed the tools and collected data from January 2018 to March 2019.

Objective 1: Structure of routine labour ward registers

We reviewed the design structure for labour ward registers to summarise: total number of data elements captured; selected indicator data elements column name, column type (specific or non-specific) and how the column was completed if the intervention was either given or not given.

Objective 2: Barriers and enablers to record and use register data

The research team, using a literature review, identified the PRISM conceptual framework (Table 1) and used these constructs to design guides for semi-structured in-depth interviews (IDI) and for focus group discussions (FGD) (Additional file 2). The guides explore routine labour ward register documentation in general, with specific open-ended questions about selected indicators (Additional files 3, 4) [33]. Tools were developed in English, translated to local languages (Bengali, Nepali and Swahili), then piloted, revised and back-translated into English.

Respondents and data collection

We purposively selected two groups of respondents: (i) Health workers (nurses/midwives/doctors) from the study hospitals routinely caring for women/newborns and are responsible for recording in ward registers; and (ii) EN-BIRTH study researchers (clinical observers, data extractors and supervisors) who were present for more than 9 months on the study site wards, for an external perspective on the register documentation process [33].

At least two IDIs were conducted in each site for each category of respondent. The sample size for the interviews was determined using saturation sampling: additional respondents were interviewed until no new information was learnt by the investigators in each site. One FGD including at least one health worker from each ward was added for triangulation. Data were collected by experienced qualitative researcher co-authors in two phases: January–June 2018 for EN-BIRTH study data collectors and January–March 2019 for EN-BIRTH study hospital health workers. Interviews were conducted in local languages in a private room and audio recorded after obtaining informed participant consent.

Data management and analysis

Data transcription, translation into English, codebook design and analysis were carried out by the same co-authors involved with tool design and data collection after all data had been collected. All transcripts were read multiple times by team members prior to developing the codebook for familiarization. A coding template in NVivo software version 12 [34] was jointly developed based on the PRISM framework (Additional files 5, 6) and the codebook. Framework analysis was used to support comparing, and to differentiate between IDI and FGD findings [35]. Two coders from each country team coded the same 2–4 interviews and compared results. Any discrepancies were discussed, which increased inter-coder reliability [36]. Differences were reconciled through discussion or involvement of another team member, and single individuals coded remaining transcripts. The multi-country team reconciled coding issues on weekly calls and the codebook was modified where necessary.

For the health worker-register interface, the EN-BIRTH team created a framework based around three categories: register design, register filling and register use. We applied this conceptual framework to identify emerging themes across all sites. Two analysis workshops and multiple multi-country calls were held to agree upon the main themes emerging from the IDIs and FGDs, and to synthesise the findings. The consolidated criteria for reporting qualitative research (COREQ) checklist guidelines were followed throughout (Additional file 7) [37].

Objective 3: Processes of care and documentation including flow and sequence

To identify how health care provision and labour ward register documentation relate to one another on labour ward, we designed a third tool called the "care-to-documentation checklist" (Additional file 8). This tool captured the process, flow and sequence of recording data in registers by selected indicators: which health worker cadre usually/sometimes provides the care; which cadre records the care; what is the order of documentation in labour ward documents (among register, patient notes, drug charts, partograph); what is the estimated time in minutes between intervention given and documentation. These close-ended questions were asked by the researcher to respondents, immediately after their IDI (but not to FGD respondents). The checklist data were entered on Excel and proportions and sequence were analysed in R version 3.6.1 [38].

Results

Objective 1: Structure of routine labour ward registers

We identified two types of registers on the labour wards: formal pre-printed and informal hand-written registers, which are typically facility-specific for programme or quality improvement purposes (Additional file 9). All study hospitals used nationally developed, formal paper-based registers; in Bangladesh, a national register was introduced during the early phase of the study, replacing previously existing, hospital-specific ones. In Muhimbili TZ, the informal "Perinatal Research Register" has been in continuous use for more than 20 years [39]. In Temeke TZ, one nurse-midwife was assigned to send summary data every day from the register to HMIS and had no other clinical responsibilities. The total number of data elements captured in formal register columns was: 58 in Bangladesh, 35 in Nepal and 48 in Tanzania (Table 2). One data element was captured per column in the register in Tanzania, but more than one in some register columns in Bangladesh and Nepal. Data elements needed as numerators for the three selected coverage indicators were captured in the Bangladesh and Tanzania registers but not in the Nepal register. In Bangladesh register columns were ticked when the intervention/practice was done and left blank when not done; in Tanzania, register columns were filled with yes/no in Swahili, except for bag-mask-ventilation, which was completed with a numerical code (Additional file 10).

Objective 2: Barriers and enablers to record and use register data

A total of 72 health workers (62 nurse-midwives and 10 medical doctors) and 65 data collectors were interviewed for this study (Table 3); background characteristics of participants are shown in Additional file 11.

As shown in Fig. 1, participants reported that these common themes could either serve as barriers or enablers to recording and using register data. The themes are shown radiating from the conceptual framework to illustrate how these themes were described as influencing one another and the hospital data culture. Each theme is summarised in turn below.

Table 2 Ward routine register designs capturing selected newborn and maternal indicators, EN-BIRTH study

Register design	BD - Azimpur		BD - Kushtia		NP - Pokhara	TZ - Temeke	TZ - Muhimbili	
	Tertiary		District		Regional	Regional	National	
Labour and Delivery Ward								
Register name	Delivery register	EmONC register	Delivery register	EmONC register	Maternity Register	Delivery book	Delivery book	Perinatal research register
Register format	Original hospital	Revised national	Original hospital	Revised national	National	National	National	Additional research
Number of data elements	25	58	24	58	35	48	48	47
Number of columns:								
total	20	45	18	45	32	48	48	39
for uterotonics	1	1	1	1	0	2	2	2
for early breastfeeding	0	1	0	1	0	2	2	2
for neonatal resuscitation	0	1	0	1	0	1	1	1

Details regarding selected indicators in Additional file 10
Note: register designs may record more than 1 data element per column. BD = Bangladesh, NP = Nepal, TZ = Tanzania

Table 3 Summary of research methods assessing barriers and enablers to labour ward register documentation, EN-BIRTH study

Method	Description of the method	Duty ward	Responsibility	Selected indicator documented explored
Heath workers:				
a) In-depth interviews and c) care-documentation checklist	Nurses/midwives (n = 3 per hospital, total n = 15)	Labour and Delivery	Care for patient and document	• Uterotonics to prevent PPH • Early initiation of breastfeeding • Neonatal bag mask ventilation
	Doctors (n = 1 per hospital, total n = 5)	Labour and Delivery included	Care for patient and document	All selected indicators
b) Focus Group Discussion	Nurses/midwives from each ward (n = 1 FGD per hospital, total n = 5)	Labour and Delivery included	Care for patient and document	All selected indicators
EN-BIRTH data collectors:				
a) In-depth interviews and c) care-documentation checklist	Data Trackers (n = 3–4 per hospital, total n = 19)	Registered patient at start of study	Observed care process and some content of documentation	Not applicable
	Clinical observers (n = 4–8 per hospital, total n = 24)	All wards	Observed care process but not content of documentation	All selected indicators
	Data Verifier/Extractor (n = 1–4 per hospital, total n = 13)	All wards	Extracted data from registers and patient notes for EN-BIRTH study	All selected indicators
	Supervisors (n = 1–2 per hospital, total n = 9)	All wards	Observed process and extracted data from registers and patient notes	All selected indicators

Further details of respondents from all wards in Additional file 2

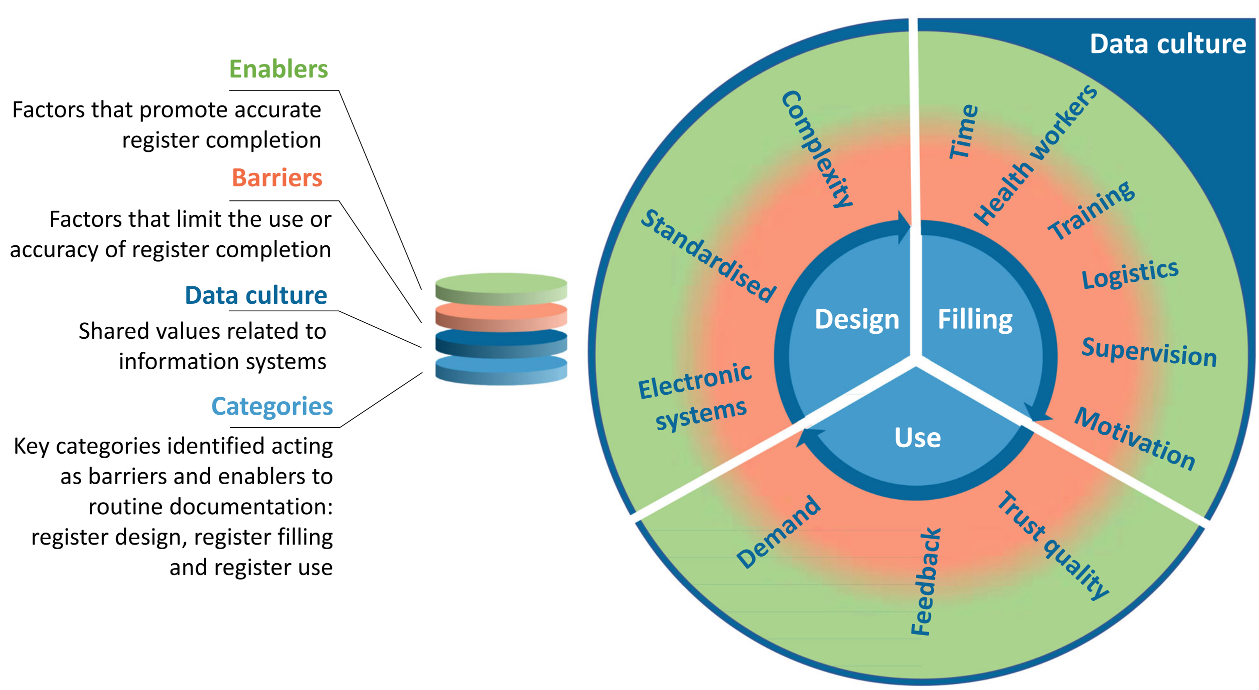


Fig. 1 Barriers and enablers to routine recording of coverage indicators in labour ward registers, EN-BIRTH study. The transition from red (barrier) to green (enabler) serves as a reminder that most factors identified by participants could serve as either a barrier or enabling factor depending on the facility-level resources and management

Register design

Three themes emerged:

Complexity The labour ward registers were described as complex by many respondents in Tanzania and Bangladesh:

> *"It is complicated somehow, first it is large and that book [register] contains a lot of details to be filled although all of them are important … ."*
> -IDI, L&D Nurse-midwife, Muhimbili TZ

Additionally, the data elements recorded in the formal labour ward register need to be duplicated in multiple documents (e.g. informal registers, patient notes), as complex registers form part of a documentation system that is not streamlined and is burdensome:

> *"We need to do the same documentation, again and again in three to four different places, which needs us to give a lot of time."*
> -IDI, L&D Nurse-midwife, Azimpur BD

Standardisation with necessary data elements Health workers from the Nepal and Bangladesh sites acknowledged all the data elements they needed were captured. However, in Tanzania, not all data elements needed to complete monthly reporting forms for HMIS were in the labour ward registers:

> *"I enter entire patient's information … and I sometimes have to add some columns where I can include some data that I know is important and should be written to help me with my end of the month report. So, if I were to just follow the register it means some data could be missed and that's the challenge that I encounter."*
> -FGD, Nurse-midwife, Muhimbili TZ

Paper or electronic All five hospitals were using paper-based registers, but respondents mentioned forthcoming transition to electronic platforms, which were anticipated to be desirable, to save time, and to improve data completeness, availability, and storage:

> *"Documentation till today is done in traditional way. However, writing that every day, is time loss. Further, if we had computerised system, it would have been very better, it could last for later."*
> -IDI, L&D Nurse-midwife, Pokhara NP

Yet many respondents expressed their need for computer training, some suggested extra staff would need to be recruited to manage digitised registers:

> *"To operate the computer for documentation, we need both manpower and proper training. For example, if we had three more staffs in this ward, two staffs will work for caring the patient and the other one will engage with documentation and can handle the computer. It will allow us to perform other things more easily."*
> -IDI, L&D Nurse-midwife, Azimpur BD

Register filling

Six themes emerged:

Health worker responsibility In all five hospitals nurse-midwives alone owned the task of labour ward register recording described as within their current nursing role. Data quality responsibility was perceived to be better when the same nurse-midwife providing the care documented in the register:

> *"For effective recording and reporting, the one who provides the care should herself do the documentation and then only it is complete and proper. A third person asked to document is not proper – there will be missing in recording and reporting. Manpower should be sufficient so the one who does the care should only perform recording and reporting."*
> -FGD, Nurse-midwife, Pokhara NP

However, task shifting of documentation to other actors was highly valued by several respondents, although difficult to obtain, especially during night shifts:

> *"It is super difficult to get support from students even the intern doctor and the trainee nurses don't help us in documenting the information in register."*
> -FGD, Nurse-midwife, Kushtia BD

Training for competence Respondents from Nepal described attending a short training on register filling as an enabling factor for register data recording. Tanzanian respondents stated they had been shown "on the job" how to fill the register and the lack of specific formal training or instructions for register filling was a barrier to documentation. In Bangladesh, only computer training had been received:

> *"We have not got any training related to register fill up. We were given only an orientation on computer but couldn't learn anything. It was too short i.e. 2 to 3 days."*
> -IDI, L&D Nurse-midwife, Azimpur BD

Time required to document Respondents expressed the large amount of time spent on documentation in general, even in the Nepal site with the lighter register design:

"If we work 7 to 8 hours duty, it usually takes around 3 hours to do documentation."
-IDI, Nurse-midwife, Pokhara NP

"In a period of 8 hours of my shift, if I have a large number of patients, I may spend more time in documentation than the time I spend in attending the patients."
-IDI, L&D Nurse-midwife, Muhimbili TZ

In all three countries, respondents related the time challenge of completing registers to the availability of the health workforce:

"Our main difficulty to fill up the register appropriately, is shortage of manpower. We have to suffer a lot to do quality documentation due to short of manpower."
-IDI, L&D Nurse-midwife, Azimpur BD

The tension between being too busy to always document immediately after care led to lower quality data:

"You find you are having say three patients and they all need care, you will start with the first one, after that you can't do the documentation, you will have to attend the second and the third, now as you go for documentation it will be difficult to remember exactly figures or details, for example it is difficult to remember exactly the time for each of them so, you will have to estimate, maybe if you have enough staff, one does the attending and another do the documentation."
-IDI, Nurse-midwife, Muhimbili TZ

Logistical resources needed New registers were usually available but sometimes the stock were locked in stores. Pens were only available in some hospitals:

"There is still a challenge of resources, for instance now we are asked to document but they don't think if pens are provided, instead you have to buy yourself. You are supposed to write … .and there are some things which I would like to write them if they would provide me with tools. Honestly resource is very challenging"
-FGD, Nurse-midwife, Muhimbili TZ

The organisation of the large registers laying on a table at the nursing station were described as a logistical barrier by some respondents:

"When she is done she will go to the nursing station to do her documentation in register book, then fills the midwifery book, the books are in different places and are far from the patient or the delivery room."
-IDI, EN-BIRTH Data Collector, Muhimbili TZ

Supervision for data quality Supervision of register filling processes was acknowledged to be an important enabler to register filling by most respondents, yet was not occurring regularly in every hospital:

"We never had any sorts of supervision about the documentation."
-IDI, L&D Nurse-midwife, Azimpur BD

"The only things that displays the work of health workers are the documentations … important for supervisors as well. If we show them the recorded data, they get to advise us about the errors and whether it [register] is complete or not. So it becomes important in supervision as well."
-IDI, L&D Nurse-midwife, Pokhara NP

Register documentation supervision was expressed as being linked to data quality:

"They normally come to verify their data on register books and if there is any problem, they tell you that here you are supposed to do this and that. This is how is being done … It is educative system because if she criticise you she must explain to you."
-FGD, Nurse-midwife, Temeke TZ

Many respondents expressed that completeness was important and the need to "fill the gaps" in registers:

"There is a big delivery book which has headings therefore, you can't skip even a single box all of them must be filled."
-FGD Nurse-midwife, Muhimbili TZ

Motivation Appreciation from supervisors was articulated by one respondent as an important motivator, and was also linked with higher quality documentation:

"We receive praise, when everything (related to documentation) is good and it works as a motivation

to continue documentation with care."
-IDI, L&D Nurse-midwife, Azimpur BD

By contrast, many health workers noted the lack of acknowledgement and/or recognition served as a motivational barrier for high quality register recording:

"There is not any formal award or recognition like that. Instead we get scolded if it's left. We are not appreciated for writing."
-FGD, Nurse-midwife, Pokhara NP

Register use

Three themes emerged regarding perceived register data utility:

Demand for data Respondents expressed varied register data demands as enablers. Nurse-midwife respondents mainly described how they themselves used the data for patient handover:

"We are documenting because even nursing itself is a continuous process … so if you did not document, the other nurse will not know where you ended, so documentation is still very important."
-FGD, Nurse-midwife, Temeke TZ

The same register data were used by supervisors for management decisions:

"Even the hospital itself insists so much on documentation... if you don't document, sometimes it becomes very difficult for the management to get revenue to know how many people should get what medicine, you have to document on health insurance and normal patients."
-FGD, Nurse-midwife, Temeke TZ

In Nepal, a doctor respondent expressed that data were used in research and for indicators:

"We also have doctors and students utilising the data. It is used for the research and general information. We create health indicators and send to central level and they also create national health indicators. And the ultimate goal for all is to know how the health indicators are. It helps to do planning accordingly."
-IDI, L&D Doctor, Pokhara NP

Feedback to health workers Provision of feedback from HMIS users of register data to those who had collected the data was perceived to be an enabler; however, respondents said feedback hardly ever happened:

"I haven't got any feedback from them (HMIS) about documentation. There sits monthly meeting in hospital with data people. We don't usually participate in that meeting."
-IDI L&D, Nurse-midwife, Azimpur BD

"It doesn't come to us directly. We don't have much information."
-FGD, Nurse-midwife, Pokhara NP

Trust in data quality Some health worker respondents stated that lack of trust in register data quality was a barrier to the usefulness of register data:

"Sometimes, variables are missing and when research needs to be done then it is not ineffective."
-IDI, L&D Doctor, Pokhara NP

"There is hardly missing areas in the register- if we find some we try to collect the information either by asking the patient again or nurse who attended the delivery. Using good quality data are important to decision make."
-IDI, L&D Nurse, Kushtia BD

Objective 3: Sequence of care and documentation

Analysis of the care-to-documentation checklist showed that the nurse-midwife who provided the intervention/practice usually also recorded in the labour ward register (Additional file 12). However, data collector respondents stated that health workers sometimes documented care provided by a colleague (Additional file 13). Among all documents to be filled, the labour ward register were described as the first to be completed in both Bangladeshi hospitals, but the order varied between first to third in the Tanzanian hospitals (Additional files 14, 15). The average estimated time between care provision and register documentation ranged from 10 to 28 min as reported by health workers and was 9 to 34 min based on data collectors' report (Additional file 16).

Discussion

EN-BIRTH study is the first LMIC multi-country assessment of barriers and enablers to labour ward register data recording. We add to previous research regarding barriers to routine facility data recording from antenatal clinics and HIV/AIDS programme data [1, 40, 41]. We found twelve consistent themes reported in all five hospitals across our conceptual framework of register design, filling and use. Figure 1 depicts the interconnected relationship between register data use, register design, and register filling. The twelve themes identified within these categories were described as either enablers or barriers by respondents in the five hospitals. We postulate that the varying interaction of these themes in each study hospital contributed to the variation in accuracy in measurement of labour ward indicators as identified in the EN-BIRTH validation study [30]. These data practice themes act within, and likely contribute to, a wider hospital data culture of accepted and normative practices, which permits health workers to collect high-quality register data that can be trusted for use.

Improved HMIS performance is increasingly recognized as a priority to improve coverage and quality of care as described in the comprehensive PRISM framework, which demonstrates the many interacting constructs needed for high-quality data for use [23, 24]. This EN-BIRTH study used the PRISM constructs to explore the barriers and enablers to recording at the service user-register interface and for health workers. We found register design complexity and the burden of data collection were common across the study sites. The sheer volume of data elements captured in these national register designs was striking. Nepal had the lightest register design, yet captured 35 data elements, compared to 48 in Tanzania and 58 in Bangladesh. Notably, data elements more than doubled when national registers were introduced in Bangladesh. Yet labour ward registers did not always match monthly reporting requirements, necessitating nurse-midwives to use their own initiative and add columns to registers, or start informal registers, to capture required data. Complexity of documentation was described as encroaching upon the time health workers can dedicate to midwifery care. Our findings align with a study describing the balance between service provision and documentation practices in Uganda [42]. Several causes contribute to this high burden of register data collection, including a lack of coordination regarding which indicators (and contributing data elements) are selected for tracking, multiple reporting flows and additional data element capture to signal rigor or research [43]. Frontline health workers have dual responsibilities of providing care and documentation of that care. With the typically high user-to-staff ratios of facilities in many LMIC settings, urgent attention to reducing any unnecessary documentation would support efforts to improve quality of care by health workers for women and babies.

Filling of registers was not systematised or consistently supported by effective logistics and supplies, even non-availability of pens and registers was cited by some respondents. Bedside care provided by the health worker was documented in one register located on a table in the labour ward. The documentation was described as done within 30 min of the care practice/intervention whilst the health worker was still responsible for the women and her baby during the critical first hour after birth. The cumulative effect of distance between point of care and point of register documentation, simultaneous responsibilities of care and documentation for a large number of data elements to be recalled could account for both under and over-reporting of interventions, as found in the EN-BIRTH observational validation study [30].

Perceived value of labour ward register data by data users in these large CEmONC hospitals was a further cross-cutting issue that likely affects data quality [30]. Data-specific training was perceived by health workers as enabling, yet few had received in-service training on how to complete registers. Supportive supervision for register recording was not a priority, as described by both health workers and research data collectors. Data completeness was expressed as more highly valued compared to data accuracy by health workers and data collectors alike. This may be driven by column filling (completeness) being feasible to visualise in registers by health workers and supervisors, and thus a signal and symbol of professionalism [44]. Although notably in Bangladesh completeness for coverage numerators cannot be calculated, as registers are designed for columns to be left blank (true zero) when interventions are not performed.

Use of register data was valued by health workers for clinical care handover or other hospital use, however none of the nurse-midwife respondents who actually fill registers mentioned use for tracking coverage or impact of services at higher levels of the health system. Increasing demand for labour ward register data use is needed. Using register data at facility level to improve quality of care or to supervise performance was mentioned could link to priority setting and health unit management also at sub-national level. National data demand includes for strategic planning and policy. Health workers around the world invest considerable time documenting large volumes of data. Nurse-midwives deserve to be informed about the value of the data they collect for wider decision making, and to be appreciated for their work in collecting it.

Enabling environments are needed for health workers to provide care and are often measured as "service readiness" [45]. Similarly, enabling "data readiness" is necessary to promote high-quality register data to flow into HMIS. An integrated approach is needed to transform routine data on labour wards, taking into account the midwife's dual role in care provision and data recording [20]. The information culture at the facility level and throughout the system is important. Decentralised data use in facilities may incentivise improving data quality [46, 47]. By increasing data visibility through feedback to frontline health workers about data use, data quality has been shown to improve in registers [14, 19, 22–24, 48–50]. However, a notable finding from our labour ward register study was the low level of two-way feedback loops between different levels of the data pyramid: nurse-midwives collecting register data and other data users higher up in the pyramid [51, 52].

Paper-based systems remain the norm in most LMIC labour wards, yet these often feed into digital systems [53]. However, care should be taken to not just digitise poor information systems. There has been rapid expansion of digital HMIS in LMIC with increased IT capability to improve data quality (automated checks, validation rules, visualizations, etc.) [1, 46, 47, 54]. Poor quality of care has been described as "too much too soon, too little too late" [55]. Similarly, in response to "too little data too late", care is needed to avoid digitisation of routine data creating "too much data too soon". Unless we turn our attention to reduce unnecessary data and improve reliability and quality of the register data, the value of digital HMIS data for clinical and programmatic decision making will not be realised. The risk is that labour ward routine register data will remain in a "vicious cycle of data quality", if data are not trusted, they are not used. If data are not used, investment in data quality suffers, and data quality deteriorates even further. Thus, simultaneous action on both data use and data quality is necessary to break this cycle. In practice, this means increasing use of current labour ward register data, whilst investing in improving data quality. Current data quality reviews typically compare HMIS monthly reports using register data as the standard [56]. Innovative ways to routinely include assessment of the quality of the source register data are important to consider. Register data assessment can be linked to routine quality improvement initiatives that use routine data, such as maternal and perinatal death surveillance and response. Checking accuracy of register data quality compared to patient case notes during such perinatal audit meetings and involving health workers could be one effective way for feedback and linking quality of data with quality of care. Without focused action to improve routine data quality, tracking progress using HMIS data towards agreed Sustainable Development Goals and ENAP targets by 2030 will be suboptimal [53].

Strengths and limitations

A strength of our study is multi-sites public hospitals in three high mortality burden LMICs. We used common tools that were co-designed by our team including the PRISM framework determinants. We interviewed health workers involved in the process themselves and, for an external perspective, EN-BIRTH research data collectors who had worked day and night on the labour ward for > 9 months. The use of open-ended and close-ended questionnaires allowed us to generate a broad range of common findings issues across sites. Our predetermined codes were based on the PRISM framework and all sites used NVivo in a collaborative analysis process.

However, our study also has limitations. There was a possible desirability bias by health workers, which might have led to either under- or over-reporting of the challenges faced. The "care-to-documentation checklist" dataset analysis was stratified by type of respondent (health worker and data collector), by indicator and by site. The qualitative data analysis presented in this paper identified common barriers and enablers for labour ward register recording across all indicators, using health worker and data collector responses together. Indicator-specific mixed-methods linked analyses will be presented in other linked papers to further explore subthemes and differences between cadres [57–62]. It was beyond the scope of this study for the EN-BIRTH data collectors to directly observe or measure the detailed process of register filling (e.g. time, logistics availability, supervision, use for reports). All hospitals were peri-urban CEmONC hospitals, which may limit generalizability to facilities at lower levels of the health system.

Research for improving measurement

Further research is needed to explore barriers and enablers in other settings and at different levels of the health system to understand the broader relevance of the themes we identified. Our exploratory research identified twelve themes that could be used to design shorter tools for routine register data capture and use, a component of HMIS that is relatively under-represented in existing tools [27, 56]. Implementation research is required for all three components we identified regarding registers in our conceptual framework (design, filling, use). To enable national or district tracking of core indicators in HMIS, the priority data elements that are being harmonized at higher levels of the data pyramid will need to be included in register design [63, 64]. Register data element availability is necessary but not sufficient; more research is required to explore whether register layout, column labelling and cell coding affect data quality. For example, facilities might consider excluding blank cells from their register design, as blank cells may indicate a health procedure either "not recorded" (incomplete) or "not done". Standardised register designs will require local ownership for adaptation, and testing of process, with considerable streamlining with other documentation, to reduce burden on frontline health workers. Research regarding improved register filling may focus on capability (capacity to engage in the register documentation), opportunity (factors that make the behaviour possible) and motivation (to energies and direct behaviour). Exploring flow of aggregated data from labour ward registers into HMIS is another gap requiring research regarding steps of aggregation. Several manual operations (e.g. manual counting, filling paper summary/tally forms, digital data entry) may reduce data quality significantly [65]. Finally, perspectives of data users beyond the patient-health worker-register interface are critical. Yet, to date, there has been little investment in improving routine register data quality to maximize the potential of this underused and widely available data source around the time of birth.

Conclusion

With more than 80% of the world's births in facilities, labour ward register data have an unrealised potential to track core indicators in facilities and higher up the health system. Our multi-country study found multiple opportunities to improve the data and the use of data: standardised design, consistent filling processes and enabling two-way feedback between different levels of the health system data pyramid. Overcoming these barriers would enable frontline health workers, especially midwives, to be valued for the register data they collect, to improve data quality and importantly to use those data to improve quality of care for the women and babies they care for.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12884-020-03517-3.

Additional file 1. National context and number of births in EN-BIRTH study hospital.

Additional file 2. Summary of qualitative research methods to assess barriers and enablers to labour/newborn ward register documentation, EN-BIRTH study.

Additional file 3. Health Worker study guides in-depth interview (IDI) focus group discussion (FGD), EN-BIRTH study.

Additional file 4. Data Collector study guides in-depth interview (IDI), EN-BIRTH study.

Additional file 5. Codebook Health Workers, EN-BIRTH study.

Additional file 6. Codebook Data Collectors, EN-BIRTH study.

Additional file 7. COREQ checklist, EN-BIRTH study.

Additional file 8. Care-to-documentation checklist, EN-BIRTH study.

Additional file 9. Labour and Delivery Registers, formal and informal, EN-BIRTH study.

Additional file 10. Labour ward routine register column design for maternal and newborn indicators, EN-BIRTH study.

Additional file 11. Demographic characteristics of respondents for barriers and enablers objective, labour/newborn wards, EN-BIRTH study.

Additional file 12. Labour ward care/documentation responsibilities by intervention, health worker respondents, EN-BIRTH study.

Additional file 13. Labour ward care/documentation responsibilities by intervention, research data collector respondents, EN-BIRTH study.

Additional file 14. Labour ward register order within all documentation, by indicator - health worker respondents, EN-BIRTH study.

Additional file 15. Labour ward register order within all documentation, by indicator - data collector respondents, EN-BIRTH study.

Additional file 16. Estimated minutes between care and documentation by indicator, care-documentation checklist, EN-BIRTH study.

Additional file 17. Ethical approval of local institutional review boards for EN-BIRTH study.

Abbreviations

BD: Bangladesh; CEmONC: Comprehensive emergency obstetric and newborn care; CIFF: Children's Investment Fund Foundation; DHS: The Demographic and Health Surveys Program; DHIS2: District Health Information Software 2; ENAP: *Every Newborn* Action Plan now branded as *Every Newborn*; EN-BIRTH: *Every Newborn*-Birth Indicators Research Tracking in Hospitals study; FGD: Focus Group Discussions; HIV: Human immunodeficiency virus; HMIS: Health Management Information Systems; icddr,b: International Centre for Diarrhoeal Disease Research, Bangladesh; IDI: In-depth interview; LMIC: low-and middle-income countries; MICS: Multiple Indicator Cluster Survey; NP: Nepal; PRISM: Performance of Routine Information System Management; TZ: Tanzania; COREQ: Consolidated criteria for reporting qualitative research

Acknowledgements

Firstly, and most importantly, we thank the women who were part of EN-BIRTH study and the health workers and data collectors. We credit the inspiration of the late Godfrey Mbaruku. We thank Claudia DaSilva, Veronica Ulaya, Mohammad Raisul Islam, Sudip Karki and Rabina Sarki for their administrative support and Sabrina Jabeen, Goutom Banik, Md. Shahidul Alam, Tamatun Islam Tanha and Md. Mohsiur Rahman for support during data collectors training.
We acknowledge the following groups for their guidance and support:
National Advisory Groups:
Bangladesh: Mohammod Shahidullah, Khaleda Islam, Md Jahurul Islam.
Nepal: Naresh P KC, Parashu Ram Shrestha.
Tanzania: Muhammad Bakari Kambi, Georgina Msemo, Asia Hussein, Talhiya Yahya, Claud Kumalija, Eliudi Eliakimu, Mary Azayo, Mary Drake, Honest Kimaro.
EN-BIRTH Expert Advisory Group: Agbessi Amouzou, Tariq Azim, Debra Jackson, Theopista John Kabuteni, Matthews Mathai, Jean-Pierre Monet, Allisyn C. Moran, Pavani K. Ram, Barbara Rawlins, Jennifer Requejo, Johan Ivar Sæbø, Florina Serbanescu, Lara Vaz.
We are also very grateful to fellow researchers who peer-reviewed this paper.

About this supplement

This article has been published as part of BMC Pregnancy and Childbirth Volume 21 Supplement 1, 2021: *Every Newborn* BIRTH multi-country validation study: informing measurement of coverage and quality of maternal and newborn care. The full contents of the supplement are available online at https://bmcpregnancychildbirth.biomedcentral.com/articles/supplements/volume-21-supplement-1.

Authors' contributions

The EN-BIRTH study was conceived by JEL, who acquired the funding and led the overall design with support from HR. For this paper, DS reviewed the literature with HR. LTD led the design of the tools and co-ordinated the multi-country group which was led by DS in Tanzania, SBZ in BD and AKS in NP. LTD and JEL conceptualised the framework. LTD summarised the registers with assistance from HR. DS and LTD led the qualitative analysis with assistance from SBZ and AKS. LTD and KP led the analysis of the care-documentation checklist with assistance from SBZ. LTD and DS designed the figures with KP. DS and LTD drafted the manuscript with inputs from SBZ and the multi-country group and JEL. Authors made substantial contributions to the conception, design, data collection or analysis or interpretation of data for the work including: icddr,b Bangladesh: SBZ with RH, JK, MTUST, RH, TM, MHR, MZHS; Golden Community, Nepal: AKC with AKS, NT; Ifakara Health Institute, Tanzania: DS with MNT; LSHTM: LTD with KP, HR, JS, CH, NS, JEL. Other authors: JIS, LMEV, JR. All authors revised the manuscript and gave final approval of the version to be published and agree to be accountable for the work. The EN-BIRTH study group authors made contributions to the conception, design, data collection or analysis or interpretation of data. This paper is published with permission from the Directors of Ifakara Health Institute, Muhimbili University of Health and Allied Sciences, icddr,b and Golden Community. The authors' views are their own, and not necessarily from any of the institutions they represent, including UNICEF.

EN-BIRTH Study Group
Bangladesh: Qazi Sadeq-ur Rahman, Ahmed Ehsanur Rahman, Tazeen Tahsina, Sojib Bin Zaman, Shafiqul Ameen, Tanvir Hossain, Abu Bakkar Siddique, Aniqa Tasnim Hossain, Tapas Mazumder, Jasmin Khan, Taqbir Us Samad Talha, Rajib Haider, Md. Hafizur Rahman, Anisuddin Ahmed, Shams El Arifeen.
Nepal: Omkar Basnet, Avinash K Sunny, Nishant Thakur, Rejina Gurung, Anjani Kumar Jha, Bijay Jha, Ram Chandra Bastola, Rajendra Paudel, Asmita Paudel, Ashish KC.
Tanzania: Nahya Salim, Donat Shamba, Josephine Shabani, Kizito Shirima, Menna Narcis Tarimo, Godfrey Mbaruku (deceased), Honorati Masanja.
LSHTM: Louise T Day, Harriet Ruysen, Kimberly Peven, Vladimir Sergeevich Gordeev, Georgia R Gore-Langton, Dorothy Boggs, Stefanie Kong, Angela Baschieri, Simon Cousens, Joy E Lawn.

Funding
The Children's Investment Fund Foundation (CIFF) is the main funder of The EN-BIRTH Study, which is administered via The London School of Hygiene & Tropical Medicine. The Swedish Research Council specifically funded the Nepal site through Lifeline Nepal and Golden Community. We acknowledge the core funders for all the partner institutions. Publication of this manuscript has been funded by CIFF. CIFF attended the study design workshop but had no role in data collection, analysis, data interpretation, report writing or decision to submit for publication. The corresponding author had full access to study data and final responsibility for publication submission decision.

Availability of data and materials
The datasets generated during and/or analysed during the current study are available on LSHTM Data Compass repository, https://datacompass.lshtm.ac.uk/955/.

Ethics approval and consent to participate
This study was granted ethical approval by institutional review boards in all operating counties in addition to the London School of Hygiene and Tropical Medicine (Additional file 17).
Voluntary informed written consent was obtained from all respondents for the qualitative interviews. Participants were assured of anonymity and confidentiality.
EN-BIRTH is study number 4833, registered at https://www.researchregistry.com

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Health Systems, Impact Evaluation and Policy, Ifakara Health Institute, Dar es Salaam, Tanzania. [2]Centre for Maternal, Adolescent, Reproductive & Child Health (MARCH), London School of Hygiene & Tropical Medicine, Keppel St, London, UK. [3]Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh. [4]Golden Community, Kathmandu, Nepal. [5]Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, King's College London, London, UK. [6]International Maternal and Child Health, Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden. [7]Department of Informatics, University of Oslo, Oslo, Norway. [8]Global Public Health Karolinska Institutet, Stockholm, Sweden. [9]International Programs, Population Reference Bureau, Washington DC, USA. [10]UNICEF Headquarters, New York, USA.

Published: 26 March 2021

References

1. Maina I, Wanjala P, Soti D, Kipruto H, Droti B, Boerma T. Using health-facility data to assess subnational coverage of maternal and child health indicators, Kenya. Bull World Health Organ. 2017;95(10):683–94.
2. World Health Organization: Stillbirths. [https://www.who.int/maternal_child_adolescent/epidemiology/stillbirth/en/]. Accessed 20 Aug 2020.
3. World Health Organization: Newborns: improving survival and well-being. [https://www.who.int/news-room/fact-sheets/detail/newborns-reducing-mortality]. Accessed 20 Aug 2020.
4. World Health Organization: Maternal mortality. [https://www.who.int/news-room/fact-sheets/detail/maternal-mortality]. Accessed 20 Aug 2020.
5. UNICEF: The State of the World's Children 2019: Statistical Tables. [https://data.unicef.org/resources/dataset/sowc-2019-statistical-tables/]. Accessed 15 Oct 2020.
6. Moxon SG, Ruysen H, Kerber KJ, Amouzou A, Fournier S, Grove J, Moran AC, Vaz LM, Blencowe H, Conroy N. Count every newborn; a measurement improvement roadmap for coverage data. BMC Pregnancy Childbirth. 2015; 15(2):S8.
7. United Nations: Sustainable Development Goal 17. [https://sdgs.un.org/goals/goal17]. Accessed 21 Aug 2020.
8. Boerma T, Requejo J, Victora CG, Amouzou A, George A, Agyepong I, Barroso C, Barros AJ, Bhutta ZA, Black RE. Countdown to 2030: tracking progress towards universal coverage for reproductive, maternal, newborn, and child health. Lancet. 2018;391(10129):1538–48.
9. Victora C, Requejo J, Boerma T, Amouzou A, Bhutta ZA, Black RE, Chopra M. Countdown to 2030 for reproductive, maternal, newborn, child, and adolescent health and nutrition. Lancet Glob Health. 2016; 4(11):e775–6.
10. UNICEF. The State of the World's Children. Children, Food and Nutrition: Growing well in a changing world. New York: UNICEF; 2019.
11. The DHS Program: the DHS program. What we do. [http://www.measuredhs.com/What-We-Do/Survey-Types/DHS.cfm]. Accessed 21 Aug 2020.
12. Moran AC, Kerber K, Sitrin D, Guenther T, Morrissey CS, Newby H, Fishel J, Yoder PS, Hill Z, Lawn JE. Measuring coverage in MNCH: indicators for global tracking of newborn care. PLoS Med. 2013;10(5):e1001415.
13. Moller A-B, Newby H, Hanson C, Morgan A, El Arifeen S, Chou D, Diaz T, Say L, Askew I, Moran AC. Measures matter: a scoping review of maternal and newborn indicators. PLoS One. 2018;13(10):e0204763.
14. Nutley T, Reynolds H. Improving the use of health data for health system strengthening. Glob Health Action. 2013;6(1):20001.
15. World Health Organization: Data quality review: a toolkit for facility data quality assessment. Module 1. Framework and metrics. 2017. [https://apps.who.int/iris/handle/10665/259224]. Accessed 20 Aug 2020.
16. Stanton C, Rawlins B, Drake M, dos Anjos M, Cantor D, Chongo L, Chavane L, da Luz VM, Ricca J. Measuring coverage in MNCH: testing the validity of Women's self-report of key maternal and newborn health interventions during the Peripartum period in Mozambique. PLoS One. 2013;8(5):e60694.
17. Maïga A, Jiwani SS, Mutua MK, Porth TA, Taylor CM, Asiki G, Melesse DY, Day C, Strong KL, Faye CM, et al. Generating statistics from health facility data: The state of routine health information systems in Eastern and Southern Africa. BMJ Glob Health. 2019;4(5):e001849.
18. DHIS2: Oslo: Health Information Systems Programme. [https://www.dhis2.org/]. Accessed 14 Oct 2020.
19. Nutley T, Gnassou L, Traore M, Bosso AE, Mullen S. Moving data off the shelf and into action: an intervention to improve data-informed decision making in cote d'Ivoire. Glob Health Action. 2014;7(1):25035.
20. Kumar M, Gotz D, Nutley T, Smith JB. Research gaps in routine health information system design barriers to data quality and use in low-and middle-income countries: a literature review. Int J Health Plann Manag. 2018;33(1):e1–9.
21. Wagenaar BH, Hirschhorn LR, Henley C, Gremu A, Sindano N, Chilengi R, Collaborative APP. Data-driven quality improvement in low-and middle-income country health systems: lessons from seven years of implementation experience across Mozambique, Rwanda, and Zambia. BMC Health Serv Res. 2017;17(Suppl 3):830.
22. Kimaro HCaN, José L, The challenges of sustainability of health information systems in developing countries: comparative case studies of Mozambique and Tanzania. Health Inform Dev Countries. 2007;1(1):1–10.
23. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009;24(3):217–28.
24. Hotchkiss DR, Aqil A, Lippeveld T, Mukooyo E. Evaluation of the performance of routine information system management (PRISM) framework: evidence from Uganda. BMC Health Serv Res. 2010;10(1):188.
25. USAID: Maternal and Child Survival Program. What Data on Maternal and Newborn Health do National Health Management Information Systems include?. 2018. [https://www.mcsprogram.org/resource/what-data-on-

maternal-and-newborn-health-do-national-health-management-information-systems-include/]. Accessed 20 Aug 2020.
26. Day LT, Gore-Langton GR, Rahman AE, Basnet O, Shabani J, Tahsina T, Poudel A, Shirima K, Ameen S, CA K, et al. Labour and delivery ward register data availability, quality, and utility - every newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries. BMC Health Serv Res. 2020;20(1):737.
27. MEASURE Evaluation: Performance of Routine Information System Management (PRISM). [https://www.measureevaluation.org/our-work/routine-health-information-systems/performance-of-routine-information-system-management-prism]. Accessed 15 Sept 2020.
28. Moran AC, Jolivet RR, Chou D, Dalglish SL, Hill K, Ramsey K, Rawlins B, Say L. A common monitoring framework for ending preventable maternal mortality, 2015–2030: phase I of a multi-step process. BMC Pregnancy Childbirth. 2016;16(1):250.
29. World Health Organization: WHO technical consultation on newborn health indicators: every newborn action plan metrics, Ferney Voltaire, France, 3-5 December 2014. 2015. [https://apps.who.int/iris/handle/10665/184225]. Accessed 17 Sept 2020.
30. Day LT, Rahman QS, Rahman AE, Salim N, KC A, Ruysen H, Tahsina T, Masanja H, Basnet O, Gore-Langton GR, et al. Assessment of the validity of the measurement of newborn and maternal health-care coverage in hospitals (EN-BIRTH): an observational study. Lancet Glob Health. 2020. https://doi.org/10.1016/S2214-109X(20)30504-0.
31. Salim N, Shabani J, Peven K, Rahman QS, KC A, Shamba D, Ruysen H, Rahman AE, KC N, Mkopi N, et al. Kangaroo mother care: EN-BIRTH multi-country validation study. BMC Pregnancy and Childbirth. 2021. https://doi.org/10.1186/s12884-020-03423-8.
32. Rahman AE, Hossain AT, Ameen S, Salim N, KC A, Day LT, Kija E, Peven K, Tahsina T, Zaman SB, et al. Antibiotic use for inpatient newborn care with suspected infection: EN-BIRTH multi-country validation study. BMC Pregnancy and Childbirth. 2021. https://doi.org/10.1186/s12884-020-03424-7.
33. Day LT, Ruysen H, Gordeev VS, Gore-Langton GR, Boggs D, Cousens S, Moxon SG, Blencowe H, Baschieri A, Rahman AE. "Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. J Global Health. 2019;9(1):010902.
34. Edhlund B, McDougall A. Nvivo 12 Essentials: Lulu. com; 2019.
35. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol. 2013;13(1):117.
36. Barry CA, Britten N, Barber N, Bradley C, Stevenson F. Using reflexivity to optimize teamwork in qualitative research. Qual Health Res. 1999;9(1):26–44.
37. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349–57.
38. R Core Team: R: A language and environment for statistical computing. [http://www.R-project.org]. Accessed 20 Aug 2020.
39. Kidanto H, Massawe S, Nystrom L, Lindmark G. Analysis of perinatal mortality at a teaching hospital in Dar Es Salaam, Tanzania, 1999-2003. Afr J Reprod Health. 2006;10(2):72–80.
40. Nyamtema AS. Bridging the gaps in the health management information system in the context of a changing health sector. BMC Med Inform Decis Mak. 2010;10:36.
41. Nicol E, Dudley L, Bradshaw D. Assessing the quality of routine data for the prevention of mother-to-child transmission of HIV: an analytical observational study in two health districts with high HIV prevalence in South Africa. Int J Med Inform. 2016;95:60–70.
42. Hutchinson E, Nayiga S, Nabirye C, Taaka L, Staedke SG. Data value and care value in the practice of health systems: a case study in Uganda. Soc Sci Med. 2018;211:123–30.
43. Stansfield S, Orobaton N, Lubinski D, Uggowitzer S, Mwanyika H. The case for a national health information system architecture; a missing link to guiding national development and implementation. Bellagio: Making the eHealth Connection; 2008.
44. Feldman MS, March JG. Information in organizations as signal and symbol. Adm Sci Q. 1981:171–86.
45. Donabedian A. The quality of care: how can it be assessed? Jama. 1988; 260(12):1743–8.
46. Inguane C, Sawadogo-Lewis T, Chaquisse E, Roberton T, Ngale K, Fernandes Q, Dinis A, Augusto O, Covele A, Hicks L. Challenges and facilitators to evidence-based decision-making for maternal and child health in Mozambique: district, municipal and national case studies. BMC Health Serv Res. 2020;20(1):1–10.
47. Wickremasinghe D, Hashmi IE, Schellenberg J, Avan BI. District decision-making for health in low-income settings: a systematic literature review. Health Policy Plann. 2016;31:ii12–24.
48. Braa J, Heywood A, Sahay S. Improving quality and use of data through data-use workshops: Zanzibar, United Republic of Tanzania. Bull World Health Organ. 2012;90:379–84.
49. Nutley T. Improving data use in decision making. An intervention to strengthen health systems; 2012.
50. Sæbø JI, Moyo CM, Nielsen P. Promoting transparency and accountability with district league tables in Sierra Leone and Malawi. Health Policy Technol. 2018;7(1):35–43.
51. Mbondji PE, Kebede D, Soumbey-Alley EW, Zielinski C, Kouvividila W, Lusamba-Dikassa P-S. Health information systems in Africa: descriptive analysis of data sources, information products and health statistics. J R Soc Med. 2014;107(1_suppl):34–45.
52. World Health Organization: Survive and Thrive: Transforming Care for Every Small and Sick Newborn. 2018. [https://www.unicef.org/reports/transforming-care-for-every-small-and-sick-newborn-2020]. Accessed 13 Aug 2020.
53. Hagel C, Paton C, Mbevi G, English M. Data for tracking SDGs: challenges in capturing neonatal data from hospitals in Kenya. BMJ Glob Health. 2020; 5(3):e002108.
54. Akanbi MO, et al. Use of electronic health records in sub-Saharan Africa: progress and challenges. J Med Trop. 2012;14(1):1.
55. Miller S, Abalos E, Chamillard M, Ciapponi A, Colaci D, Comandé D, Diaz V, Geller S, Hanson C, Langer A, et al. Beyond too little, too late and too much, too soon: a pathway towards evidence-based, respectful maternity care worldwide. Lancet. 2016;388(10056):2176–92.
56. World Health Organization. Data quality review: a toolkit for facility data quality assessment. Module 3: Data verification and system assessment. Geneva; 2017. [https://www.who.int/healthinfo/tools_data_analysis/dqr_data_verification/en/]. Accessed 20 Aug 2020
57. Kong S, Day LT, Zaman SB, Peven K, Salim N, Sunny AK, Shamba D, Rahman QS, KC A, Ruysen H, et al. Birthweight: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03355-3.
58. Ruysen H, Shabani J, Hanson C, Day LT, Pembe AB, Peven K, Rahman QS, Thakur N, Sharma K, Tahsina T, et al. Uterotonics for prevention of postpartum haemorrhage: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03420-x.
59. KC A, Peven K, Ameen S, Msemo G, Basnet O, Ruysen H, Zaman SB, Mkony M, Sunny AK, Rahman QS, et al. Neonatal resuscitation: EN-BIRTH multi-country validation study. BMC Pregnancy and Childbirth. 2021. https://doi.org/10.1186/s12884-020-03422-9.
60. Tahsina T, Hossain AT, Ruysen H, Rahman AE, Day LT, Peven K, Rahman QS, Khan J, Shabani J, Kc A, et al. Immediate newborn care and breastfeeding: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03421-w.
61. Zaman SB, Siddique AB, Ruysen H, KC A, Peven K, Ameen S, Thakur N, Rahman QS, Salim N, Gurung R, et al. Chlorhexidine for facility-based umbilical cord care: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03338-4.
62. Peven K, Day LT, Ruysen H, Tahsina T, KC A, Shabani J, Kong S, Ameen S, Basnet O, Haider R, et al. Stillbirths including intrapartum timing: EN-BIRTH multi-country validation study. BMC Pregnancy and Childbirth. 2021. https://doi.org/10.1186/s12884-020-03238-7.
63. World Health Organization: Analysis and use of health facility data - Guidance for RMNCAH programme managers. [https://www.who.int/publications/m/item/analysis-and-use-of-health-facility-data-guidance-for-rmncah-programme-managers]. Accessed 20 Aug 2020.
64. DHIS2: RMNCAH configuration package for data collection and dashboards. [https://www.dhis2.org/who-package-downloads#rmncah]. Accessed 20 Aug 2020.
65. Sæbø JI. Global scaling of health information infrastructures: circulating translations. Oslo: Faculty of Mathematics and Natural Sciences, University of Oslo; 2013. Oslo.

Publisher's Note

Chapter 7 – Objective 5: Birthweight measurement

This manuscript assesses routine birthweight in EN-BIRTH study hospitals: accuracy, gaps and opportunities to measure coverage and quality of care.

The chapter was published in March 2021 in BMC Pregnancy and Childbirth. The manuscript was published under a creative commons license (Creative Commons Attribution 4.0 International License) and no further permissions are needed.

The published manuscript is included in full below and supplementary material referenced in the paper is available at https://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/s12884-020-03355-3#Sec26

7.1 List of Figures

7.2 List of Tables

7.3 Citation

Kong S, Day LT, Zaman SB, et al.
Birthweight: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* 2021; **21**(Suppl 1): 240. https://doi.org/10.1186/s12884-020-03355-3

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	034282	**Title**	Dr
First Name(s)	Louise Tina		
Surname/Family Name	Day		
Thesis Title	Quality of care and quality of data for hospital births – tension or traction?		
Primary Supervisor	Associate Professor Cally Tann		

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?	Birthweight: EN-BIRTH multi-country validation study. MC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):240. https://doi.org/10.1186/s12884-020-03355-3. Kong, S., Day, L. T., Zaman, S. B., Peven, K., Salim, N., Sunny, A. K., Shamba, D., Rahman, Q. S., K, C. A., Ruysen, H., El Arifeen, S., Mee, P., Gladstone, M. E., Blencowe, H., Lawn, J. E. and EN-BIRTH Study Group		
When was the work published?	March 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 jointly designed the objectives with my co-first author. I co-designed data cleaning and analysis plan working with my co-first author colleague who ran the statistical analysis under my leadership. I jointly drafted the manuscript whilst mentoring co-first author. I led on results, discussion and conclusion. I am the corresponding author, lead on the manuscript revision in response to peer-review, including co-ordinating collaborative inputs from co-authors.

SECTION E

[redacted]

Student Signature	
Date	

Supervisor Signature	
Date	

Kong et al. BMC Pregnancy and Childbirth 2021, 21(Suppl 1):240
https://doi.org/10.1186/s12884-020-03355-3

BMC Pregnancy and Childbirth

From *Every Newborn* BIRTH multi-country validation study: informing measurement of coverage and quality of maternal and newborn care

RESEARCH

Open Access

Birthweight: EN-BIRTH multi-country validation study

Stefanie Kong[1†], Louise T. Day[1*†], Sojib Bin Zaman[2], Kimberly Peven[1,3], Nahya Salim[4,5], Avinash K. Sunny[6], Donat Shamba[5], Qazi Sadeq-ur Rahman[2], Ashish K.C.[7], Harriet Ruysen[1], Shams El Arifeen[2], Paul Mee[8], Miriam E. Gladstone[1], Hannah Blencowe[1†], Joy E. Lawn[1†] and EN-BIRTH Study Group

Abstract

Background: Accurate birthweight is critical to inform clinical care at the individual level and tracking progress towards national/global targets at the population level. Low birthweight (LBW) < 2500 g affects over 20.5 million newborns annually. However, data are lacking and may be affected by heaping. This paper evaluates birthweight measurement within the *Every Newborn* Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study.

Methods: The EN-BIRTH study took place in five hospitals in Bangladesh, Nepal and Tanzania (2017–2018). Clinical observers collected time-stamped data (gold standard) for weighing at birth. We compared accuracy for two data sources: routine hospital registers and women's report at exit interview survey. We calculated absolute differences and individual-level validation metrics. We analysed birthweight coverage and quality gaps including timing and heaping. Qualitative data explored barriers and enablers for routine register data recording.

Results: Among 23,471 observed births, 98.8% were weighed. Exit interview survey-reported weighing coverage was 94.3% (90.2–97.3%), sensitivity 95.0% (91.3–97.8%). Register-reported coverage was 96.6% (93.2–98.9%), sensitivity 97.1% (94.3–99%). Routine registers were complete (> 98% for four hospitals) and legible > 99.9%. Weighing of stillbirths varied by hospital, ranging from 12.5–89.0%. Observed LBW rate was 15.6%; survey-reported rate 14.3% (8.9–20.9%), sensitivity 82.9% (75.1–89.4%), specificity 96.1% (93.5–98.5%); register-recorded rate 14.9%, sensitivity 90.8% (85.9–94.8%), specificity 98.5% (98–99.0%). In surveys, "don't know" responses for birthweight measured were 4.7%, and 2.9% for knowing the actual weight. 95.9% of observed babies were weighed within 1 h of birth, only 14.7% with a digital scale. Weight heaping indices were around two-fold lower using digital scales compared to analogue. Observed heaping was almost 5% higher for births during the night than day. Survey-report further increased observed birthweight heaping, especially for LBW babies. Enablers to register birthweight measurement in qualitative interviews included digital scale availability and adequate staffing.

(Continued on next page)

* Correspondence: Louise-Tina.Day@lshtm.ac.uk
†Stefanie Kong and Louise T Day are joint first authors.
†Hannah Blencowe and Joy E Lawn are joint senior authors.
[1]Centre for Maternal, Adolescent, Reproductive & Child Health (MARCH), London School of Hygiene & Tropical Medicine (LSHTM), London, UK
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusions: Hospital registers captured birthweight and LBW prevalence more accurately than women's survey report. Even in large hospitals, digital scales were not always available and stillborn babies not always weighed. Birthweight data are being captured in hospitals and investment is required to further improve data quality, researching of data flow in routine systems and use of data at every level.

Keywords: Birth, Newborn, Maternal, Stillbirth, Coverage, Validity, Survey, Health management information systems, Birthweight, Low birthweight

Key findings

What is known and what is new about this study?

- An estimated 20.5 million low birthweight (LBW) babies are born each year, and tracking progress in the highest burden countries still relies on population-based surveys, which are known to have missing data and substantial heaping (preference for recording weights ending in 00). Improving birthweight data in both routine systems and surveys is essential.
- EN-BIRTH is the largest multi-country, multi-site study (> 23,000 births) to assess availability, validity and quality of birthweight data in both survey and routine registers. Qualitative data explored barriers and enablers for routine register recording of birthweight.

Survey–what did we find and what does it mean?

- Survey-reported birthweight coverage underestimated observed coverage by nearly 5% and LBW prevalence by 1%.
- Survey-reported birthweight heaping was 1.5 times higher than the observed heaping.
- Women with stillborn babies reported a much lower coverage of weighing than observed.

Register–what did we find and what does it mean?

- Routine hospital registers were highly complete (> 96%) and legible (> 99%).
- Register-recorded birthweight coverage underestimated observed by 2.2%.
- LBW prevalence underestimated observed by only 0.7%.
- Register-reported birthweight heaping at 2500 g further increased observed heaping by 1.4% for digital scales and 1.1% for analogue.

Gap analysis for quality of care

- Nearly all (95.9%) babies were weighed within 1 h, however, only 14.7% were weighed on digital scales. Stillbirths were weighed much less often, despite birthweight data being fundamental to classifying and intervening to prevent stillbirth.
- Substantial heaping of observed birthweights included those at 2500 g, so the LBW rate will likely be inaccurate.
- Birthweight heaping indices were approximately two-fold lower using digital compared to analogue scales and also 3–5% lower during day shifts compared to night shifts.

What next and research gaps?

- Routine register-records outperformed exit-survey report accuracy for measurement of birthweight and LBW in these hospitals. Further research is needed to assess if survey-reported accuracy decreases over time.
- Investment is needed to explore how digital scales, standardised register process and design can improve birthweight routine data measurement quality further.
- Improving data flow of currently available hospital birthweight data into Health Management Information Systems (HMIS) has potential to close the large LBW data gap in high-burden LMIC settings.

Background

Birthweight closely correlates with newborn survival and lifelong health. The World Health Organization (WHO) recommends measuring birthweight within the first hour of life, ideally using calibrated digital scales with 10 gramme (g) precision [1]. Low birthweight rate has agreed global targets and data are needed to track progress [2]. Among neonatal deaths, 80% have low birthweight (LBW) defined as < 2500 g [3, 4]. An estimated 20.5 million LBW neonates were born in 2015; 91% were born in low- and middle-income countries (LMICs), with almost half in south Asia (48%) and a quarter in sub-Saharan Africa (24%) [3, 5]. LBW survivors continue to have a higher risk of morbidity, including stunting, lower intelligence quotient, and cardiovascular disease later in life [6–8]. Stillborn babies, estimated at > 2 million per year and 84% in LMICs, have similar contributing factors to placental failure as LBW livebirths, yet are not visible as standard birthweight indicator definitions use a livebirth denominator [9].

Tracking coverage of birthweight measurement is recommended and LBW rate is one of only four newborn health measures in WHO's 100 core health indicators [10]. Global nutrition targets set by WHO include a 30% reduction of LBW infants from 2012 to 2025 [2], but the required annual rate of reduction is currently off target [11]. Birthweight data are essential to reach the target neonatal mortality rate (NMR) of Sustainable Development Goal (SDG) 3.2 by 2030 [12]. NMR and stillbirth rates stratified by birthweight group need to be used for perinatal death surveillance and response in settings where accurate gestational age and cause of death assessment is not possible [13]. At an individual level, birthweight data ensures that at-risk newborns receive the immediate care they need and serves as the first measurement for monitoring a child's growth to promote health outcomes throughout the life-course.

Birthweight data are not available for almost one-third (39.7 million) of newborns – the majority in LMICs [3]. Available birthweight data in high mortality burden countries are mostly from population-based surveys, notably the Demographic and Health Surveys (DHS) Program and the United Nations Children's Fund (UNICEF) Multiple

Indicator Cluster Surveys (MICS) [14, 15]. As > 80% of births globally are now in facilities [15], potentially more birthweight data can be made available through routine Health Management Information Systems (HMIS) [4, 14]. When birthweight data are available, concerns about quality, including heaping, limit use and usefulness. Previous birthweight-related indicator validation studies in LMICs have mostly focused on household survey measurement [16–19], with few addressing routine facility measurement [20]. The validity of birthweight measurement through routine hospital registers in LMIC has not previously been studied. The barriers and enablers that affect the quality of birthweight data in routine hospital registers in LMIC are not known.

The *Every Newborn* Action Plan, agreed by all United Nations member states and > 80 development partners, includes an ambitious measurement improvement roadmap [12, 21] with urgent focus to improve data for use towards high-quality care around the time of birth [12, 22]. As part of this roadmap, the *Every Newborn* – Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study aimed to validate the measurement of selected newborn and maternal indicators for routine tracking of coverage and quality of facility-based care [23, 24].

Objectives

This paper is part of a supplement based on the EN-BIRTH multi-country validation study, '*Informing measurement of coverage and quality of maternal and newborn care*', and focuses on birthweight with three objectives:

1. **Determine accuracy/validity of NUMERATOR** for survey-reported and register-recorded birthweight indicator measurement compared to direct observation.
2. **Analyse GAPs in coverage and quality of birthweight measurement**: timeliness, scale choice, proportion of implausible values and heaping/rounding inaccuracy.
3. **Identify BARRIERS and ENABLERS** for routine register recording of birthweight by evaluating register design, filling and use.

Methods

The EN-BIRTH study was a mixed-methods observational study and detailed information regarding the EN-BIRTH research protocol and overall validation results have been published separately [23, 24]. This is the first analysis of the EN-BIRTH birthweight data. A study on birthweight measurement processes and perceived value is published elsewhere in the supplement [25]. Data were collected between June 2017 and July 2018 in five public comprehensive emergency obstetric and newborn care (CEmONC) hospitals in three high burden countries: Maternal and Child Health Training Institute (MCHTI), Azimpur and Kushtia District Hospital in Bangladesh (BD); Pokhara Academy of Health Sciences in Nepal (NP); Temeke Regional Hospital and Muhimbili National Referral Hospital in Tanzania (TZ) (Additional files 1 and 2). Results are reported in accordance with STROBE Statement checklists for observational studies (Additional file 3).

Study participants were consenting women recruited on admission to labour and delivery ward and their newborn babies. We use the term "newborn" in this paper to cover both live births and stillbirths (total births). Exclusion criteria at admission were imminent birth and no fetal heartbeat heard. Trained research clinical observers collected the birthweight from the weighing scale (external gold standard) as the health worker weighed the newborn. Data were time-stamped when documenting birthweight in grammes and type of weighing scale (digital or analogue). Separate groups of data extractors captured birthweight data from existing routine labour ward registers and women's responses to exit-survey after discharge. Data were captured using a custom-built android tablet-based application [26] (Additional file 5).

Implausible observed birthweights (< 350 g or > 6000 g) were excluded from all analyses. Calculations were done for each hospital then combined using a random effects meta-analysis approach. We used 95% confidence intervals to indicate uncertainty when applying our results to a different population. We calculated I^2 and τ^2 to assess heterogeneity between hospitals. Results were stratified by mode of birth (vaginal/caesarean), birth outcome (live birth/stillbirth), and type (single/multiple (twins or higher)) and association determined using chi-squared test.

Analyses were undertaken using Stata version 16 [27] and R statistical programming version 3.5.0 used for graphs [28].

Assessing biases in the data

To determine the reliability of our gold standard, we calculated Cohen's Kappa coefficient for 5% of the sample observed by both supervisors and data collectors [23]. To assess any change in routine register recording practices due to study observer presence, we compared absolute differences between completeness of register extracted study data with one-year pre-study register data collected retrospectively [29]. We also calculated Kappa coefficients for a 5% sample of double-extracted study register data.

Objective 1: Determine numerator for indicator measurement accuracy/validity

We evaluated measurement of two aspects of birthweight data:

a) ***Birthweight coverage*** defined as the number of facility total births (live births and stillbirths) that were weighed, among total births, expressed as a percentage.
b) ***LBW prevalence*** defined as the number of facility total births (live births and stillbirths) whose birthweights were < 2500 g, among total births weighed, expressed as a percentage.

To assess data accuracy, we compared both routine register-recorded coverage and exit interview survey-reported coverage with the gold standard, observed coverage (Fig. 1). Population-based surveys (e.g. DHS and MICS) typically measure coverage from "yes" responses and combine "don't know" with "no" responses as "no coverage." Thus, we analysed survey-reported coverage in this way and also with "don't know" excluded to evaluate effect on accuracy. We interpreted register "not recorded" to mean the baby had not been weighed. LBW classification was calculated using available numeric birthweight data from all three sources.

We calculated absolute differences between observed, register-recorded and exit survey-reported coverage. Cut-off ranges were adapted from WHO's Data Quality Review (DQR) methods (over/underestimate by 0–5%,6–10%, 11–15%, 16–20% and > 20%) [30, 31].

To understand how coverage measurement affected low and normal birthweight categorisation, we calculated "validity ratios". Similar to verification ratios in DQR methods [30], a ratio higher than 1.0 implies overestimation of survey-reported or register-recorded coverage compared to observed, and a ratio lower than 1.0 implies an underestimate. Cut-off ranges adapted from DQR methods were used for heat-maps [30].

Individual-level validity "diagnostic test" methods were calculated using two-way tables. When column totals were ≥ 10, we calculated sensitivity, specificity, negative predictive value, positive predictive value, area under the curve and inflation factor; otherwise we present percent agreement [23, 32]. Individual-level agreement was assessed using Bland-Altman plots [33].

Objective 2: Gaps in coverage and quality of birthweight measurement

We calculated gap analyses for high-quality birthweight among (A) total births as the total eligible population; (B) birthweight coverage; (C) right timeliness of measurement - weighed ≤1 h after birth; (D) right device - digital scales.

Data completeness for registers was assessed. Birthweight heaping and rounding were evaluated for observed, survey-reported and register-recorded data in two ways: First, the proportion of total birthweights that

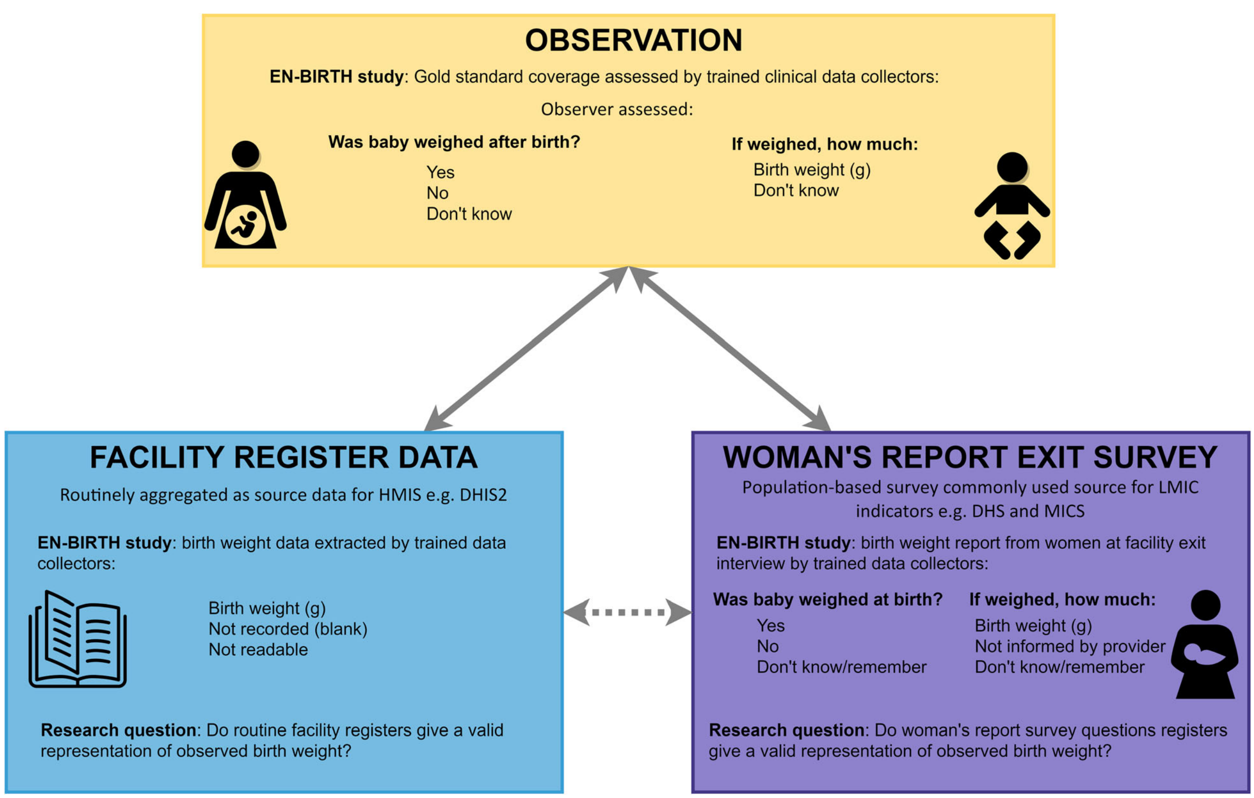


Fig. 1 Birthweight validation design, EN-BIRTH study. Adapted from EN-BIRTH protocol [23]

were multiples of 500 g; second, the proportion of heaped weight values (e.g. 2500 g) relative to all weight values within the adjacent 500 g bracket (e.g. 2250-2750 g). We stratified by type of weighing device and time of birth by midwifery shift time (day/night). To demonstrate the effect of heaping on LBW prevalence in routine register documentation, we adjusted LBW prevalence by reallocating 25% of babies with an exact birthweight of 2500 g to the LBW category and compared with exit-survey findings using the same method [34].

Objective 3: Barriers and enablers to routine register recording

We evaluated barriers and enablers to recording of birthweight in routine registers as part of the wider barriers and enablers objective of the EN-BIRTH study. The structure of the routine labour ward register was correlated with completeness and accuracy of measurement [31].

We designed three tools: a) semi-structured in-depth interview (IDI) guide, b) semi-structured focus group discussion (FGD) guide, c) "care-to-documentation checklist."

Experienced qualitative researchers conducted IDIs with two purposively sampled groups of respondents in each EN-BIRTH study hospital: 1) hospital midwives and doctors involved in birthweight measurement and 2) study data collectors. To triangulate results, FGDs were carried out with health workers. The sample size was determined using saturation sampling. Qualitative data were thematically analysed by categorizing pre-identified codes based on the Performance of Routine Information System Management (PRISM) conceptual framework [35] using NVivo 12 for data management. The care-to-documentation checklist was completed after the IDI and captured details regarding: which health worker cadre weighs the baby; who documents the birthweight; into which documents (patient notes, registers, partograph, etc.); the typical order of documentation; estimation of how long between weighing the baby and documentation. Data were entered into Microsoft Excel and analysed in R version 3.6.1 [28]. This paper specifically presents emerging themes regarding birthweight recording across three topics: 1) Register design 2) Register filling and 3) Register use. Detailed methods and results of all emerging themes for register recording of all EN-BIRTH selected indicators are available in an associated paper [36].

Results

Among the total 23,471 births observed, 22,617 (96.3%) newborns were weighed after birth and implausible weights were 0.01% (Additional file 4). Exit-survey interviews were completed by 88.4% of their mothers and register data were extracted for 95.3% (Fig. 2).

Background characteristics are shown in Table 1. 12.1% of mothers were adolescents < 20 years and almost half of women (48.4%) had completed secondary education. Live births were 97.3% and twins/triplets 3.9%. The proportion of babies delivered by caesarean section varied widely, from 7.2% in Temeke TZ to 73.2% in Azimpur BD. Hospital register design in Bangladesh was updated during the study as part of a national standardisation – we present revised register results in the multi-site tables and figures and report the effect of this natural experiment in Additional file 6.

Inter-rater reliability was very high for both observation and data extraction (Additional file 7). Routine register completeness comparison before and during study showed decrease in completeness by < 1.5%, except in Kushtia BD, which increased from 66.1% to 85.2% (Additional file 8).

Coverage data by observation, survey-report and register-record are shown in Fig. 3. Coverage comparisons and individual-level metrics are shown in Tables 2 and 3. Any association with delivery mode, multiple births, and stillbirth are shown in Additional files 9, 10 and 11.

Objective 1: Numerator validation

Birthweight coverage

Survey-reported coverage, 94.3% (90.2–97.3%), underestimated the observed coverage of 98.8%. Exit-survey heterogeneity was low, $\tau^2 = 0.03$ (Additional file 12). "Don't know" responses were 4.5% (2.1–8.4%) and pooled individual-level validation results were mixed: sensitivity 95.0% (91.3–97.8%), specificity 43.3%(15.1–74.0%). There was no evidence of a difference in survey-reported coverage by delivery mode or single/multiple pregnancy. Across the sites, stillbirth observed birthweight coverage ranged from 12.5–98.3%, and survey-report underestimated by 8.2–46.6% (Additional file 10).

Register-recorded coverage of 96.6% (93.2–98.9%) underestimated the observed coverage of 98.8%. Heterogeneity was low, $\tau^2 = 0.03$ (Additional file 12). In Temeke TZ, coverage was overestimated by 0.1% and in the other four hospitals underestimated by 0.3–12.1%. Sensitivity was > 88% and specificity ranged from 3.5% in Muhimbili TZ to 82.0% in Kushtia BD. Register-recorded coverage was significantly higher among babies born from vaginal deliveries compared to caesarean section, as well as live births compared to stillbirths (Additional files 10 and 11).

Low birthweight (LBW) prevalence

Observed LBW prevalence overall was 15.6%, lowest in Temeke TZ 7.6% and highest in Muhimbili TZ 28.1%. Survey-reported LBW coverage, 14.3 (8.9–20.9%), underestimated observed coverage of 15.6%. "Don't know"

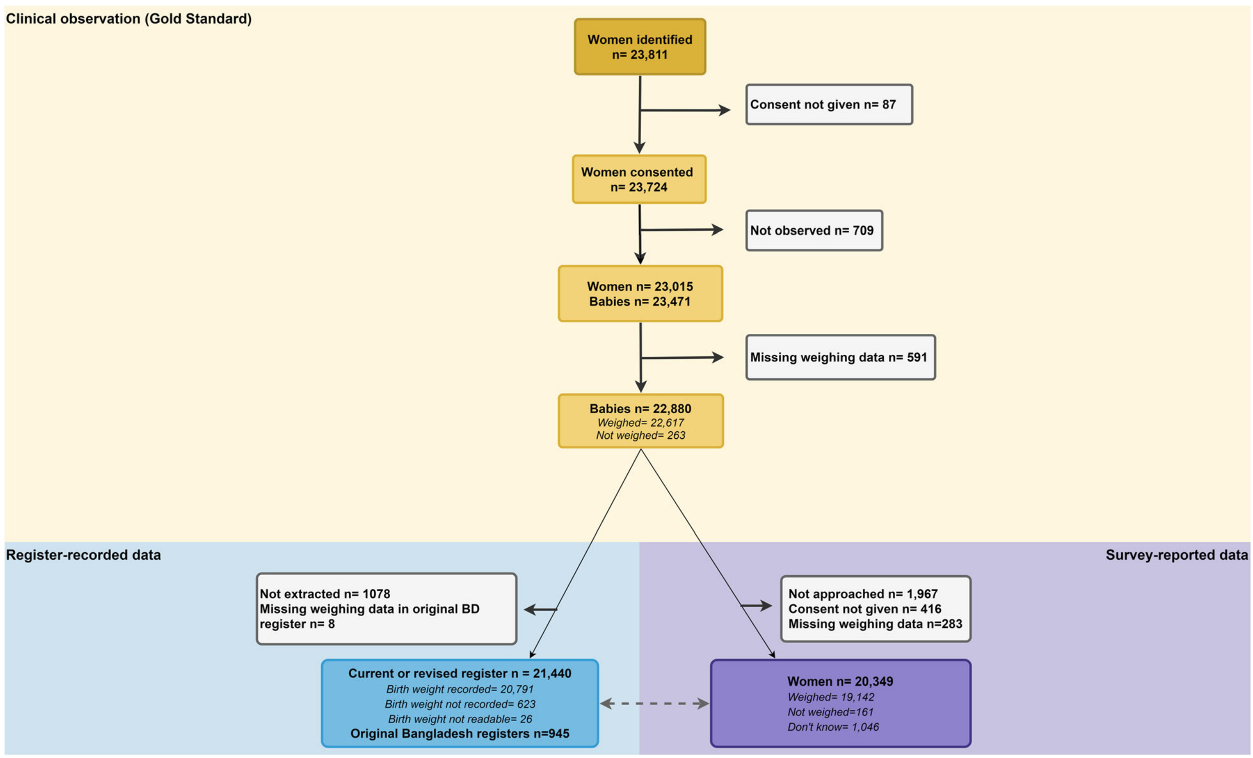


Fig. 2 Flow diagram for birthweight cases, EN-BIRTH study (n=23,471). Adapted from EN-BIRTH protocol [23]

survey responses were 2.9% (1.8–4.3%). Sensitivity was 82.9% (75.1–89.4%) and specificity 96.4% (93.5–98.5%). LBW observed among stillborn babies ranged widely from 0.0–75.5%, both over- and underestimated by survey-report in different sites.

Register-recorded LBW coverage of 14.9% (8.8–22.3%) underestimated observed coverage, 15.6%. Register sensitivity was 90.8% (85.9–94.8%) and specificity 98.5% (98.0–99.0%). Both survey-reported and register-recorded LBW coverage were higher among caesarean sections, stillbirths, and twins/triplets.

Survey-reported validity ratios for LBW babies were poor to good (0.78–1.62) and very good to excellent (0.91–1.08) for normal birthweight (Fig. 4). Register-recorded validity ratios were excellent (0.99–1.03) for both LBW and normal birthweight newborns.

Bland-Altman plots showed agreement between observed and survey-reported birthweight was reasonable, with mean difference = 6.3 g (2.7, 9.9), and for register-recorded was high, with mean difference = – 1.39 g (– 4.4, 1.6) (Additional file 13).

Objective 2: Gaps in coverage and quality of birthweight measurement

Figure 5 shows gap analyses linked to coverage measurement. Almost all newborns (95.9%) were observed to be weighed within the right time (C), 1 h of birth. Digital scales as the right device (D) were used in only three of the hospitals: Azimpur BD (74.2%), Muhimbili TZ (29.3%) and rarely in Temeke TZ (2.0%) (Additional file 14).

Register-recorded birthweight was legible (Fig. 6). Completeness was very high (> 98%) in all hospitals, except in Kushtia BD (85.5%). Completeness was higher in Bangladesh revised registers compared to the original: Azimpur BD = 98.4% from 57.4% and Kushtia BD = 85.2% from 43.8% (Additional file 6).

Birthweight heaping and rounding

Observer-assessed birthweight heaping was two-fold lower by digital (15.7%) compared to analogue scales (36%). Survey-report further increased heaping (digital 25.3%, analogue 43.4%). Register-record increased heaping by only 1.4% for digital scales and 1.1% for analogue (Table 4). Heaping indices were consistently lower for digital than analogue scales across all 500 g increments (Table 5), and higher during night than day shifts (Table 4). Re-allocation of 25% of 2500 g birthweights to the LBW category increased LBW prevalence by 2.0% for register-record and 2.5% for survey-report (Additional file 15).

Table 1 Characteristics of babies and women observed in labour and delivery wards, EN-BIRTH study ($n = 23,471$ births)

	Bangladesh		Nepal	Tanzania		All sites
	Azimpur Tertiary n (%)	Kushtia District n (%)	Pokhara Regional n (%)	Temeke Regional n (%)	Muhimbili National n (%)	n (%)
Total Women	2910	2412	7370	6748	3575	23,015
Women's Age						
< 18 years	25 (0.9)	3 (0.1)	311 (4.2)	26 (0.4)	8 (0.2)	373 (1.6)
18–19 years	475 (16.3)	197 (8.2)	817 (11.1)	767 (11.4)	159 (4.4)	2415 (10.5)
20–24 years	1158 (39.8)	954 (39.6)	3080 (41.8)	2314 (34.3)	722 (20.2)	8228 (35.8)
25–29 years	867 (29.8)	736 (30.5)	2114 (28.7)	1697 (25.1)	1134 (31.7)	6548 (28.5)
30–34 years	297 (10.2)	373 (15.5)	827 (11.2)	1146 (17)	924 (25.8)	3567 (15.5)
35+ years	88 (3)	149 (6.2)	221 (3)	798 (11.8)	628 (17.6)	1884 (8.2)
Maternal education						
No Education	39 (1.3)	77 (3.2)	268 (3.6)	202 (3)	66 (1.8)	652 (2.8)
Primary incomplete	111 (3.8)	127 (5.3)	252 (3.4)	81 (1.2)	45 (1.3)	616 (2.7)
Primary complete	339 (11.6)	347 (14.4)	302 (4.1)	31 (0.5)	5 (0.1)	1024 (4.4)
Secondary incomplete	985 (33.8)	954 (39.6)	1637 (22.2)	4053 (60.1)	1299 (36.3)	8928 (38.8)
Secondary complete or higher	1273 (43.7)	870 (36.1)	4509 (61.2)	2346 (34.8)	2146 (60)	11,144 (48.4)
Missing	163 (5.6)	37 (1.5)	402 (5.5)	35 (0.5)	14 (0.4)	651 (2.8)
Parity						
Nullipara	1350 (46.4)	1038 (43)	4402 (59.7)	2917 (43.2)	1363 (38.1)	11,070 (48.1)
Multipara	56 (1.9)	5 (0.2)	6 (0.1)	13 (0.2)	3 (0.1)	83 (0.4)
Missing	1504 (51.7)	1369 (56.8)	2961 (40.2)	3816 (56.6)	2207 (61.8)	11,857 (51.5)
Total Baby	2936	2459	7442	6869	3765	23,471
Live Birth	2895 (99.5)	2302 (96.6)	7171 (98.1)	6606 (97.3)	3490 (94.5)	22,464 (97.3)
Baby's condition at L&D discharge						
Alive	2895 (99.5)	2302 (96.6)	7171 (98.1)	6606 (97.3)	3490 (94.5)	22,464 (97.3)
Stillbirth	11 (0.3)	74 (3)	126 (1.7)	153 (2.2)	186 (3)	550 (2.2)
Neonatal death	1 (0)	6 (0.3)	4 (0.1)	28 (0.4)	19 (0.5)	58 (0.3)
Missing	2 (0.1)	2 (0.1)	6 (0.1)	5 (0.1)	0 (0)	15 (0.1)
Baby number						
Single	2864 (98.3)	2296 (96.1)	7185 (98)	6561 (96.4)	3336 (90)	22,242 (96.1)
Twin	48 (1.6)	86 (3.6)	140 (1.9)	242 (3.6)	336 (9.1)	852 (3.7)
Triplets	3 (0.1)	6 (0.3)	3 (0)	0 (0)	33 (0.9)	45 (0.2)
Mode of birth						
Normal vertex delivery	784 (26.7)	1453 (59.1)	5889 (79.1)	6307 (91.8)	1616 (42.9)	16,049 (68.4)
Vaginal breech/ Vacuum/ Forceps	1 (0)	0 (0)	351 (4.7)	10 (0.1)	10 (0.3)	372 (1.6)
Caesarean section	2142 (73)	996 (40.5)	1163 (15.6)	489 (7.1)	2105 (55.9)	6895 (29.4)
Missing	9 (0.3)	10 (0.4)	39 (0.5)	63 (0.9)	34 (0.9)	155 (0.7)
Birthweight						
Extremely LBW < 1000 g	1 (0)	7 (0.3)	27 (0.4)	13 (0.2)	44 (1.2)	92 (0.4)
Very LBW 1000-1499 g	1 (0)	27 (1.2)	38 (0.5)	22 (0.3)	159 (4.5)	247 (1.1)
LBW 1500-2499 g	351 (12.2)	437 (19.1)	830 (11.4)	466 (7.1)	794 (22.2)	2878 (12.7)
All LBW < 2500 g (sum of above)	353 (12.2)	471 (20.6)	895 (12.3)	501 (7.6)	997 (27.9)	3217 (14.2)
Not LBW ≥2500 g	2528 (87.5)	1804 (78.9)	6274 (86.5)	6051 (91.7)	2549 (71.4)	19,206 (85)

Table 1 Characteristics of babies and women observed in labour and delivery wards, EN-BIRTH study ($n = 23,471$ births) *(Continued)*

	Bangladesh		Nepal	Tanzania		All sites
	Azimpur	Kushtia	Pokhara	Temeke	Muhimbili	
	Tertiary	District	Regional	Regional	National	
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Missing	7 (0.2)	11 (0.5)	83 (1.1)	46 (0.7)	24 (0.7)	171 (0.8)
Sex						
Male	1465 (50.4)	1220 (51.3)	3903 (53.6)	3481 (51.5)	1833 (50.2)	11,902 (51.8)
Female	1441 (49.6)	1154 (48.5)	3369 (46.2)	3265 (48.3)	1813 (49.6)	11,042 (48.1)
Ambiguous	1 (0)	4 (0.2)	13 (0.2)	7 (0.1)	6 (0.2)	31 (0.1)

Objective 3: Barriers and enablers to routine recording

All study hospital labour ward registers had a specific column to record birthweight, usually recorded in kilogrammes to 1 decimal place, despite the Bangladesh revised register column heading specifying the unit in grammes (Fig. 6).

IDIs were conducted with 40 nurse-midwives/doctors and 65 EN-BIRTH study data collectors and one FGD was conducted in each hospital ($n = 5$). Emerging themes functioning as both barriers or enablers in the five hospitals are shown in Fig. 7.

Register design

All respondents stated the labour ward register was adequately designed for birthweight measurement. Complexity of documentation systems was expressed by respondents as a barrier, since birthweight is also written in several other formal and informal documents. The order of birthweight documentation was first into the register in Bangladesh, while in Nepal and Tanzania birthweights were recorded in one to three other documents before the register (Additional file 16).

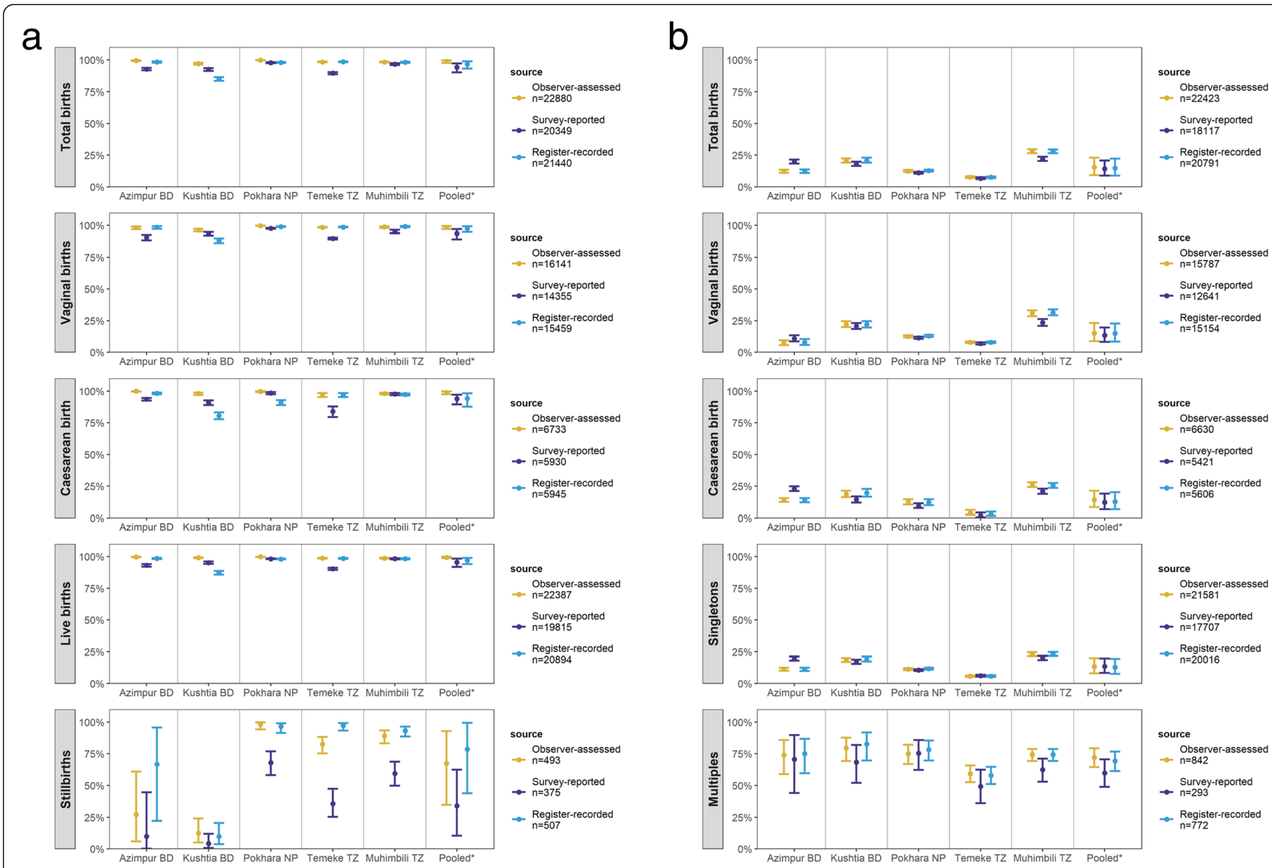


Fig. 3 **a** Coverage rates for babies weighed at birth and **b** prevalence of low birthweight newborns measured by observation, exit-survey and register, EN-BIRTH study. *Random effects. **a** n = 22,880 births, **b** n = 22,423 births. *BD* Bangladesh, *NP* Nepal, *TZ* Tanzania

Table 2 Individual-level validation in surveys and registers for weighing coverage, EN-BIRTH study (n = 23,471 births)

	Bangladesh				Nepal		Tanzania				All sites pooled (Random effects)	
	Azimpur		Kushtia		Pokhara		TZ - Temeke		TZ - Muhimbili			
	Tertiary		District		Regional		Regional		National			
Baby weighed - Survey reported		95% CI		95% CI		95% CI		95% CI		95% CI		95% CI
Observer coverage (%)	99.5	(99.1, 99.7)	97.1	(96.3, 97.7)	99.8	(99.7, 99.9)	98.4	(98.1, 98.7)	98.4	(97.9, 98.8)	98.8	(97.7, 99.6)
Survey reported coverage (%)	92.8	(91.8, 93.7)	92.5	(91.3, 93.5)	97.8	(97.4, 98.1)	89.6	(88.7, 90.4)	96.7	(96, 97.3)	94.3	(90.2, 97.3)
"Don't know" responses (%)	6.8	(5.9, 7.7)	5.4	(4.5, 6.3)	2.0	(1.7, 2.4)	9.5	(8.7, 10.3)	2.2	(1.7, 2.8)	4.7	(2.1, 8.4)
Sensitivity (%)	93.1	(92.1, 94)	95.4	(94.4, 96.2)	97.9	(97.5, 98.2)	89.8	(88.9, 90.5)	97.1	(96.4, 97.7)	95.0	(91.3, 97.8)
Specificity (%)	57.1	(28.9, 82.3)	84.6	(73.5, 92.4)	25.0	(5.5, 57.2)	27.3	(17, 39.6)	20.9	(10.0, 36.0)	43.3	(15.1, 74.0)
Percent agreement (%)	92.9	(91.9, 93.8)	95.1	(94.1, 95.9)	97.8	(97.4, 98.1)	89.0	(88.2, 89.8)	95.8	(95, 96.6)	91.8	(88.4, 94.7)
Baby weighed - Register recorded												
Observer coverage (%)	99.5	(99.1, 99.7)	97.1	(96.3, 97.7)	99.8	(99.7, 99.9)	98.4	(98.1, 98.7)	98.4	(97.9, 98.8)	98.8	(97.7, 99.6)
Register recorded coverage (%)	98.4	(97.8, 98.9)	85.0	(83.4, 86.5)	98.0	(97.7, 98.4)	98.5	(98.2, 98.8)	98.1	(97.6, 98.5)	96.6	(93.2, 98.9)
Not recorded (%)	1.6	(1.2,2.2)	14.8	(13.3,16.4)	1.9	(1.6,2.2)	1.3	(1.1,1.6)	1.8	(1.4,2.2)	3.2	(1.0, 6.7)
Not readable (%)	–	–	0.2	(0.1,0.6)	0.1	(0,0.2)	0.1	(0.1,0.3)	0.1	(0.1,0.3)	0.1	(0, 0.2)
Sensitivity (%)	*	*	87.7	(86.2, 89.1)	98.2	(97.9, 98.5)	98.8	(98.5, 99.1)	98.4	(97.9, 98.8)	97.1	(94.3, 99.0)
Specificity (%)	*	*	82.0	(68.6, 91.4)	15.4	(1.9, 45.4)	9.3	(4.3, 16.9)	3.5	(0.4, 12.1)	24.1	(0.6, 61.9)
Percent agreement (%)	*	*	87.6	(85.8, 88.7)	98.1	(97.6, 98.3)	97.5	(96.9, 97.7)	96.9	(96.1, 97.3)	95.2	(92.2, 97.5)

*Validity statistics suppressed where < 10 count in either column of two-by-two table
– No observations
Percent agreement was calculated as the sum of true positives and true negatives divided by the total number of newborns: (TP + TN)/n. For survey-reported weighing coverage, we combined "don't know" with "no" answers. Survey validity results with "don't know" responses excluded are presented in Additional file 12. Two-way tables are presented in Additional file 19

Register filling

All respondents stated recording birthweight in labour ward registers is standard practice. Birthweight is usually written down by the same nurse-midwife who weighed the newborn, but only after providing all other care around the time of birth for mother and baby. Estimated time from weighing the newborn to birthweight register documentation averaged 4–31 min, up to a maximum of 1–3 h (Additional file 17). Shortage of time was a frequently measured barrier to high quality register documentation. EN-BIRTH data collectors described seeing that when busy, health workers may record the birthweight on a separate piece of paper, or ask the mother or another colleague to remember the weight, and transfer this weight later into formal documents. The baby may be weighed again if later no one can recall the birthweight.

The enabler of additional actors only available during the day shift was mentioned.

> *"Most of the time documentation was done appropriately because there were students who could offer assistance during the day. But it was very difficult during night shift because the midwife should do everything by herself like getting the birthweight, resuscitation … when it comes to recording she will find that she has forgotten most of the information."*
> -Health worker, Muhimbili TZ

EN-BIRTH study clinical observers commented on the barrier that health workers did not trust the precision of the weighing scales and sometimes used their personal judgement and rounded birthweights:

> *"If [the analogue scale] shows 4 kilo 300 grammes, they assume it [is] 4 kilo, 500 grammes."*
> -Data collector, Azimpur BD

Register use

Health workers acknowledged the importance of birthweight data and described its use for clinical care only:

> *"Information recording is critical and exact [numbers] should be recorded … we take special care on managing babies with low birthweight, high birthweight … [which] can require paediatrics consultation."*
> -Health worker, Pokhara NP

No respondent mentioned birthweight data for use higher up the health system. A barrier to use was expressed in the low level of trust in the birthweight data quality:

Table 3 Individual-level validation in surveys and registers for LBW prevalence, EN-BIRTH study (23,471 births)

	Bangladesh				Nepal		Tanzania				All sites pooled (Random effects)	
	Azimpur		Kushtia		Pokhara		TZ - Temeke		TZ - Muhimbili			
	Tertiary		District		Regional		Regional		National			
Low birthweight - Survey-reported		95% CI		95% CI		95% CI		95% CI		95% CI		95% CI
Observer prevalence (%)	12.3	(11.1, 13.5)	20.7	(19.1, 22.4)	12.5	(11.7, 13.3)	7.6	(7, 8.3)	28.1	(26.6, 29.6)	15.6	(9.3, 23.1)
Survey reported prevalence (%)	19.8	(18.3, 21.5)	18.1	(16.5, 19.8)	11.1	(10.3, 11.8)	6.7	(6, 7.5)	22.0	(20.4, 23.7)	14.3	(8.9, 20.9)
"Birthweight not informed by provider" (%)	0.9	(0.6,1.4)	0.2	(0.1,0.5)	0.0	(0,0.1)	7.3	(6.6,8.1)	0.9	(0.6,1.4)	1.1	(0.0, 4.3)
"Don't know" (%)	4.3	(3.6,5.1)	0.9	(0.6,1.4)	2.7	(2.3,3.1)	4.4	(3.9,5)	3.2	(2.6,4)	2.9	(1.8, 4.3)
Sensitivity (%)	89.0	(84.9, 92.3)	81.0	(76.9, 84.7)	87.4	(84.8, 89.8)	63.3	(56.8, 69.4)	88.8	(85.8, 91.4)	82.9	(75.1, 89.4)
Specificity (%)	89.7	(88.4, 91.0)	97.4	(96.5, 98.1)	98.6	(98.3, 98.9)	96.6	(96.0, 97.1)	97.5	(96.7, 98.2)	96.4	(93.5, 98.5)
Percent agreement (%)	85.0	(83.5, 86.3)	93.1	(92, 94.2)	94.7	(94.2, 95.3)	83.7	(82.6, 84.7)	91.8	(90.7, 92.8)	81.5	(74.3, 87.8)
Low birthweight - Register-recorded												
Observer prevalence (%)	12.3	(11.1, 13.5)	20.7	(19.1, 22.4)	12.5	(11.7, 13.3)	7.6	(7, 8.3)	28.1	(26.6, 29.6)	15.6	(13.9, 14.8)
Register recorded prevalence (%)	12.3	(11, 13.8)	21.1	(19.2, 23)	12.8	(12, 13.6)	7.5	(6.9, 8.2)	28.1	(26.6, 29.6)	14.9	(8.8, 22.3)
Sensitivity (%)	93.3	(89.6, 96.0)	88.9	(85.2, 91.9)	94.0	(92.2, 95.5)	81.2	(77.4, 84.6	94.2	(92.5, 95.6)	90.8	(85.9, 94.8)
Specificity (%)	99.2	(98.6, 99.5)	97.3	(96.3, 98.1)	99.0	(98.7, 99.2)	98.5	(98.1, 98.8)	98.2	(97.6, 98.6)	98.5	(98.0, 99.0)
Percent agreement (%)	98.3	(96.2, 97.7)	87.6	(82, 85.3)	98.1	(96.1, 96.9)	97.5	(95.4, 96.4)	96.9	(94.6, 96.1)	91.8	(87.6, 95.1)

Don't know % = proportion of women who answered "Don't Know" when asked the weight of their child

"Some nurses do not record the details after they have helped a mother to deliver … if [documents are] not fully filled so people start to estimate, so this leads to non - accurate data about the weight of a child … we sometimes fill not actual data."
-Health worker, Temeke TZ

Discussion

Birthweight measurement in our five CEmONC study hospitals was almost universal and routine facility registers measured coverage of weighing at birth and LBW classification more accurately than exit interview surveys. These findings align with our qualitative study in one EN-BIRTH hospital, Temeke TZ, which reported birthweight is highly valued by both health workers and mothers [25].

Routine registers' high completeness and accuracy for birthweight across all five hospitals was especially notable. Importantly, we found register records for LBW babies had both high sensitivity and specificity > 90%, which was even higher than a study from Nigeria that reported sensitivity 62% and specificity 85% [20]. Birthweight coverage for babies of any birthweight (LBW and not LBW) similarly had high overall sensitivity of 97.1%; however, specificity was very low (4–15%) in three hospitals. We postulate this might be due to the baby being weighed and register documented

		Bangladesh		Nepal	Tanzania		All sites pooled (Random effects)
		Azimpur	Kushtia	Pokhara	Temeke	Muhimbili	
	Prevalence	Tertiary	District	Regional	Regional	National	
Ratio: Survey: Observed	Low birthweight	1.62	0.87	0.89	0.88	0.78	0.97
	Normal birthweight	0.91	1.03	1.02	1.01	1.08	1.01
Ratio: Register: Observed	Low birthweight	1.01	1.02	1.03	0.99	1	1.01
	Normal birthweight	1	1	1	1	1	1

< 0.80	OR	>1.20	Poor
0.80 to 0.84	OR	1.16 to 1.20	Moderate
0.85 to 0.89	OR	1.11 to 1.15	Good
0.90 to 0.94	OR	1.06 to 1.10	Very Good
0.95 to 0.99	OR	1.00 to 1.05	Excellent

Fig. 4 Validity ratios for survey-reported and register-recorded low/normal birthweight prevalence compared to observation, EN-BIRTH study. Heat-mapped using WHO's Data Quality Review (DQR) 5%, 10%, 15% and 20% cutoffs [30]

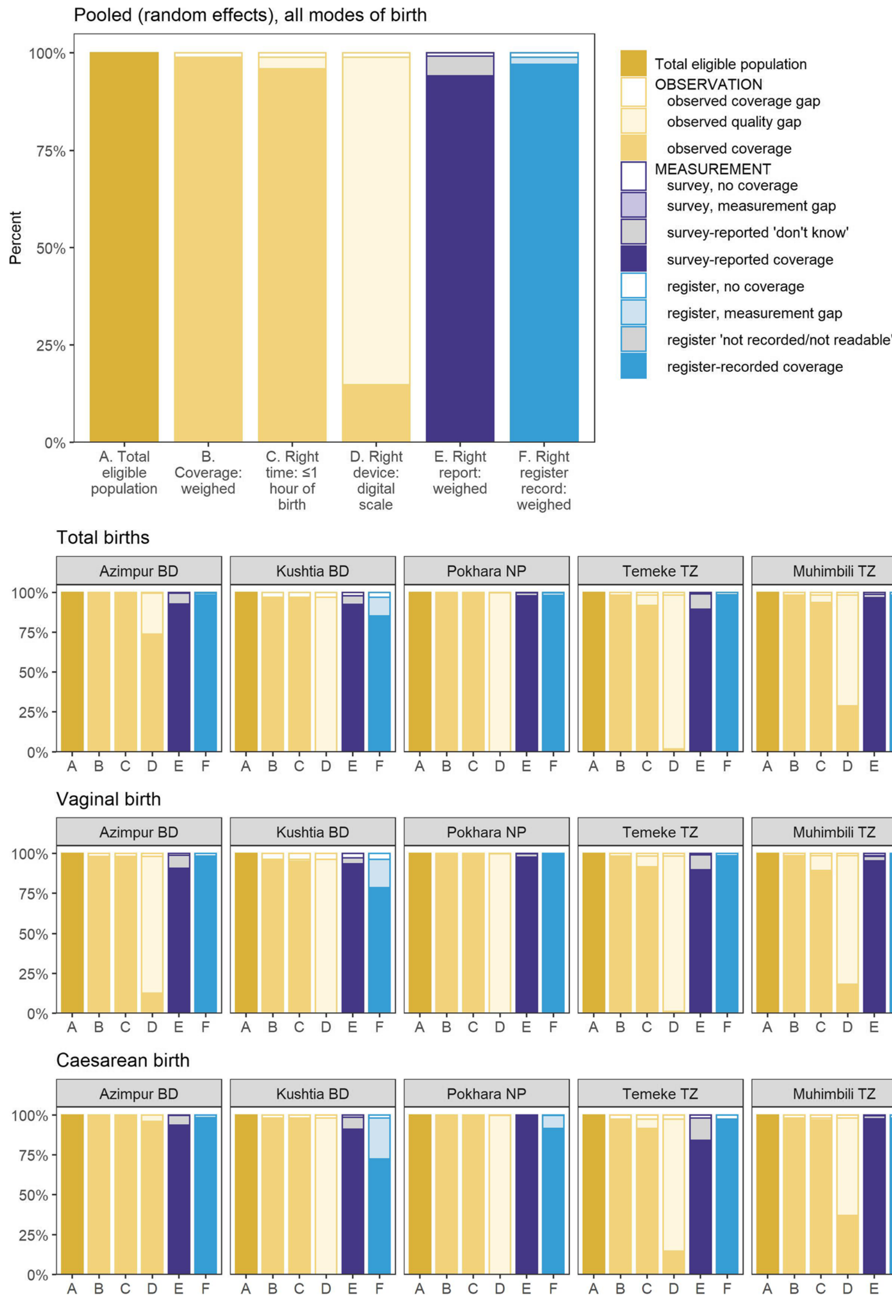


Fig. 5 Gap analysis for coverage and quality of weighing practice at birth, EN-BIRTH study (n = 23,471). Stratified by vaginal and caesarean births in EN-BIRTH study (observer assessed n = 23,471, survey reported n = 20,349, and register recorded n = 21,440). *BD* Bangladesh, *NP* Nepal, *TZ* Tanzania

after observation had ceased (higher false positives). The exception was Kushtia BD's higher specificity of 82.0%, which may relate to the lower register completeness overall (85.2%) (higher true negatives). Register birthweight for LBW babies outperforming all birthweight babies may reflect the extra care given by health workers to the more vulnerable babies – for example, weighing more quickly after birth and thus being captured by the EN-BIRTH observers.

Survey-reported birthweight at the point of hospital discharge soon after birth was also accurate compared to observation. Our results align with a systematic review of 40 studies that showed high agreement between survey-recalled and register-recorded birthweights as the standard [37]. For weighing coverage, survey-report compared to observation had high sensitivity but lower specificity. Similar to registers, this could be due to mothers' correct report of baby weighing after observation stopped. Survey-

Figure 4. Routine register design and data quality dimensions for birthweight by site, EN-BIRTH study

		Bangladesh		Nepal	Tanzania	
		Azimpur	**Kushtia**	**Pokhara**	**Temeke**	**Muhimbili**
		Tertiary	**District**	**Regional**	**Regional**	**National**
Register design	Column allotted data element	specific column	specific column	specific column	specific column	specific column
	Column heading	Weight (grammes)	Weight (grammes)	Weight	Weight of the child in gram/Kg	BWT
	Data format if birthweight recorded	#.# kg	#.# kg	# kg #.# kg #.## kg	# #.#	#.# #.# kg #.## kg ### g #.### ####
	Data format if birth weight not recorded	blank	blank	blank	dash	dash
Completeness	Data element recorded in register	98.4%	85.2%	98.1%	98.7%	98.3%
External Consistency	*Indicator:* Observed coverage	99.5%	97.1%	99.8%	98.4%	98.4%
	Indicator: Measured coverage - register recorded	98.4%	85.0%	98.0%	98.5%	98.1%
	Measurement gap: Register recorded and observed	1.1% underestimate	12.1% underestimate	1.8% underestimate	0.1% overestimate	0.3% underestimate

Key

>20%	Poor
16-20%	Moderate
11-15%	Good
6-10%	Very Good
0-5%	Excellent

Fig. 6 Routine register design and data quality dimensions for birthweight by site, EN-BIRTH study. For basis of ranges, see WHO Data Quality Review [31]

report for LBW babies again outperformed their counterparts, likely for the same reasons of extra care given to LBW babies. This is in contrast to previous studies that revealed mixed but generally low accuracy for LBW prevalence, ranging from a sensitivity of 45% in a study conducted in Nepal to 71% in Kenya [16, 18, 19, 38]. These validation studies evaluated survey report from soon after birth to household survey 22 months later.

Quality of birthweight measurement was mixed. Whilst liveborn babies had timely birthweights, we found quality gaps for other dimensions, especially the widely recognized heaping on multiples of 500 g [5, 29, 34]. The EN-BIRTH study design permitted exploration of cumulative heaping at different measurement capture points: the birthweight observation, exit interview and register-record. We found heaping increased slightly between observation and register-record despite the reality that usually the same health worker weighs and documents. Notably, heaping doubled when the same data were captured from women's report at exit interview. Obtaining a precise birthweight for all babies is fundamental. For instance, a baby whose true birthweight of 2480 g if rounded to 2500 g would not be correctly identified as LBW and fail to receive appropriate care. The same logic applies to identifying newborns weighing 2000 g or less, for whom kangaroo mother care is recommended.

The stillbirth birthweight gap was a striking finding in all hospitals except Pokhara NP. If gestational age is uncertain, the definition of stillbirth uses birthweight, vital for the minimum dataset for perinatal death surveillance and response to reduce preventable death [39]. As such, we suggest tracking coverage of stillbirth birthweight has potential as an indicator of respectful maternal and newborn care. More in-depth analyses regarding stillbirths in the EN-BIRTH study is reported separately [40].

Digital scale measurement gave lower heaping indices across all weights compared to analogue scales in our study. A 1980s Canadian study had postulated that digit bias was attributed to the use of analogue scales; however, a British study later found that significant rounding and truncation persisted even with digital scales [41, 42]. Few published studies have explored the relationship between type of scale and LBW estimates. We found less heaping at 2500 g using digital scales, implying more babies would have been correctly classified as LBW. One previous study in India also found that the percentage of LBW babies identified by digital scales (29.5%) was higher compared to analogue scales (23%) [43].

In our study, two of five CEmONC hospitals were not, or rarely using, digital scales despite the relative low cost of these devices. This high usage of analogue scales remains a concern because heaping and rounding may be attributed to the instrument's imprecision and/or the health workers' subsequent lack of confidence in the measurement. Increasing the availability of digital scales at hospitals is important; however, some nurses stated their preference to use analogue scales because they were more familiar with these devices [44]. Thus, beyond providing digital scales, training and supportive supervision are required to improve quality of birthweight measurement. Our findings provide additional support to inform health system decisions to invest in digital scales for all facilities providing care at birth and improve accuracy of birthweight, especially LBW measurement.

High-quality care must be consistently provided during both day and night shifts. Our qualitative interview findings of lower availability of health workers under increased time pressure during night shifts lends explanation for poorer quality birthweight measurement

Table 4 Heaping index of observer-assessed, survey-reported, and register-recorded birthweights stratified by time of birth, EN-BIRTH study

		Bangladesh				Nepal		Tanzania				All sites pooled[a]	
		Azimpur		Kushtia		Pokhara		Temeke		Muhimbili			
		Tertiary		District		Regional		Regional		National			
All		Digital	Analogue	Digital	Analogue	Digital	Analogue	Digital	Analogue	Digital	Analogue	Digital	Analogue
Observation	All birthweights within 350-6000 g, n	2140	741	0	2275	0	7169	133	6418	1037	2509	3310	19,112
	All heaped birthweights within 350-6000 g, (%)	(5.7)	(44.0)	(0.0)	(38.3)	(0.0)	(36.8)	(34.6)	(41.9)	(13.6)	(22.2)	(15.7)	(36.0)
Survey	All birthweights within 350-6000 g, n	1858	628	0	2105	0	6363	74	4307	747	1683	2679	15,086
	All heaped birthweights within 350-6000 g, (%)	(13.5)	(51.8)	(0.0)	(45.8)	(0.0)	(39.9)	(47.3)	(52.6)	(23.4)	(29.7)	(25.3)	(43.4)
Register	All birthweights within 350-6000 g, n	1658	529	0	1757	0	6698	129	6183	1013	2458	2800	17,631
	All heaped birthweights within 350-6000 g, (%)	(4.2)	(42.7)	(0.0)	(41.0)	(0.0)	(37.3)	(38.8)	(43.8)	(17.0)	(23.8)	(17.1)	(37.1)
Day shifts (07:00 h-20:59 h)													
Observation	All birthweights within 350-6000 g, n	1661	481	0	1731	0	4856	82	3866	683	1594	2426	12,528
	All heaped birthweights within 350-6000 g, (%)	(4.9)	(40.5)	(0.0)	(37.4)	(0.0)	(35.4)	(26.8)	(40.7)	(12.9)	(22.4)	(12.9)	(34.7)
Survey	All birthweights within 350-6000 g, n	1446	408	0	1587	0	4320	44	2592	513	1078	2003	9985
	All heaped birthweights within 350-6000 g, (%)	(13.2)	(49.0)	(0.0)	(46.3)	(0.0)	(38.5)	(40.9)	(52.2)	(22.8)	(28.5)	(22.5)	(42.2)
Register	All birthweights within 350-6000 g, n	1275	351	0	1313	0	4495	83	3718	670	1558	2028	11,439
	All heaped birthweights within 350-6000 g, (%)	(3.1)	(39.6)	(0.0)	(40.0)	(0.0)	(35.8)	(28.9)	(42.5)	(15.8)	(24.1)	(13.6)	(35.7)
Night shift (21:00 h-06:59 h)													
Observation	All birthweights within 350-6000 g, n	479	260	0	544	0	2313	51	2552	354	915	884	6584
	All heaped birthweights within 350-6000 g, (%)	(8.8)	(50.4)	(0.0)	(41.2)	(0.0)	(39.6)	(47.1)	(43.8)	(15.0)	(21.9)	(20.0)	(38.6)
Survey	All birthweights within 350-6000 g, n	412	220	0	518	0	2043	30	1714	234	605	676	5100
	All heaped birthweights within 350-6000 g, (%)	(14.3)	(56.8)	(0.0)	(44.4)	(0.0)	(42.7)	(56.7)	(53.2)	(24.8)	(31.9)	(27.5)	(45.3)
Register	All birthweights within 350-6000 g, n	383	178	0	444	0	2203	46	2464	343	899	772	6190
	All heaped birthweights within 350-6000 g, (%)	(8.1)	(48.9)	(0.0)	(44.1)	(0.0)	(40.3)	(56.5)	(45.9)	(19.2)	(23.5)	(23.5)	(39.8)

[a]Percentages are pooled (random effects)
Total births stratified by birth during day or night. Further calculations are shown in Additional files 20 and 21

Table 5 Heaping index of observer-assessed, survey-reported, and register-recorded birthweights, EN-BIRTH study

		Bangladesh				Nepal		Tanzania				All sites pooled (Random effects)	
		Azimpur		Kushtia		Pokhara		Temeke		Muhimbili			
		Tertiary		District		Regional		Regional		National			
		Digital	Analogue	Digital	Analogue	Digital	Analogue	Digital	Analogue	Digital	Analogue	Digital	Analogue
Observation	1000 g heaping index, within 751-1249 g n = 162, (%)	*	(0.0)	(0.0)	(33.3)	(0.0)	(38.5)	*	(26.7)	(9.5)	(20.7)	(9.5)	(20.9)
	1500 g heaping index, within 1251-1749 g, n = 452, (%)	(0.0)	(0.0)	(0.0)	(23.5)	(0.0)	(41.1)	(100.0)	(33.8)	(14.5)	(24.9)	(8.7)	(27.5)
	2000 g heaping index, within 1751-2249 g, n = 1293, (%)	(2.3)	(25.9)	(0.0)	(37.4)	(0.0)	(52.7)	(33.3)	(45.5)	(15.2)	(22.4)	(6.8)	(37.1)
	2500 g heaping index, within 2251-2749 g, n = 5295, (%)	(7.4)	(51.7)	(0.0)	(37.5)	(0.0)	(43.5)	(33.3)	(40.2)	(11.9)	(24.2)	(12.1)	(39.1)
	3000 g heaping index, within 2751-3249 g, n = 7930, (%)	(5.6)	(49.5)	(0.0)	(39.6)	(0.0)	(44.3)	(32.6)	(46.3)	(15.1)	(23.8)	(15.1)	(40.4)
	3500 g heaping index, within 3251-3749 g, n = 4678, (%)	(4.2)	(49.5)	(0.0)	(42.3)	(0.0)	(42.5)	(36.7)	(41.1)	(13.0)	(21.3)	(14.9)	(38.7)
	4000 g heaping index, within 3751-4249 g, n = 993, (%)	(6.4)	(43.8)	(0.0)	(29.9)	(0.0)	(30.8)	(38.5)	(33.1)	(17.1)	(15.3)	(16.3)	(28.3)
Survey	1000 g heaping index, within 751-1249 g, n = 74, (%)	(0.0)	*	(0.0)	(28.6)	(0.0)	(47.1)	*	(0.0)	(28.6)	(29.0)	(20.6)	(30.7)
	1500 g heaping index, within 1251-1749 g, n = 230, (%)	(0.0)	(0.0)	(0.0)	(34.6)	(0.0)	(41.7)	*	(21.4)	(25.7)	(30.0)	(16.3)	(30.7)
	2000 g heaping index, within 1751-2249 g, n = 1284, (%)	(6.6)	(17.0)	(0.0)	(38.4)	(0.0)	(51.9)	(66.7)	(56.4)	(21.4)	(29.1)	(15.2)	(38.9)
	2500 g heaping index, within 2251-2749 g, n = 4006, (%)	(16.9)	(68.0)	(0.0)	(46.7)	(0.0)	(46.6)	(71.4)	(51.3)	(26.4)	(33.9)	(26.5)	(49)
	3000 g heaping index, within 2751-3249 g, n = 6322, (%)	(13.6)	(54.5)	(0.0)	(47.6)	(0.0)	(45.3)	(47.8)	(57.9)	(19.8)	(30.8)	(21.5)	(47.1)
	3500 g heaping index, within 3251-3749 g, n = 3773, (%)	(17.8)	(65.8)	(0.0)	(50.4)	(0.0)	(46.8)	(43.3)	(52.0)	(27.0)	(30.2)	(26.6)	(48.1)
	4000 g heaping index, within 3751-4249 g, n = 900, (%)	(17.8)	(56.3)	(0.0)	(34.8)	(0.0)	(30.4)	(40.0)	(37.7)	(23.4)	(21.3)	(22)	(32.7)
Register	1000 g heaping index, within 751-1249 g, n = 155, (%)	*	*	(0.0)	(12.5)	(0.0)	(31.3)	*	(35.3)	(5.6)	(28.8)	(5.6)	(28.6)
	1500 g heaping index, within 1251-1749 g, n = 424, (%)	(0.0)	(0.0)	(0.0)	(23.4)	(0.0)	(41.7)	(100.0)	(29.2)	(21.3)	(28.7)	(16)	(28.3)
	2000 g heaping index, within 1751-2249 g, n = 1187	(1.6)	(46.7)	(0.0)	(43.0)	(0.0)	(48.7)	(0.0)	(45.8)	(22.2)	(22.6)	(4.5)	(40.4)
	2500 g heaping index, within 2251-2749 g, n = 4745, (%)	(5.8)	(49.1)	(0.0)	(39.2)	(0.0)	(43.4)	(36.4)	(41.5)	(15.5)	(27.1)	(15.1)	(39.7)
	3000 g heaping index, within 2751-3249 g, n = 7205, (%)	(4.3)	(51.3)	(0.0)	(43.2)	(0.0)	(45.6)	(39.5)	(49.6)	(16.5)	(23.7)	(16.7)	(42.3)
	3500 g heaping index, within 3251-3749 g, n = 4318, (%)	(2.7)	(46.8)	(0.0)	(43.5)	(0.0)	(44.8)	(40.9)	(42.2)	(14.3)	(22.5)	(15.5)	(39.3)
	4000 g heaping index, within 3751-4249 g, n = 938, (%)	(2.9)	(25.0)	(0.0)	(31.3)	(0.0)	(28.6)	(35.7)	(32.0)	(24.4)	(16.0)	(17.7)	(26)

- Undefined (n/0), *Indeterminate (0/0)
Heaping index, disaggregated by type of scale, was defined as the proportion of babies with birthweights of a specific value (e.g. 2500 g) relative to the number of babies within 500 g range of this value inclusive of this value (e.g. 2251–2749). For observed birthweights, only plausible values were included. For surveys and registers, "Don't know" and "Not recorded/Not readable" responses were excluded
Further calculations are shown in Additional file 22

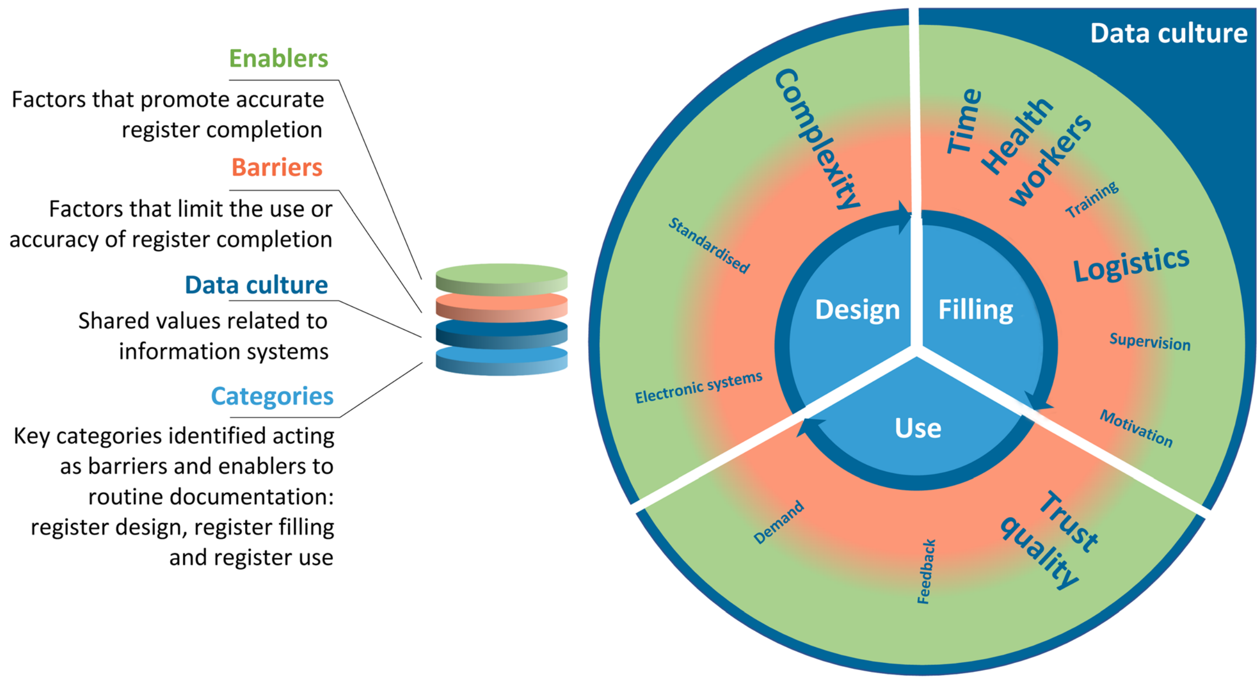


Fig. 7 Barriers and enablers to routine register recording of birthweight, EN-BIRTH study. This figure illustrates the overall barriers and enablers to facility-based data collection identified by EN-BIRTH participants. The bold text are the issues specific to birthweight. The transition from red to green is a reminder that most factors identified by participants could serve as either a barrier or enabling factor depending on the facility-level resources and management

at night. We suggest that available hospital birthweight data, stratified by day/night time of birth, could be explored as a tracer indicator for measuring quality of care. Additionally, these data can be used to assess the needs for consistent staffing during all shifts, so midwives have sufficient support to complete care and documentation tasks in a timely manner.

We identified opportunities to improve quality of birthweight register data. In Bangladesh, although original and revised register designs both included birthweight, register-recorded completeness improved substantially after introduction of the revised register design. The improvement was seen in both hospitals in Bangladesh; however, it was lower in Kushtia BD, illustrating that design alone is not sufficient. In Azimpur BD, health workers continued to record birthweight in kilogrammes to one decimal place, despite the revised register instructions to measure in grammes. Logistical challenges of revised register stock-outs in Kushtia BD necessitated using original registers again during data collection. Improving feedback loops between health workers and those at other levels of the health system using facility birthweight data is critical. Feedback could increase understanding of how birthweight data are used, why accurate measurement is essential and how to address the opportunities to improve quality of birthweight measurement in LMIC settings.

Strengths and limitations

A major strength of this study was the multi-site, multi-country design using direct observation as gold standard to compare to register records and survey report. The large sample size of > 23,000 facility births enabled diagnostic validation testing with stratification by normal and low birthweight and by mode of birth. Our observational gold standard was assessed by duplicate observation, and the effect of register recording completeness due to the presence of researchers was assessed by comparison with pre-study data extraction. Another strength is our inclusion of stillbirths, lending insight into an important public health issue, as often only live births are included when calculating birthweight indicators [44, 45]. Although the changes in the Bangladesh registers midway were unexpected, this provided the opportunity to examine the results of a "natural experiment."

However, our study also had limitations. We did not observe whether scales were calibrated prior to birthweight, which could contribute to heaping. The clinical observers read the scale at the same time as the health worker and thus could have also contributed to the observed heaping. The data collection tablet app platform collected birthweight only in grammes, while health workers recorded in registers either kilogrammes or grammes. This may have introduced information bias, affecting birthweight in terms of accuracy and reliability and a missed opportunity to compare any effect of unit of measurement on birthweight data quality. For the purposes of calculating the heaping indices, we assumed that all the birthweights of interest were heaped despite a proportion of them being truly a multiple of 500 g. We could not apply a correction for multiplicity.

Our findings of highly accurate register-recorded birthweights in CEmONC hospitals may not be generalizable to facilities at other levels of the health system. Moreover, our study intentionally focused on facility delivery; while the global facility delivery rate is > 80%, in the EN-BIRTH study countries, it is only 37% in Bangladesh, 57% in Nepal and 63% in Tanzania [15, 46]. The validity of birthweight measurement in population-based studies has been addressed in a parallel study [47].

Research gaps

Globally, there remains a large gap between facility births and availability of birthweight data in routine systems in both south Asia (19.6%) and sub-Saharan Africa (48.3%) [48]. Further research regarding data flow and quality of aggregated facility birthweight data from facilities at all levels of the health system is critical.

Implementation research is also needed to explore how hospital birthweight data quality can be improved: using standardized weighing technique training to reduce heaping, utilizing calibrated digital scales and streamlining documentation. Even when stillbirths were weighed, women were not able to accurately report that weighing had happened. More research is required to better understand how information is provided to women following a stillbirth, and even if women are routinely allowed to see their stillborn baby. Since EN-BIRTH only assessed women's report at hospital exit, follow-up studies are needed to determine if exit survey-reported accuracy decays over time, considering household surveys are usually every 2–5 years. Studies could be conducted to explore if household survey estimates of LBW are improved if birthweight is recorded on health cards given to parents, which they can show at the time of the survey [49].

Conclusions

We found high individual-level validity for coverage of weighing at birth and LBW classification in both registers and surveys, with the former outperforming the latter. Our results provide evidence supporting the use of both these data sources to increase the availability of birthweight data in LMICs. Surveys will remain an important data source especially in the most vulnerable populations, where deliveries mostly occur at home. Given the increase in facility births worldwide, birthweight data recorded in registers and incorporated into routine administrative systems can provide essential information for programs and policies. Currently, registers are an underused source of information. However, registers could offer a cost-efficient way to generate more frequent coverage measurements compared to intermittent population-based surveys. Register data completeness are already high. Closing data quality gaps for birthweight heaping will require standardised processes and ensuring facilities have sufficient staffing to carry out care and documentation in a timely manner. Only then will each and every newborn – even the smallest, sickest, and most marginalized – be counted and weighed, and countries have better data to track how many survive and thrive.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12884-020-03355-3.

Additional file 1. EN-BIRTH study sites — National mortality rates and hospital context.

Additional file 2. EN-BIRTH study data collection dates by site and time elapsed between birth and exit survey.

Additional file 3. STROBE statement.

Additional file 4. EN-BIRTH data collection flow.

Additional file 5. EN-BIRTH implausible birthweights.

Additional file 6. Comparison of original vs. revised Bangladesh registers in EN-BIRTH.

Additional file 7. Inter-observer agreement (Kappa) for gold standard observational data, EN-BIRTH study.

Additional file 8. Labour ward register data extraction completeness comparison pre-study and during-study for EN-BIRTH.

Additional file 9. Weighing coverage and LBW prevalence, EN-BIRTH study (figure).

Additional file 10. Weighing coverage and LBW prevalence in EN-BIRTH study (table).

Additional file 11. Chi-squared test results comparing EN-BIRTH weighing coverage and LBW prevalence, disaggregated.

Additional file 12. EN-BIRTH study birthweight validation results.

Additional file 13. Bland-Altman plots comparing observed EN-BIRTH birthweights with survey-reported and register-recorded.

Additional file 14. Types of weighing scales used in EN-BIRTH study, Total denotes babies who were observed to be weighed.

Additional file 15. Adjusted LBW prevalence in exit surveys and routine registers, EN-BIRTH study.

Additional file 16. Interview results with data collectors and health workers on barriers and enablers checklist, EN-BIRTH study.

Additional file 17. EN-BIRTH interview results with data collectors and health workers on estimated time to complete documentation.

Additional file 18. Ethical approval of local institutional review boards, EN-BIRTH study.

Additional file 19. Weighing and low birthweight indicators individual-level validation showing two-way tables, EN-BIRTH study.

Additional file 20. EN-BIRTH birthweight heaping index and measurement ratios, day shift. Day shift = 07:00–20:59.

Additional file 21. EN-BIRTH birthweight heaping index and measurement ratios, night shift. Night shift = 21:00–06:59.

Additional file 22. EN-BIRTH birthweight heaping index.

Abbreviations

BD: Bangladesh; CEmONC: Comprehensive emergency obstetric and newborn care; CIFF: Children's Investment Fund Foundation; DHS: The Demographic and Health Surveys Program; EN-BIRTH: *Every Newborn*-Birth Indicators Research Tracking in Hospitals study; FGD: Focus Group Discussions; g: Grammes; HMIS: Health Management Information Systems; icddr,b: International Centre for Diarrheal Disease Research, Bangladesh; IHI: Ifakara Health Institute, Tanzania; LBW: Low Birthweight; LMIC: Low-Middle Income Country/countries; LSHTM: London School of Hygiene & Tropical Medicine; MARCH Centre: Maternal, Adolescent, Reproductive & Child Health Centre, LSHTM; MCHTI: Maternal and Child Health Training Institute, Azimpur, Bangladesh; MUHAS: Muhimbili University of Health and Allied Sciences, Tanzania; MICS: Multiple Indicator Cluster Survey; NMR: Neonatal Mortality Rate; NP: Nepal; PRISM: Performance of Routine Information System Management; SDG: Sustainable Development Goal; TZ: Tanzania; UNICEF: United Nations Children's Fund; WHO: World Health Organization

Acknowledgements

Firstly, and most importantly, we thank the women, their families, the health workers and data collectors. We credit the inspiration of the late Godfrey Mbaruku. We thank Claudia DaSilva, Veronica Ulaya, Mohammad Raisul Islam, Sudip Karki and Rabina Sarki for their administrative support and Sabrina Jabeen, Goutom Banik, Md. Shahidul Alam, Tamatun Islam Tanha and Md. Mohsiur Rahman for support during data collectors training.

We acknowledge the following groups for their guidance and support:

National Advisory Groups:

Bangladesh: Mohammod Shahidullah, Khaleda Islam, Md Jahurul Islam.

Nepal: Naresh P KC, Parashu Ram Shrestha.

Tanzania: Muhammad Bakari Kambi, Georgina Msemo, Asia Hussein, Talhiya Yahya, Claud Kumalija, Eliudi Eliakimu, Mary Azayo, Mary Drake, Honest Kimaro.

EN-BIRTH validation collaborative group:

Bangladesh: Md. Ayub Ali, Bilkish Biswas, Rajib Haider, Md. Abu Hasanuzzaman, Md. Amir Hossain, Ishrat Jahan, Rowshan Hosne Jahan, Jasmin Khan, M A Mannan, Tapas Mazumder, Md. Hafizur Rahman, Md. Ziaul Haque Shaikh, Aysha Siddika, Taslima Akter Sumi, Md. Taqbir Us Samad Talha.

Tanzania: Evelyne Assenga, Claudia Hanson, Edward Kija, Rodrick Kisenge, Karim Manji, Fatuma Manzi, Namala Mkopi, Mwifadhi Mrisho, Andrea Pembe.

Nepal: Jagat Jeevan Ghimire, Rejina Gurung, Elisha Joshi, Avinash K Sunny, Naresh P. KC, Nisha Rana, Shree Krishna Shrestha, Dela Singh, Parashu Ram Shrestha, Nishant Thakur.

LSHTM: Hannah Blencowe, Sarah G Moxon.

EN-BIRTH Expert Advisory Group: Agbessi Amouzou, Tariq Azim, Debra Jackson, Theopista John Kabuteni, Matthews Mathai, Jean-Pierre Monet, Allisyn C Moran, Pavani K Ram, Barbara Rawlins, Jennifer Requejo, Johan Ivar Sæbø, Florina Serbanescu, Lara Vaz.

We are also very grateful to fellow researchers who peer-reviewed this paper.

About this supplement

This article has been published as part of BMC Pregnancy and Childbirth Volume 21 Supplement 1, 2021: *Every Newborn* BIRTH multi-country validation study: informing measurement of coverage and quality of maternal and newborn care. The full contents of the supplement are available online at https://bmcpregnancychildbirth.biomedcentral.com/articles/supplements/volume-21-supplement-1.

Authors' contributions

The EN-BIRTH study was conceived by JEL, who acquired the funding and led the overall design with support from HR. Each of the three country research teams input to design of data collection tools and review processes, data collection and quality management with technical coordination from HR, GRGL, and DB. The iccdr,b team (notably AER, TT, TH, QSR, SA and SBZ) led the development of the software application, data dashboards and database development with VG and the LSHTM team. IHI (notably DS) coordinated work on barriers and enablers for data collection and use, working closely with LTD. QSR was the main lead for data management working closely with OB, KS and LTD. For this paper, SK and LTD led the analyses and first draft of manuscript working closely with KP, PM, HB, and JEL. All authors (SK, LTD, SBZ, KP, NS, AKS, DS, QSR, AKC, HR, SEA, PM, MEG, HB, JEL) revised the manuscript and gave final approval of the version to be published and agree to be accountable for the work. The EN-BIRTH study group authors made contributions to the conception, design, data collection or analysis or interpretation of data. This paper is published with permission from the Directors of Ifakara Health Institute, Muhimbili University of Health and Allied Sciences, icddr,b and Golden Community. The authors' views are their own, and not necessarily from any of the institutions they represent. EN-BIRTH Study Group: Bangladesh: Qazi Sadeq-ur Rahman, Ahmed Ehsanur Rahman, Tazeen Tahsina, Sojib Bin Zaman, Shafiqul Ameen, Tanvir Hossain, Abu Bakkar Siddique, Aniqa Tasnim Hossain, Tapas Mazumder, Jasmin Khan, Taqbir Us Samad Talha, Rajib Haider, Md. Hafizur Rahman, Anisuddin Ahmed, Shams El Arifeen. Nepal: Omkar Basnet, Avinash K Sunny, Nishant Thakur, Rejina Gurung, Anjani Kumar Jha, Bijay Jha, Ram Chandra Bastola, Rajendra Paudel, Asmita Paudel, Ashish KC. Tanzania: Nahya Salim, Donat Shamba, Josephine Shabani, Kizito Shirima, Menna Narcis Tarimo, Godfrey Mbaruku (deceased), Honorati Masanja. LSHTM: Louise T Day, Harriet Ruysen, Kimberly Peven, Vladimir Sergeevich Gordeev, Georgia R Gore-Langton, Dorothy Boggs, Stefanie Kong, Angela Baschieri, Simon Cousens, Joy E Lawn.

Funding

The Children's Investment Fund Foundation (CIFF) are the main funder of the EN-BIRTH Study and funding is administered via The London School of Hygiene & Tropical Medicine. The Swedish Research Council specifically funded the Nepal site through Lifeline Nepal and Golden Community. We acknowledge the core funders for all the partner institutions. Publication of this manuscript has been funded by CIFF. CIFF attended the study design workshop but had no role in data collection, analysis, data interpretation, report writing or decision to submit for publication. The corresponding author had full access to study data and final responsibility for publication submission decision.

Availability of data and materials

The datasets generated during and/or analysed during the current study are available on LSHTM Data Compass repository, https://datacompass.lshtm.ac.uk/955/.

Ethics approval and consent to participate

This study was granted ethical approval by institutional review boards in all operating counties in addition to the London School of Hygiene & Tropical Medicine (Additional file 18).

Voluntary informed written consent was obtained from all observed participants, their families for newborns, and respondents for the qualitative interviews. Participants were assured of anonymity and confidentiality. All women were provided with a description of the study procedures in their preferred language at admission, and offered the right to refuse, or withdraw consent at any time during the study. Facility staff were identified before data collection began and no health worker refused to be observed whilst providing care.

EN-BIRTH is study number 4833, registered at https://www.researchregistry.com.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Centre for Maternal, Adolescent, Reproductive & Child Health (MARCH), London School of Hygiene & Tropical Medicine (LSHTM), London, UK. [2]Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh. [3]Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, King's College

London, London, UK. [4]Department of Paediatrics and Child Health, Muhimbili University of Health and Allied Sciences (MUHAS), Dar Es Salaam, Tanzania. [5]Department of Health Systems, Impact Evaluation and Policy, Ifakara Health Institute (IHI), Dar es Salaam, Tanzania. [6]Golden Community, Lalitpur, Nepal. [7]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden. [8]Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Public Health, London School of Hygiene and Tropical Medicine, London, UK.

Published: 26 March 2021

References

1. World Health Organization: Reproductive Health Indicators: Guidelines for Their Generation, Interpretation, and Analysis for Global Monitoring. https://www.who.int/reproductivehealth/publications/monitoring/924156315x/en/. Accessed 14 Aug 2020.
2. World Health Organization: Global Targets 2025. https://www.who.int/nutrition/global-target-2025/en/]. Accessed 9 Dec 2019.
3. UNICEF, WHO: UNICEF-WHO low birthweight estimates: levels and trends 2000–2015. https://www.unicef.org/reports/UNICEF-WHO-low-birthweight-estimates-2019/. Accessed 14 Aug 2020.
4. World Health Organization: Survive and Thrive: Transforming Care for Every Small and Sick Newborn. https://www.unicef.org/reports/transforming-care-for-every-small-and-sick-newborn-2020. Accessed 13 Aug 2020.
5. Blencowe H, Krasevec J, de Onis M, Black RE, An X, Stevens GA, Borghi E, Hayashi C, Estevez D, Cegolon L, et al. National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. Lancet Glob Health. 2019;7(7):e849–60.
6. Gu H, Wang L, Liu L, Luo X, Wang J, Hou F, Nkomola PD, Li J, Liu G, Meng H, et al. A gradient relationship between low birth weight and IQ: a meta-analysis. Sci Rep. 2017;7(1):18035.
7. Christian P, Lee SE, Donahue Angel M, Adair LS, Arifeen SE, Ashorn P, Barros FC, Fall CH, Fawzi WW, Hao W, et al. Risk of childhood undernutrition related to small-for-gestational age and preterm birth in low- and middle-income countries. Int J Epidemiol. 2013;42(5):1340–55.
8. Risnes KR, Vatten LJ, Baker JL, Jameson K, Sovio U, Kajantie E, Osler M, Morley R, Jokela M, Painter RC, et al. Birthweight and mortality in adulthood: a systematic review and meta-analysis. Int J Epidemiol. 2011;40(3):647–61.
9. UN-IGME. A Neglected Tragedy - The global burden of stillbirths 2020. https://www.unicef.org/media/84851/file/UN-IGME-the-global-burden-of-stillbirths-2020.pdf. Accessed 25 Oct 2020.
10. World Health Organization: Global Reference List of 100 Core Health Indicators (plus health-related SDGs). https://www.who.int/healthinfo/indicators/2018/en/. Accessed 13 Aug 2020.
11. Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller A-B, Kinney M, Lawn J, Born Too Soon Preterm Birth Action G. Born too soon: the global epidemiology of 15 million preterm births. Reprod Health. 2013;10(Suppl 1):S2.
12. UNICEF, WHO: Every Newborn; An Action Plan to End Preventable Deaths. https://www.who.int/maternal_child_adolescent/newborns/every-newborn/en/. Accessed 22 Dec 2018.
13. MEASURE Evaluation: Birthweight Specific Neonatal Mortality Rate (BWSNMR). https://www.measureevaluation.org/prh/rh_indicators/womens-health/nb/birth-weight-specific-mortality-rate-bwsmr. Accessed 9 Dec 2019.
14. Lawn JE, Gravett MG, Nunes TM, Rubens CE, Stanton C, the GRG. Global report on preterm birth and stillbirth (1 of 7): definitions, description of the burden and opportunities to improve data. BMC Pregnancy Childbirth. 2010;10(1):S1.
15. UNICEF: The State of the World's Children 2019. Children, Food and Nutrition: Growing well in a changing world. https://www.unicef.org/sowc/. Accessed 13 Aug 2020.
16. Blanc AK, Warren C, McCarthy KJ, Kimani J, Ndwiga C, RamaRao S. Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya. J Glob Health. 2016;6(1):010405.
17. Blanc AK, Diaz C, McCarthy KJ, Berdichevsky K. Measuring progress in maternal and newborn health care in Mexico: validating indicators of health system contact and quality of care. BMC Pregnancy Childbirth. 2016;16(1):255.
18. McCarthy KJ, Blanc AK, Warren CE, Kimani J, Mdawida B, Ndwidga C. Can surveys of women accurately track indicators of maternal and newborn care? A validity and reliability study in Kenya. J Glob Health. 2016;6(2): 020502.
19. Chang KT, Mullany LC, Khatry SK, Le-Clerq SC, Munos MK, Katz J. Validation of maternal reports for low birthweight and preterm birth indicators in rural Nepal. J Glob Health. 2018;8(1):2047–978.
20. Bhattacharya AA, Umar N, Audu A, Felix H, Allen E, Schellenberg JRM, Marchant T. Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: a case study from Gombe state, Nigeria. PLoS One. 2019;14(1):e0211265.
21. Moxon SG, Ruysen H, Kerber KJ, Amouzou A, Fournier S, Grove J, Moran AC, Vaz LM, Blencowe H, Conroy N. Count every newborn; a measurement improvement roadmap for coverage data. BMC Pregnancy Childbirth. 2015;15(2):S8.
22. Jolivet RR, Moran AC, O'Connor M, Chou D, Bhardwaj N, Newby H, Requejo J, Schaaf M, Say L, Langer A. Ending preventable maternal mortality: phase II of a multi-step process to develop a monitoring framework, 2016–2030. BMC Pregnancy Childbirth. 2018;18(1):258.
23. Day LT, Ruysen H, Gordeev VS, Gore-Langton GR, Boggs D, Cousens S, Moxon SG, Blencowe H, Baschieri A, Rahman AE, et al. "Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. J Glob Health. 2019;9(1):010902.
24. Day LT, Rahman QS, Rahman AE, Salim N, KC A, Ruysen H, Tahsina T, Masanja H, Basnet O, Gore-Langton GR, et al. Assessment of the validity of the measurement of newborn and maternal health-care coverage in hospitals (EN-BIRTH): an observational study. Lancet Glob Health. 2020. https://doi.org/10.1016/S2214-109X(20)30504-0.
25. Gladstone ME, Salim N, Ogillo K, Shamba D, Gore-Langton GR, Day LT, Blencowe H, Lawn JE. Birthweight measurement processes and perceived value: qualitative research in one EN-BIRTH study hospital in Tanzania. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03356-2.
26. Ruysen H, Rahman AE, Rahman QS, Zaman SB, Hossain T, Basnet O, Shirima K, Gordeev VS, Arifeen SE, Lawn JE. Electronic data collection for multi-country, hospital-based clinical observation of maternal and newborn care: experiences from the EN-BIRTH study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03426-5.
27. StataCorp. Stata Statistical Software: Release 16. College Station: Stata Press; 2019.
28. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013.
29. Day LT, Gore-Langton GR, Rahman AE, Basnet O, Shabani J, Tahsina T, Poudel A, Shirima K, Ameen S, Ashish KC, et al. Labour & Delivery register data availability, quality, and utility: Every Newborn-Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study baseline analysis in three countries. BMC Health Serv Res. 2020;20(1):737.
30. World Health Organization: Data quality review: a toolkit for facility data quality assessment. Module 2: Desk review of data quality. https://apps.who.int/iris/handle/10665/259225. Accessed 01 July 2019.
31. World Health Organization: Data quality review: a toolkit for facility data quality assessment. Module 2: Desk review of data quality. https://apps.who.int/iris/handle/10665/259225. Accessed 1 July 2019.
32. Munos MK, Blanc AK, Carter ED, Eisele TP, Gesuale S, Katz J, Marchant T, Stanton CK, Campbell H, Improving Coverage Measurement G. Validation studies for population-based intervention coverage indicators: design, analysis, and interpretation. J Glob Health. 2018;8(2):020804.
33. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8(2):135–60.
34. Blanc AK, Wardlaw T. Monitoring low birth weight: an evaluation of international estimates and an updated estimation procedure. Bull World Health Organ. 2005;83(3):178–85.
35. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009;24(3):217–28.
36. Shamba D, Day L, Zaman S, Khan J, Talha T, Rahman M, Kayastha A, Thakur N, Tarimo M, Singh N. Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03517-3.
37. Shenkin SD, Zhang MG, Der G, Mathur S, Mina TH, Reynolds RM. Validity of recalled v. recorded birth weight: a systematic review and meta-analysis. J Dev Orig Health Dis. 2017;8(2):137–48.
38. Duffy S, Crangle M. Delivery room logbook - fact or fiction? Trop Dr. 2009; 39(3):145–9.
39. 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(2):e98–e108.
40. Peven K, Day LT, Ruysen H, Tahsina T, Ashish KC, Shabani J, Kong S, Ameen S, Basnet O, Haider R, et al. Stillbirths including intrapartum timing: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03238-7.
41. Edouard L, Senthilselvan A. Observer error and birthweight: digit preference in recording. Public Health. 1997;111(2):77–9.
42. Emmerson AJ, Roberts SA. Rounding of birth weights in a neonatal intensive care unit over 20 years: an analysis of a large cohort study. BMJ Open. 2013;3(12):e003650.
43. Rekha C, Whelan RM, Reddy P, Reddy PS. Evaluation of adjustment methods used to determine prevalence of low birth-weight babies at a rural hospital in Andhra Pradesh, India. Indian J Public Health. 2013;57(3):177–80.
44. UNICEF: Low birthweight. https://data.unicef.org/topic/nutrition/low-birthweight/. Accessed 20 Jan 2020.
45. Croft TN, Marshall AMJ, Allen CK, et al. Guide to DHS Statistics: DHS-7 (version 2). https://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_DHS-7_v2.pdf. Accessed 20 Sept 2020.
46. World Health Organization: WHO GHO and Data Portal for Global Strategy. http://apps.who.int/gho/data/node.gswcah. Accessed 7 Jan 2020.
47. Biks GA, Blencowe H, Hardy VP, Misganaw B, Angaw DA, Wagnew A, Abebe SM, Guadu T, Martins J, Fisker AB, et al. Birthweight data completeness and quality in population-based surveys: EN-INDEPTH study. BMC Popul Health Metrics. 2020. https://doi.org/10.1186/s12963-020-00229-w.
48. Day LT, Blencowe H, Carvajal-Aguirre L, Chavula K, Guenther T, Gupta G, Jackson D, Kinney M, Monet J-P, Moran A, et al. Survive and tehrive: transforming care for every small and sick newborn https://www.unicef.org/reports/transforming-care-for-every-small-and-sick-newborn-2020. Accessed 13 Aug 2020.
49. UNICEF: Low birthweight. https://data.unicef.org/topic/nutrition/low-birthweight/. Accessed 20 Jan 2020.

Publisher's Note

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

Chapter 8 – Objective 6: Neonatal resuscitation measurement

This manuscript assesses measurement opportunities for neonatal resuscitation: indicator definitions and quality of care.

The chapter was published in March 2021 in BMC Pregnancy and Childbirth. The manuscript was published under a creative commons license (Creative Commons Attribution 4.0 International License) and no further permissions are needed.

The published manuscript is included in full below and supplementary material referenced in the paper is available at https://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/s12884-020-03422-9#Sec24

8.1 List of Figures

8.2 List of Tables

8.3 Citation

Kc, A., Peven, K., Ameen, S., Msemo, G., Basnet, O., Ruysen, H., Zaman, S. B., Mkony, M., Sunny, A. K., Rahman, Q. S., Shabani, J., Bastola, R. C., Assenga, E., Kc, N. P., El Arifeen, S., Kija, E., Malla, H., Kong, S., Singhal, N., Niermeyer, S., Lincetto, O., **Day, L. T., Lawn, J. E**. and EN-BIRTH Study Group. Neonatal resuscitation: EN-BIRTH multi-country validation study.
BMC Pregnancy Childbirth. Vol. 21, 2021/03/27 edn; 2021: 235. https://doi.org/10.1186/s12884-020-03422-9

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	034282	**Title**	Dr
First Name(s)	Louise Tina		
Surname/Family Name	Day		
Thesis Title	Quality of care and quality of data for hospital births – tension or traction?		
Primary Supervisor	Associate Professor Cally Tann		

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?	Neonatal resuscitation: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. vol. 21, 2021/03/27 edn; 2021: 235. https://doi.org/10.1186/s12884-020-03422-979 Kc, A., Peven, K., Ameen, S., Msemo, G., Basnet, O., Ruysen, H., Zaman, S. B., Mkony, M., Sunny, A. K., Rahman, Q. S., Shabani, J., Bastola, R. C., Assenga, E., Kc, N. P., El Arifeen, S., Kija, E., Malla, H., Kong, S., Singhal, N., Niermeyer, S., Lincetto, O., Day, L. T., Lawn, J. E. and EN-BIRTH Study Group.		
When was the work published?	March 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 jointly designed the objectives with the co-first authors and co-senior authors. I co-designed data cleaning and analysis plan linked to my knowledge and experience of neonatal resuscitation algorithms, working with my co-first author colleagues who ran the statistical analysis. I made substantial contributions to the drafted manuscript especially the background, results, discussion and conclusion whilst mentoring my colleague one of the co-first authors, Kimberly Peven. I made substantial contributions to the paper revisions in response to peer-review.

SECTION E

Student Signature	[redacted]
Date	

Supervisor Signature	
Date	

KC et al. BMC Pregnancy and Childbirth 2020, 21(Suppl 1):235
https://doi.org/10.1186/s12884-020-03422-9

BMC Pregnancy and Childbirth

From Every Newborn BIRTH multi-country validation study: informing measurement of coverage and quality of maternal and newborn care

RESEARCH

Open Access

Neonatal resuscitation: EN-BIRTH multi-country validation study

Ashish KC[1*‡], Kimberly Peven[2,3‡], Shafiqul Ameen[4], Georgina Msemo[5,6], Omkar Basnet[7], Harriet Ruysen[2], Sojib Bin Zaman[4], Martha Mkony[5], Avinash K. Sunny[7], Qazi Sadeq-ur Rahman[4], Josephine Shabani[5], Ram Chandra Bastola[8,9], Evelyne Assenga[10], Naresh P. KC[11], Shams El Arifeen[4], Edward Kija[5], Honey Malla[7], Stefanie Kong[2], Nalini Singhal[12], Susan Niermeyer[13], Ornella Lincetto[14], Louise T. Day[2†], Joy E. Lawn[2†] and EN-BIRTH Study Group

Abstract

Background: Annually, 14 million newborns require stimulation to initiate breathing at birth and 6 million require bag-mask-ventilation (BMV). Many countries have invested in facility-based neonatal resuscitation equipment and training. However, there is no consistent tracking for neonatal resuscitation coverage.

Methods: The EN-BIRTH study, in five hospitals in Bangladesh, Nepal, and Tanzania (2017–2018), collected time-stamped data for care around birth, including neonatal resuscitation. Researchers surveyed women and extracted data from routine labour ward registers. To assess accuracy, we compared gold standard observed coverage to survey-reported and register-recorded coverage, using absolute difference, validity ratios, and individual-level validation metrics (sensitivity, specificity, percent agreement). We analysed two resuscitation numerators (stimulation, BMV) and three denominators (live births and fresh stillbirths, non-crying, non-breathing). We also examined timeliness of BMV. Qualitative data were collected from health workers and data collectors regarding barriers and enablers to routine recording of resuscitation.

Results: Among 22,752 observed births, 5330 (23.4%) babies did not cry and 3860 (17.0%) did not breathe in the first minute after birth. 16.2% ($n = 3688$) of babies were stimulated and 4.4% ($n = 998$) received BMV. Survey-report underestimated coverage of stimulation and BMV. Four of five labour ward registers captured resuscitation numerators. Stimulation had variable accuracy (sensitivity 7.5–40.8%, specificity 66.8–99.5%), BMV accuracy was higher (sensitivity 12.4–48.4%, specificity > 93%), with small absolute differences between observed and recorded BMV. Accuracy did not vary by denominator option. < 1% of BMV was initiated within 1 min of birth. Enablers to register recording included training and data use while barriers included register design, documentation burden, and time pressure.

(Continued on next page)

* Correspondence: ashish.k.c@kbh.uu.se
‡Ashish KC and Kimberly Peven are joint first authors.
†Joy E. Lawn and Louise T. Day are joint senior authors.
[1]International Maternal and Child Health, Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusions: Population-based surveys are unlikely to be useful for measuring resuscitation coverage given low validity of exit-survey report. Routine labour ward registers have potential to accurately capture BMV as the numerator. Measuring the true denominator for clinical need is complex; newborns may require BMV if breathing ineffectively or experiencing apnoea after initial drying/stimulation or subsequently at any time. Further denominator research is required to evaluate non-crying as a potential alternative in the context of respectful care. Measuring quality gaps, notably timely provision of resuscitation, is crucial for programme improvement and impact, but unlikely to be feasible in routine systems, requiring audits and special studies.

Keywords: Birth, Neonatal resuscitation, Coverage, Quality, Measurement, Validity, Survey, Hospital records, Health management information systems

Key findings

What is known and what is new about this study?

- Neonatal resuscitation programmes are being scaled up globally, yet coverage of resuscitative interventions is not routinely tracked. Resuscitation coverage and quality measures have not yet been validated in either population-based surveys or routine facility registers.
- Challenges exist for measurement of resuscitation coverage indicators:
  - *Numerator:* Which action during clinical resuscitation (e.g. stimulation or bag-mask-ventilation [BMV]) is both measurable and valid?
  - *Denominator:* What is measurable and useful (e.g. live births plus fresh stillbirths or non-breathing, or non-crying babies)?
- EN-BIRTH is the first observational study (> 23,000 births) to assess validity of neonatal resuscitation coverage measurement, in both exit survey of women's report and routine register records. Using time-stamped data, we analysed coverage and quality of neonatal resuscitation in five hospitals in Bangladesh, Nepal, and Tanzania.

Survey — what did we find and what does it mean?

- *Numerator options*: Survey-reported coverage of BMV (0.3–1.9%) markedly under-estimated observed coverage (0.7–7.1%). BMV had low sensitivity (< 21%) and high specificity (> 98%). Newborn stimulation was reported by < 3% of women, very much lower than observed coverage (5.2–21.0%).
- *Denominator options*: Crying at birth had low "don't know" responses (< 3%) in exit survey. Compared to observed crying within a minute of birth, sensitivity was high (> 95%); however, specificity was low (< 22%). Survey-reported BMV coverage validity was consistently low for all denominators assessed.

Register — what did we find and what does it mean?

- *Numerator options:* Stimulation and BMV were recorded by 4 of 5 labour ward registers, yet accuracy varied between hospitals even with the same register design. BMV sensitivity ranged from 12.4–48.4% and specificity was high (> 93%). For stimulation, sensitivity was low at 7.5–40.8% and specificity was more variable (range 66.8–99.5).
- *Denominators:* Livebirths and fresh stillbirths were recorded in all registers. The "non-crying/non-breathing" combined denominator was only in the Bangladesh registers and could not be validated. Register-recorded BMV coverage was consistent whichever denominators was applied.

Gap analysis for quality of care and measurement

- Most newborns (71.4–94.7%) who did not respond to stimulation did receive BMV, but only 1% within the recommended 1 min after birth.

Key findings *(Continued)*

What next and research gaps?

- Population-based surveys are not likely to be useful for measuring neonatal resuscitation coverage, given low validity of exit-survey report. Additionally, household surveys would be underpowered since resuscitation is required by a small proportion of babies.
- Routine hospital registers have potential to track resuscitation coverage indicators, but implementation research is needed to standardise design and processes, including data flow to Health Management Information Systems. BMV is the most accurate numerator, true denominator measurement is complex and requires more research, including assessment of non-crying.
- Data use with feedback loops and support to frontline healthcare workers could help improve data quality and quality of care. Local clinical quality improvement and special studies are important to reduce quality gaps, particularly for timely BMV, and help meet global goals to end preventable deaths.

Background

Annually, 7–14 million newborns (5–10%) are estimated to require stimulation to initiate breathing at birth and 6 million newborns require bag-mask-ventilation (BMV) [1, 2]. Intrapartum-related events (previously termed "birth asphyxia") are a leading cause of neonatal mortality, accounting for 11% of under-five deaths [2, 3]. Such intrapartum-related events can cause stillbirths just before birth and neonatal deaths just after. The majority (> 84%) of stillbirths are in low- and middle-income countries (LMICs) and an estimated 50% are intrapartum [4, 5]. Resuscitation is recommended for all babies who do not breathe after birth since live births may be misclassified as stillbirths [6, 7]. Meeting Sustainable Development Goal (SDG) targets by 2030 for ending preventable neonatal deaths requires universal coverage of high quality care around birth for women and their babies, including resuscitation for those who do not breathe at birth [8, 9]. Globally ~ 80% of births are now in facilities [10], with many LMICs scaling up neonatal resuscitation programs [11–13]. However, lack of measurement for coverage and quality of neonatal resuscitation impedes tracking of progress [14].

The definition of coverage requires a *numerator* capturing the intervention (or a component) divided by a target *denominator* regarding clinical need. A good indicator may not include all of the clinical intervention but should "indicate" well and also not incentivise undesirable practices. Resuscitation coverage measurement has specific challenges. Clinical algorithms have multiple actions that could be used as numerators, notably: stimulation of the baby or the action of BMV. Suction is indicated for some babies, but inappropriate suctioning can be harmful, thus should be avoided for a measurement focus [15].

Resuscitation algorithms start at birth for all babies, including fresh stillbirths, being dried and assessed for crying or breathing. WHO guidance on basic resuscitation focuses on the baby who is not breathing spontaneously or is depressed [16]. A global partnership called "Helping Babies Breathe," (HBB) widely used for neonatal resuscitation training in LMICs, uses crying during thorough drying as a rapid and objective assessment, then evaluating breathing (Fig. 1) [17]. In line with WHO guidelines, if the baby is not crying and not breathing, then stimulation is provided to improve or initiate breathing, and clearing of the airway if it is blocked with secretions. If the baby is not breathing after these actions BMV should begin within 1 min of birth.

Most data on maternal and newborn health care coverage in LMICs relies on population-based surveys, notably Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS), none of which capture neonatal resuscitation. Routine facility data are currently an underutilised source for neonatal resuscitation coverage for routine Health Management Information Systems (HMIS). Interventions around the time of birth are typically recorded in one or more facility documents: individual patient records, labour and delivery ward registers, and intervention-specific registers (e.g., neonatal resuscitation register) [18]. Previous research has demonstrated availability of some neonatal resuscitation data in routine labour ward registers [19, 20]. Use of HMIS data aggregated from registers is impeded by concerns regarding data quality [21], but to date no validation studies have been undertaken regarding either survey or routine register data for neonatal resuscitation coverage indicators.

The *Every Newborn* Action Plan, agreed by all 195 United Nations member states, includes an ambitious measurement improvement roadmap [9] to validate coverage indicator measurement for care and outcomes around the time of birth. The *Every Newborn*–Birth Indicators Research Tracking in Hospitals (EN-BIRTH) study was undertaken in three countries (Tanzania, Bangladesh, and Nepal) and aimed to assess validity of measurement of selected newborn and maternal indicators for routine facility-based tracking of coverage, quality of care, and outcomes [22].

Objectives

This paper is part of a supplement based on the EN-BIRTH multi-country validation study, *'Informing measurement of coverage and quality of maternal and newborn care'*, and focuses on neonatal resuscitation measurement with four objectives:

1. **Assess NUMERATOR accuracy/validity** for neonatal resuscitation coverage indicator (stimulation and BMV) measurement by exit survey of women's report and routine labour ward registers compared to direct observation (gold standard).
2. **Compare DENOMINATOR options** for resuscitation coverage measurement: including all births (except macerated stillbirths), non-crying babies and non-breathing babies.
3. **Analyse GAPS in coverage, quality of care and measurement** in relation to recommendations, notably timely initiation of BMV.
4. **Evaluate BARRIERS AND ENABLERS** to routine labour ward register recording for resuscitation regarding register design, filling, and use.

Methods

EN-BIRTH was an observational, mixed methods study comparing data from clinical observers (gold standard) to survey-reported and register-recorded coverage of perinatal care and outcomes (Fig. 2). Detailed information regarding the research protocol, methods, and analysis has been published separately [22, 23]. Data were collected from July 2017–July 2018 in five public CEmONC hospitals in three high mortality burden countries: Maternal and Child Health Training Institute, Azimpur and Kushtia General Hospital in Bangladesh (BD); Pokhara Academy of Health Sciences in Nepal (NP); Temeke Regional Hospital and Muhimbili National Referral Hospital in Tanzania (TZ). (Additional file 1). Baseline health facility assessments established that all five hospitals had capacity to resuscitate newborns. Resuscitation guidelines used in all five hospitals were based on HBB [17]. Participants were consenting women admitted in labour for care around birth. Exclusion criteria included imminent birth and no fetal heart beat heard on admission. Clinically trained researchers observed participants 24 h per day and recorded data on the baby's condition at birth (e.g., crying/breathing) and care (e.g., stimulation and BMV). The observers received refresher training in HBB as part of their clinical observation training before the study started [22]. Data were collected with a custom-built android tablet-based application, including timestamps for observations. Research data collectors interviewed women after discharge before exit from hospital regarding their baby's condition after birth and care received. Resuscitation and

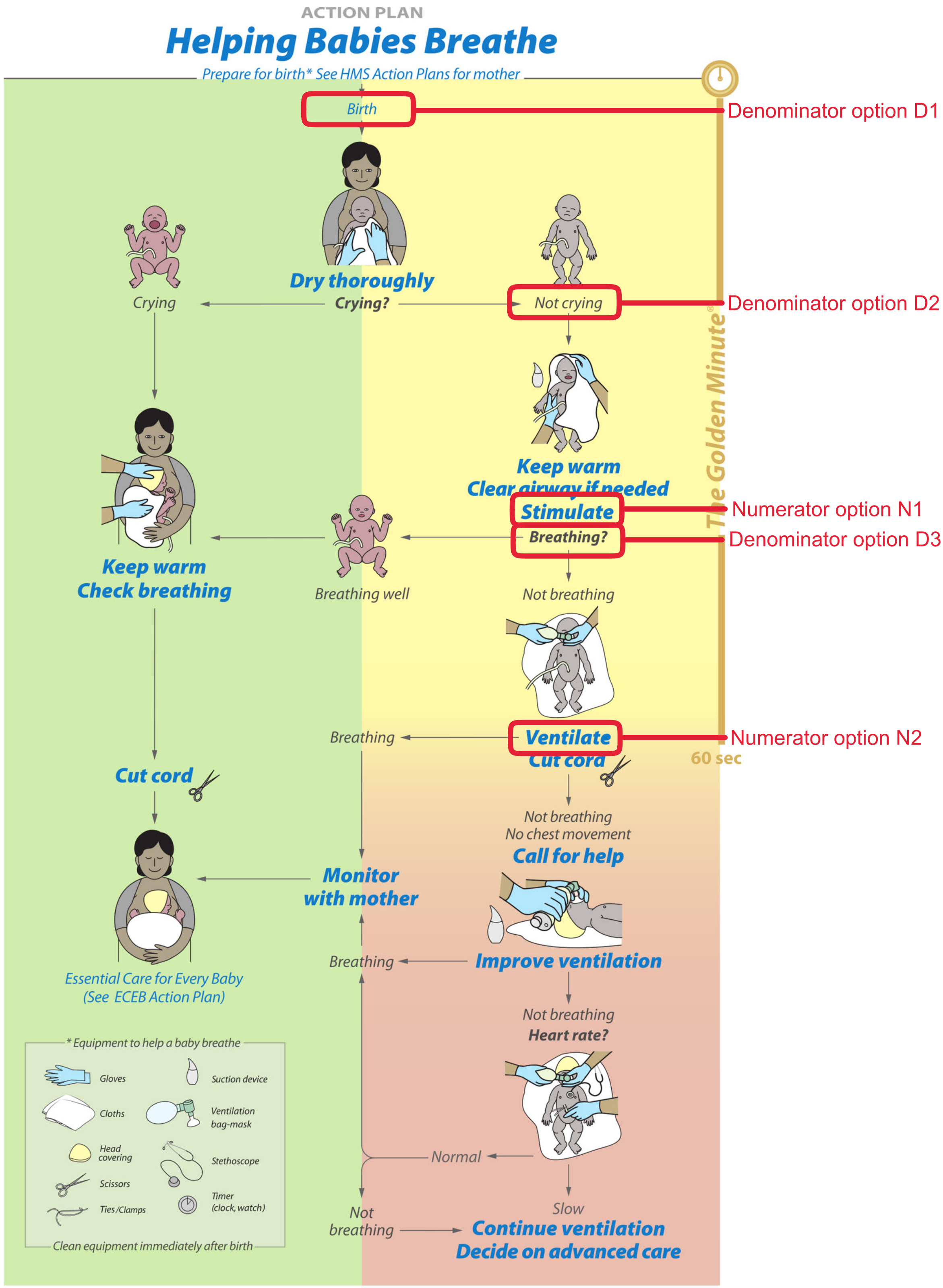


Fig. 1 Helping Babies Breathe algorithm decision points to measure neonatal resuscitation coverage

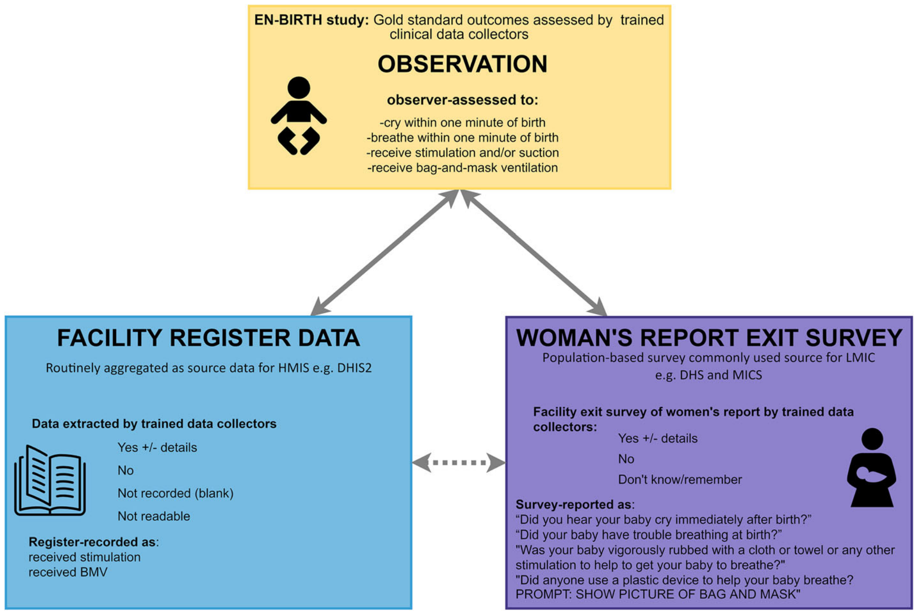


Fig. 2 Neonatal resuscitation validation design, EN-BIRTH study. EN-BIRTH validation design comparing observation gold standard with register-recorded and women's report on exit survey; EN-BIRTH data collection tools (observation checklist, register data extraction tool and exit survey tool) are published separately [22]

outcome data were extracted from routine hospital registers. Metadata definitions of selected indicator options for validity testing are shown in Additional file 2. To determine the reliability of the observational data (gold standard) supervisors duplicated observation (and register data extraction) for a subset of 5% to calculate Cohen's Kappa coefficients. Health workers and data collectors were interviewed about barriers and enablers to use of routine registers in recording of perinatal care and outcomes.

Results are reported in accordance with STROBE Statement checklists for cross-sectional studies (Additional file 3). Quantitative analysis was undertaken using R version 3.6.1 [24].

Objective 1: Numerator for indicator measurement validation

Livebirths and fresh stillbirths (hereafter referred to as "newborns"), were considered to require initial assessment for resuscitation, whilst macerated stillbirths were excluded. We explored accuracy of two possible numerator options N1) Stimulation and N2) BMV in both survey and register data compared to observation data.

In exit surveys, where a woman reported her newborn had difficulty breathing at birth, she was asked about resuscitation practices. In line with common survey indicator reporting, where women replied, "don't know" we considered the survey-reported stimulation/BMV response as "no".

We compared observed coverage (gold standard) of stimulation and BMV to survey-reported and register-recorded coverage. We calculated absolute differences between measured coverage (survey or register) and observed coverage to understand under- or over-estimation at the population level. Using two-way tables, we calculated individual-level validity statistics: sensitivity, specificity, and percent agreement ((true positive + true negative)/total) of register-recorded and survey-reported BMV coverage to measure observed coverage. Area under the curve, inflation factor, positive predictive value, and negative predictive value were also calculated. All calculations were stratified by hospital with 95% confidence intervals. Pooled results for validity analyses were calculated using random effects meta-analysis, presented with i^2, τ^2, and heterogeneity statistic (Q).

Objective 2: Denominator comparisons

We explored neonatal resuscitation coverage measurement using three possible denominator options: D1) all newborns (total births excluding macerated stillbirths), D2) newborns not crying within the first minute after birth and D3) newborns not breathing within the first minute after birth.

We compared these denominators using validity ratios (measured:observed coverage), similar to verification ratios in data quality review methods [25], for survey-reported and register-recorded BMV coverage. Validity ratios > 1 show overestimation of survey-reported or register-recorded coverage compared to observed, while ratios < 1 show underestimation. Results were heat-mapped using standard data quality review cut-offs (over/underestimate by 0–5%, 6–10%, 11–15%, 16–20 and > 20%).

Objective 3: Gap analysis for coverage and quality of care, and measurement

We examined gaps in coverage and timely neonatal resuscitation amongst a subset of newborns with a clinical need for resuscitation within 1 min of birth. These newborns were not breathing in the first minute after birth and did not respond to stimulation (or suction when performed). For this (A) eligible population subset, we analysed four gaps for neonatal resuscitation: (B) coverage gap for BMV, (C) quality of care gap between *any* BMV coverage, and timely coverage (within 1 min), (D) measurement gap for survey-report, and (E) measurement gap for register-record.

Objective 4: Barriers and enablers to routine recording

Qualitative data collection tools for focus group discussions and in-depth interviews were informed by the Performance of Routine Information System Management Series (PRISM) conceptual framework [26]. Detailed qualitative methods and overall results are available in an associated paper [27]. A purposive sample of nurses, midwives, doctors, and EN-BIRTH data collectors from each of the five hospitals participated. Analysis identified themes based on three domains: register design, filling, and use [26]. In addition, respondents were asked questions regarding the order in which resuscitation is documented in registers, patient notes, and other documents as well as how long after resuscitation is documentation entered in the labour ward register. This paper presents emerging themes regarding recording of neonatal resuscitation.

Results

Among 23,811 eligible women across the five participating hospitals, 23,724 consented to participate (Fig. 3). Among 23,471 observed births, 22,752 were live births (22,522) or fresh stillbirths (230). Data extraction was completed for 21,101 newborns (92.7%), and exit surveys were conducted with 20,245 women (90.7%). Reasons for women's non-participation in exit survey included refusals and exit from facility prior to research

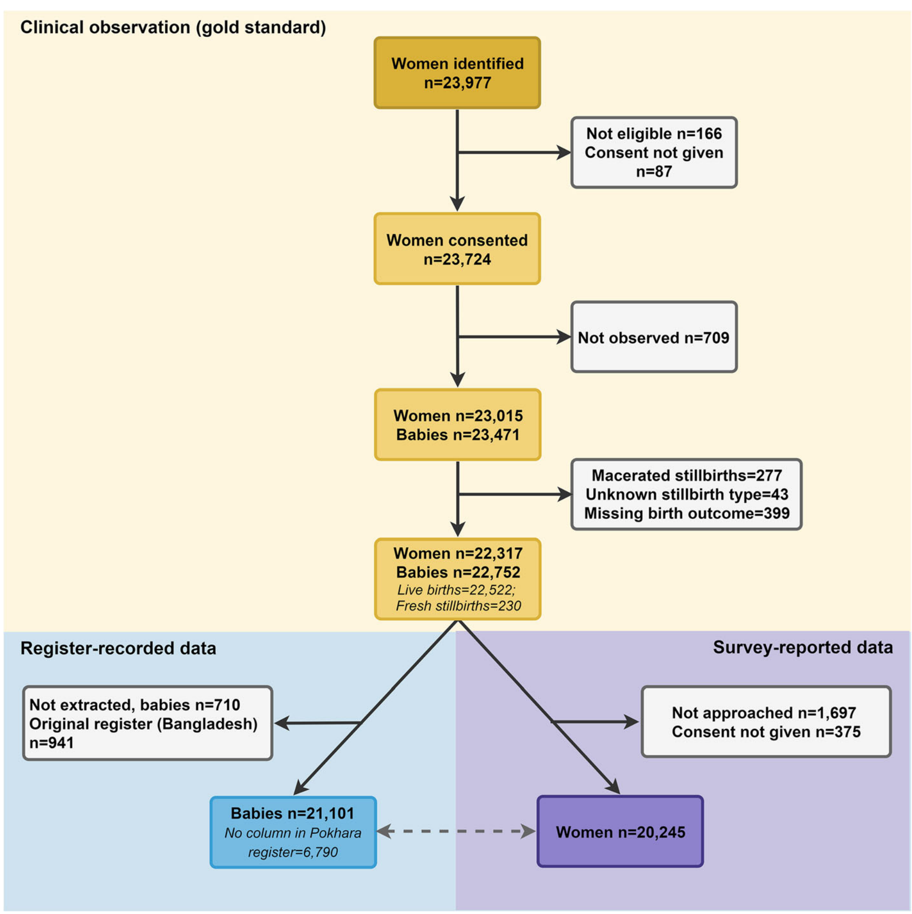


Fig. 3 Flow diagram of cases for neonatal resuscitation analysis, EN-BIRTH study (*n* = 22,752)

Table 1 Characteristics of babies and women, EN-BIRTH study (n = 22,752 births)

	Bangladesh		Nepal	Tanzania		All sites[a]
	Azimpur Tertiary	Kushtia District	Pokhara Regional	Temeke Regional	Muhimbili National	
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
a) Total babies observed	2903	2352	7211	6702	3584	22,752
Birth outcome - Live Birth	2896 (99.7)	2308 (98.2)	7175 (99.5)	6634 (99)	3509 (97.9)	22,522 (99.0)
Newborn condition at L&D discharge						
Alive	2895 (99.7)	2302 (97.9)	7171 (99.4)	6606 (98.6)	3490 (97.4)	22,464 (98.7)
Fresh stillbirth	7 (0.2)	44 (1.9)	36 (0.5)	68 (1.0)	75 (2.1)	230 (1.0)
Neonatal death	1 (0.0)	6 (0.3)	4 (0.1)	28 (0.4)	19 (0.5)	58 (0.3)
Mode of birth						
Normal vaginal birth	766 (26.4)	1369 (58.2)	5812 (80.6)	6213 (92.7)	1513 (42.2)	15,673 (68.9)
Vaginal breech/ Vacuum/ Forceps	1 (0.0)	0 (0.0)	342 (4.7)	10 (0.1)	9 (0.3)	362 (1.6)
Caesarean Section	2136 (73.6)	983 (41.8)	1057 (14.7)	478 (7.1)	2060 (57.5)	6714 (29.5)
Birthweight of baby < 2500 g	353 (12.1)	471 (20.0)	840 (11.7)	480 (7.2)	938 (26.2)	3082 (13.5)
Sex Female/Girl baby	1439 (49.6)	1143 (48.6)	3329 (46.2)	3229 (48.4)	1760 (49.5)	10,900 (48.1)
b) Total women observed[b]	2879	2309	7145	6584	3400	22,317
Women's age						
< 18 years	25 (0.9)	2 (0.1)	305 (4.3)	25 (0.4)	7 (0.2)	364 (1.6)
18–19 years	467 (16.2)	189 (8.2)	800 (11.2)	752 (11.4)	152 (4.5)	2360 (10.6)
20–24 years	1150 (39.9)	901 (39)	2989 (41.8)	2263 (34.4)	687 (20.2)	7990 (35.8)
25–29 years	856 (29.7)	714 (30.9)	2051 (28.7)	1655 (25.1)	1087 (32.0)	6363 (28.5)
30–34 years	294 (10.2)	358 (15.5)	790 (11.1)	1117 (17.0)	883 (26.0)	3442 (15.4)
35+ years	87 (3.0)	145 (6.3)	210 (2.9)	772 (11.7)	584 (17.2)	1798 (8.1)
Women's education						
No education	37 (1.3)	75 (3.2)	259 (3.6)	196 (3.0)	63 (1.9)	630 (2.8)
Primary incomplete	110 (3.8)	117 (5.1)	244 (3.4)	76 (1.2)	40 (1.2)	587 (2.6)
Primary complete	333 (11.6)	329 (14.2)	289 (4.0)	28 (0.4)	4 (0.1)	983 (4.4)
Secondary incomplete	976 (33.9)	917 (39.7)	1589 (22.2)	3956 (60.1)	1224 (36.0)	8662 (38.8)
Secondary complete or higher	1263 (43.9)	837 (36.2)	4381 (61.3)	2295 (34.9)	2055 (60.4)	10,831 (48.5)
Missing	160 (5.6)	34 (1.5)	383 (5.4)	33 (0.5)	14 (0.4)	624 (2.8)
Parity						
Nullipara	1333 (46.3)	981 (42.5)	4272 (59.8)	2848 (43.3)	1290 (37.9)	10,724 (48.1)
Multipara	1493 (51.9)	1323 (57.3)	2866 (40.1)	3723 (56.5)	2106 (61.9)	11,511 (51.6)
Missing	53 (1.8)	5 (0.2)	7 (0.1)	13 (0.2)	4 (0.1)	82 (0.4)

[a]Individually weighted
[b]Data were collected from women's registration and survey report

team approach. Table 1 shows characteristics of newborns in the EN-BIRTH study sample, by hospital. Overall, 98.7% were alive at discharge from labour and delivery, 1% were fresh stillbirths, and less than 1% were born alive but died on the labour ward. Nearly one-third of births (29.5%) were by caesarean section, highest (73.6%) in Azimpur BD.

Among 22,752 newborns (denominator option D1), 3688 (16.2%) were stimulated (numerator option N1) and 998 (4.4%) received BMV (numerator option N2) (Fig. 4). Within the first minute after birth, 5330 were observed as non-crying (denominator option D2), and among these 3860 were also observed as non-breathing (denominator option D3).

Assessing biases in the data

Duplicate case observation inter-rater reliability showed substantial agreement (> 0.71) for resuscitation elements (Additional file 4). Register extraction agreement was

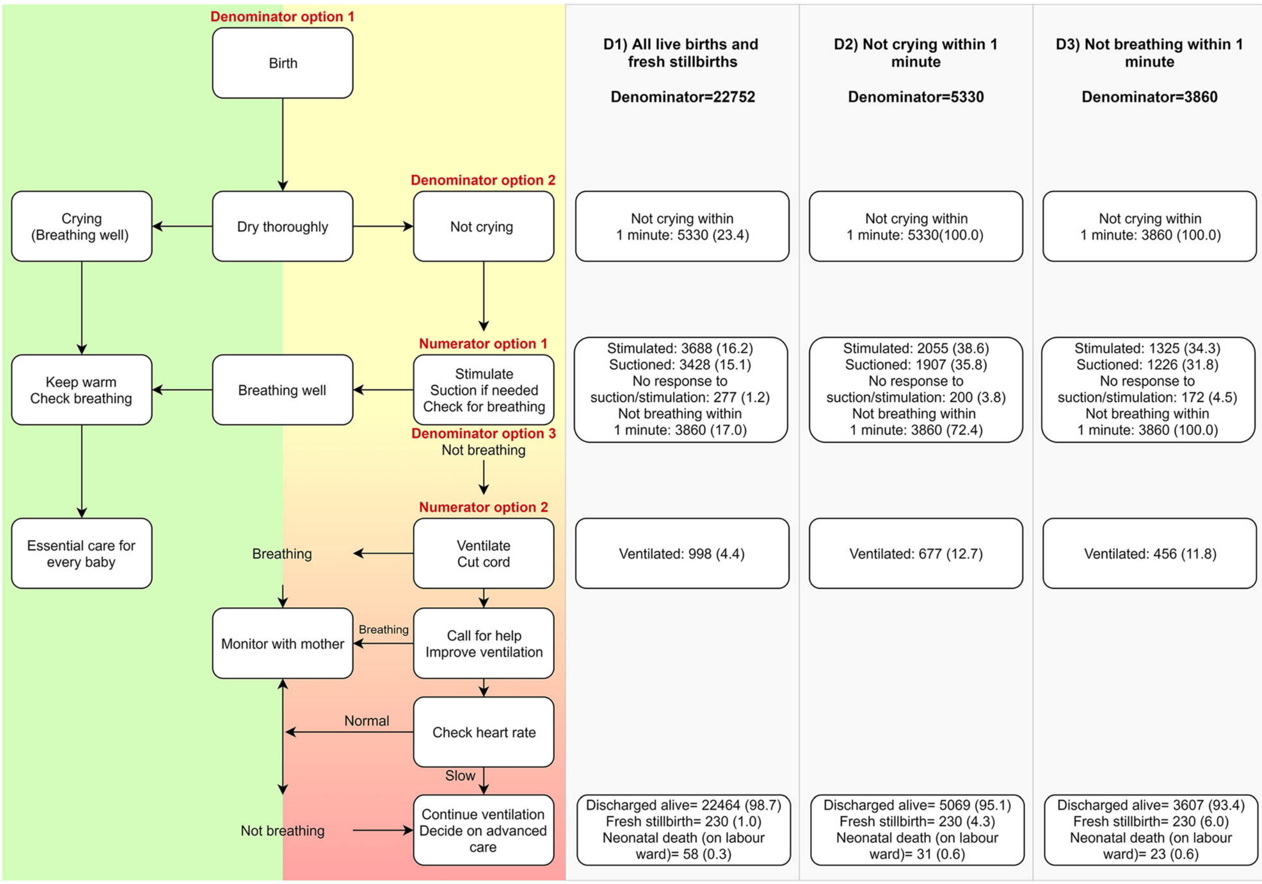


Fig. 4 Neonatal resuscitation numerators and denominators, EN-BIRTH study (individually weighted, observation data, *n* = 22,752)

Table 2 Individual-level validation in exit surveys and registers for stimulation at birth indicator, EN-BIRTH study (n = 22,752)

	Bangladesh Azimpur Tertiary	Bangladesh Kushtia District	Nepal Pokhara Regional	Tanzania Temeke Regional	Tanzania Muhimbili National	All sites Pooled (random effects)
Stimulation - Survey reported - live births + fresh stillbirths						
Observer coverage %	5.2 (4.4,6.1)	10.2 (9.0,11.5)	20.6 (19.6,21.5)	15.8 (15.0,16.7)	21 (19.7,22.4)	13.9 (8.7,20.1)
Survey reported coverage %	0.7 (0.4,1.1)	2.2 (1.7,2.9)	0.7 (0.5,0.9)	0.6 (0.5,0.9)	1.4 (1.0,1.9)	1.0 (0.6,1.6)
"Don't know" responses %	3.3 (2.7,4.1)	0.8 (0.5,1.3)	1.4 (1.1,1.7)	0.3 (0.1,0.4)	4 (3.3,4.9)	1.7 (0.6,3.3)
Sensitivity % (95% CI)	7.2 (3.5,12.9)	13.2 (9.1,18.5)	2.5 (1.7,3.6)	3.6 (2.4,5.2)	3.1 (1.7,5.1)	4.8 (2.5,7.8)
Specificity % (95% CI)	99.7 (99.3,99.8)	99 (98.4,99.4)	99.8 (99.7,99.9)	99.8 (99.6,99.9)	99 (98.5,99.4)	99.5 (99.2,99.8)
Percent agreement (TN + TP/n) %	95.0	90.6	80.2	86.8	81.6	86.9 (80.9,92.0)
Stimulation - Register recorded - live births and fresh stillbirths						
Observer coverage %	5.2 (4.4,6.1)	10.2 (9.0,11.5)	–	15.8 (15.0,16.7)	21.0 (19.7,22.4)	12.4 (6.7,19.6)
Register recorded coverage %	0.8 (0.5,1.3)	7.7 (6.6,9.0)	–	17.4 (16.5,18.3)	34.8 (33.3,36.4)	12.3 (2.3,28.7)
Not recorded %	98.7 (98.1,99.1)	91.8 (90.5,93.0)	–	20.0 (19.0,21.0)	30.8 (29.2,32.3)	65.9 (20.7,97.8)
Not readable %	0.5 (0.3,0.9)	0.4 (0.2,0.9)	–	0.3 (0.2,0.5)	0.3 (0.1,0.5)	0.4 (0.3,0.5)
Sensitivity % (95% CI)	7.5 (3.3,14.2)	15.2 (10.5,21.0)	–	39.5 (36.5,42.6)	40.8 (37.2,44.5)	24.8 (13.3,38.5)
Specificity % (95% CI)	99.5 (99.1,99.8)	93.1 (91.8,94.2)	–	86.8 (85.9,87.7)	66.8 (65,68.5)	89.4 (73.8,98.5)
Percent agreement (TN + TP/n) %	95.1	85.6	–	79.3	61.3	81.9 (67.4,92.9)

Full denominator details presented in Additional file 14 lower and varied greatly between sites, ranging from – 0.035 to 0.939.

Objective 1: Numerator for indicator measurement validation

Numerator option 1: stimulation

Observed coverage of stimulation ranged from 5.2% in Azimpur BD to 21.0% in Muhimbili TZ. Survey-report gave large underestimates for stimulation with survey-reported coverage ranging from 0.6–2.2%. Sensitivity was very low (< 14%) while specificity was high (> 98%) (Table 2; additional validity details in Additional file 5 and Additional file 6).

Register-recorded coverage (0.8–34.8%) underestimated coverage in the Bangladesh hospitals and overestimated coverage in the Tanzania hospitals (Fig. 5). While sensitivity was low (< 41%), specificity was high across most sites (66.8–99.5%).

Numerator option 2: BMV

Observed BMV ranged from 0.7% in Azimpur BD to 7.1% in Muhimbili TZ. Survey-reported coverage (0.3–1.9%) underestimated observed coverage (Fig. 6). Sensitivity was < 21% while specificity was high across all hospitals (> 98%). Register-recorded coverage (0.9–7.2%) was closer to observed coverage. While sensitivity ranged from 12.4–48.4%, specificity was > 93% across all hospitals (Table 3; additional validity details in Additional files 7 and 8).

Objective 2: Denominator for indicator measurement comparison

Denominator option 1: all newborns (live births and fresh stillbirths)

The validation of birth outcomes is reported separately [28]. Survey validity ratios for BMV coverage measurement using this all newborn denominator performed poorly (0.11–0.71) and register validity ratios were moderate to poor (0.70–1.22) (Fig. 7).

Denominator option 2: non-crying newborns

Survey-reported prevalence of crying at birth (90.5–95.8%) was higher than observed prevalence of crying within 1 min of birth (72.0–86.7%) with very few "don't know" responses (< 3%). While sensitivity was high (> 95%) specificity was low (< 22%) (Table 4; additional validity details in Additional files 9 and 10).

Survey validity ratios for BMV using this non-crying denominator performed poorly (0.13–0.58), while sensitivity was low (< 16%), specificity was high (> 98%). Register validity ratios ranged from poor to very good (0.40–0.92). While sensitivity was low (11.1–46.8%), specificity was high (> 91%).

Denominator option 3: non-breathing newborns

Prevalence of not breathing within the first minute ranged from 11.7% in Azimpur BD to 21.0% in Pokhara NP. The survey validity ratio for BMV coverage measurement using this non-breathing denominator performed poorly (0.14–0.49). Sensitivity ranged from 0

		Bangladesh Azimpur Tertiary	Bangladesh Kushtia District	Nepal Pokhara Regional	Tanzania Temeke Regional	Tanzania Muhimbili National
Stimulation						
Stimulation-register design: Column allotted data element		specific column	specific column		specific column	specific column
Column heading		Stimulation	Stimulation		Helping Babies Breathe	Helping Babies Breathe
Data element completed if stimulated		Tick	Tick		2 (= Stimulation)	2 (= Stimulation)
Data element completed if not stimulated		blank (clear instruction)	blank (clear instruction)		No (in Swahili)	No (in Swahili)
Other elements recorded in column					1=suction, 3=BMV	1=suction, 3=BMV
Denominator D1: All live births and fresh stillbirths						
Completeness	Data element recorded in register	not possible to analyse	not possible to analyse		80.0%	69.2%
External Consistency						
Indicator:	Observed coverage %	5.2%	10.2%		15.8%	21.0%
Indicator:	Measured coverage - register recorded %	0.8%	7.7%		17.4%	34.8%
Measurement gap:	Register recorded and observed	4.4% underestimate	2.5% underestimate		1.6% overestimate	13.8% overestimate
Bag Mask Ventilation						
BMV-register design: Column allotted data element		specific column	specific column		specific column	specific column
Column heading		Bag Mask Ventilated	Bag Mask Ventilated		Helping Babies Breathe	Helping Babies Breathe
Data element completed if Bag Mask Ventilated		Tick	Tick		3 (= BMV)	3 (= BMV)
Data element completed if not Bag Mask Ventilated		blank (clear instruction)	blank (clear instruction)		No (in Swahili)	No (in Swahili)
Other elements recorded in column					1=suction, 2=stimulation	1=suction, 2=stimulation
Denominator D1: All live births and fresh stillbirths						
Completeness	Data element recorded in register	not possible to analyse	not possible to analyse		91.0%	54.6%
External Consistency						
Indicator:	Observed coverage %	0.7%	5.9%		6.4%	7.1%
Indicator:	Measured coverage - register recorded %	0.9%	7.2%		5.4%	5.0%
Measurement gap:	Register recorded and observed	0.2% overestimate	1.3% overestimate		1.0% underestimate	2.1% underestimate

Key

no column for data element	
specific column	
>20%	Poor
16-20%	Moderate
11-15%	Good
6-10%	Very Good
0-5%	Excellent

Fig. 5 Hospital register design and completeness for stimulation and bag-mask-ventilation, EN-BIRTH study (*n* = 22,752)

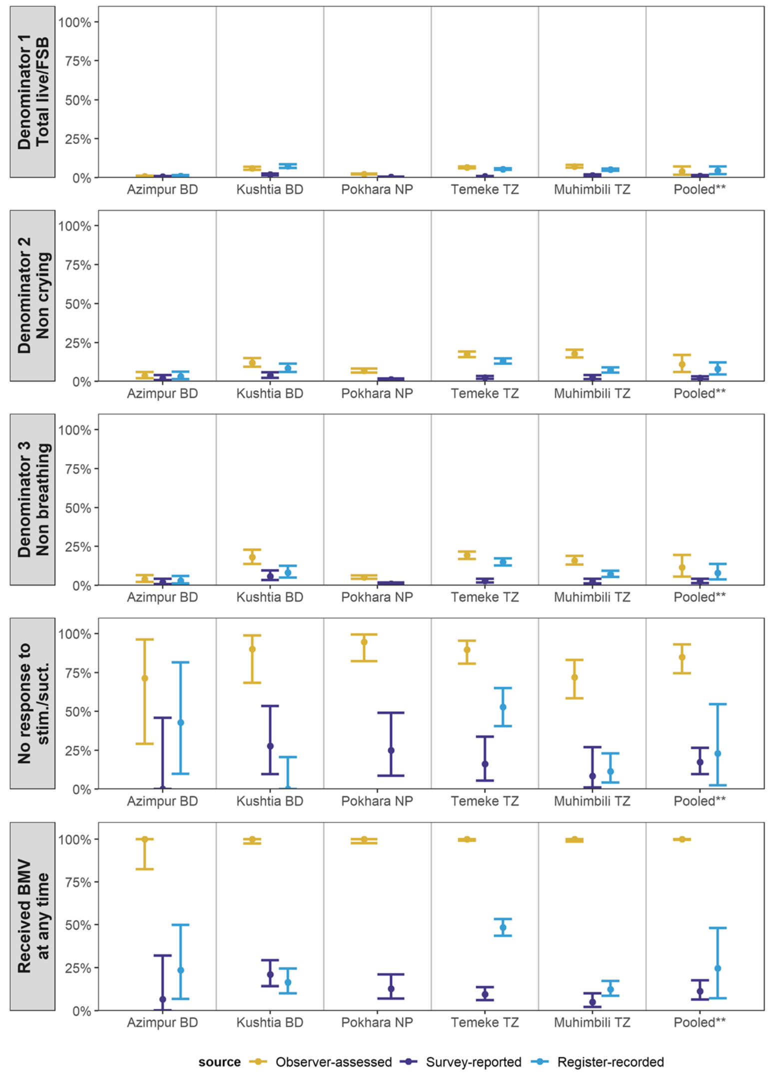


Fig. 6 Coverage (and 95%CI) of bag-mask-ventilation measured by observation, register, and exit survey, EN-BIRTH study (*n* = 22,752). *Random effects meta-analysis; BD = Bangladesh, NP = Nepal, TZ = Tanzania, stim. = stimulation, suct. = suction, BMV = bag mask ventilation, FSB = fresh stillbirth; BMV = bag mask ventilation; Full denominator details presented in Additional file 14

to 20.8% while specificity was > 97% across hospitals. Register validity ratios were better, but still classified as poor (0.45–0.78). While sensitivity ranged from 11.1–51.3%, specificity was high across all hospitals (> 92%).

Objective 3: Coverage and quality gap analysis

Among the subset proxy for true clinical need [newborns who did not cry/breathe in the first minute with no response to stimulation (or suction if needed)], most received BMV, ranging from 71.4% in Azimpur BD to 94.7% in Pokhara NP (Fig. 8) but timely coverage was very low (1%). Survey-reported coverage (< 28%) substantially underestimated true coverage. Register-recorded coverage also underestimated true coverage and ranged widely from 0.0% in Kushtia BD to 52.9% in Temeke TZ.

Among newborns receiving any BMV on the labour ward, the proportion receiving the first ventilation breath within 1 min of birth ranged from 0.2% in Temeke TZ to 8.0% in Pokhara NP. Across the three denominators explored, time to initiation of BMV was similar (Fig. 9).

Objective 4: Barriers and enablers to routine recording

Register design

Labour ward registers varied in design, between the five hospitals (Fig. 5). Bangladesh labour ward registers had three specific columns for recording neonatal resuscitation: (i) “baby did not breathe/cry after birth” (tick box for ‘yes’ and tick box for ‘no’), (ii) “stimulation” (instructions to tick for ‘yes’ and leave blank for ‘no’) and (iii) “BMV” (instructions to tick for ‘yes’ and leave

Table 3 Individual-level validation in registers and exit surveys of bag-mask-ventilation indicator, EN-BIRTH study (n = 22,752)

	Bangladesh Azimpur Tertiary	Bangladesh Kushtia District	Nepal Pokhara Regional	Tanzania Temeke Regional	Tanzania Muhimbili National	All sites Pooled (random effects)
D1) Live births and fresh stillbirths						
Survey-reported						
Observer coverage %	0.7 (0.4,1.0)	5.9 (5.0,7.0)	2.1 (1.8,2.5)	6.4 (5.9,7.1)	7.1 (6.3,8.0)	4.0 (1.7,7.0)
Survey reported coverage %	0.5 (0.3,0.9)	1.9 (1.4,2.6)	0.3 (0.2,0.5)	0.7 (0.5,1.0)	1.4 (1.0,1.9)	0.9 (0.5,1.5)
"Don't know" responses %	3.5 (2.8,4.3)	1.0 (0.7,1.6)	1.6 (1.3,1.9)	0.2 (0.1,0.4)	4.2 (3.5,5.1)	1.8 (0.6,3.6)
Sensitivity % (95% CI)	6.7 (0.2,31.9)	21 (14.2,29.2)	12.9 (7.0,21.0)	9.5 (6.2,13.8)	5.0 (2.0,10.1)	11.3 (6.4,17.5)
Specificity % (95% CI)	99.5 (99.2,99.7)	99.2 (98.7,99.5)	99.9 (99.7,99.9)	99.7 (99.5,99.8)	98.9 (98.3,99.2)	99.5 (99.1,99.8)
Percent agreement (TN + TP/n) %	99.0	94.9	98.4	95.6	93.7	96.6 (94.3,98.4)
Register-recorded						
Observer coverage %	0.7 (0.4,1.0)	5.9 (5.0,7.0)	–	6.4 (5.9,7.1)	7.1 (6.3,8.0)	4.5 (1.8,8.4)
Register recorded coverage %	0.9 (0.6,1.5)	7.2 (6.1,8.4)	–	5.4 (4.9,6.0)	5.0 (4.3,5.8)	4.3 (2.1,7.1)
Not recorded %	98.9 (98.3,99.3)	92 (90.7,93.1)	–	9.0 (8.3,9.7)	45.4 (43.8,47.1)	66.0 (15.7,99.3)
Not readable %	0.2 (0.1,0.5)	0.8 (0.5,1.3)	–	0.3 (0.2,0.4)	0.2 (0.1,0.5)	0.3 (0.2,0.6)
Sensitivity % (95% CI)	23.5 (6.8,49.9)	16.4 (10.2,24.4)	–	48.4 (43.6,53.4)	12.4 (8.6,17.2)	24.6 (7.2,48.1)
Specificity % (95% CI)	99.2 (98.8,99.6)	93.4 (92.2,94.4)	–	97.6 (97.2,97.9)	95.5 (94.8,96.2)	96.8 (94.3,98.6)
Percent agreement (TN + TP/n) %	98.7	89.0	–	94.4	89.7	93.6 (88.8,97.2)
D2) Non-crying						
Survey-reported						
Observer coverage %	3.6 (2.1,6.2)	12.0 (9.5,15.1)	6.8 (5.7,8.2)	17.4 (15.7,19.2)	17.8 (15.4,20.4)	10.9 (6.1,17.0)
Survey reported coverage %	2.0 (0.9,4.3)	3.8 (2.3,5.9)	1.2 (0.7,2)	2.5 (1.7,3.4)	2.5 (1.5,4.2)	2.3 (1.5,3.3)
"Don't know" responses %	9.7 (6.9,13.4)	1.8 (0.9,3.5)	3.7 (2.8,4.9)	0.7 (0.4,1.3)	8.0 (6.0,10.6)	4.1 (1.5,8.0)
Sensitivity % (95% CI)	9.1 (0.2,41.3)	16.1 (8.0,27.7)	14.7 (7.3,25.4)	10.8 (6.8,16.0)	7.4 (3.0,14.7)	11.6 (8.7,14.8)
Specificity % (95% CI)	98.2 (96.2,99.3)	98 (96.2,99.1)	99.5 (99,99.8)	98.9 (98.1,99.4)	98.4 (96.9,99.3)	98.7 (98.0,99.3)
Percent agreement (TN + TP/n) %	95.4	87.9	95.3	86.8	84.1	90.4 (85.1,94.6)
Register-recorded						
Observer coverage %	3.6 (2.1,6.2)	12 (9.5,15.1)	–	17.4 (15.7,19.2)	17.8 (15.4,20.4)	12.1 (6.9,18.5)
Register recorded coverage %	3.3 (1.6,6.4)	8.5 (6.2,11.5)	–	13 (11.5,14.7)	7.2 (5.7,9.1)	7.9 (4.4,12.2)
Not recorded %	96.7 (93.6,98.4)	90.9 (87.8,93.3)	–	10.1 (8.8,11.7)	47.1 (43.9,50.4)	64.4 (19.2,97.4)
Not readable %	0.0 (0.0,1.7)	0.7 (0.2,2.1)	–	0.3 (0.1,0.7)	0 (0,0.5)	0.2 (0.0,0.6)
Sensitivity % (95% CI)	25.0 (5.5,57.2)	11.1 (3.7,24.1)	–	46.8 (41.1,52.5)	13.5 (8.7,19.7)	23.7 (6.5,47.2)
Specificity % (95% CI)	97.7 (95.1,99.2)	91.8 (88.8,94.3)	–	94.1 (92.8,95.3)	94.1 (92.2,95.7)	94.4 (92.4,96.1)
Percent agreement (TN + TP/n) %	94.5	83.9	–	85.9	80.0	86.3 (81.1,90.8)
D3) Non-breathing						
Survey-reported						
Observer coverage %	3.9 (2.2,6.7)	18.0 (13.8,23)	5.1 (4.1,6.4)	19.2 (16.9,21.7)	15.9 (13.3,18.9)	11.5 (5.5,19.5)
Survey reported coverage %	1.9 (0.8,4.3)	5.8 (3.4,9.6)	1.0 (0.5,1.8)	2.7 (1.7,4.1)	2.3 (1.2,4.3)	2.5 (1.3,4.1)
"Don't know" responses %	10.6 (7.5,14.7)	2.3 (0.9,5.2)	3.2 (2.3,4.4)	0.6 (0.2,1.6)	7.1 (5.0,10.0)	4.1 (1.5,7.9)
Sensitivity % (95% CI)	0.0 (0.0,30.8)	20.8 (10.5,35.0)	16.7 (7.0,31.4)	11.3 (6.2,18.6)	8.2 (2.7,18.1)	12.8 (8.0,18.5)
Specificity % (95% CI)	98.0 (95.7,99.3)	97.6 (94.5,99.2)	99.6 (99.0,99.9)	98.8 (97.7,99.5)	98.7 (96.9,99.6)	98.7 (97.8,99.4)
Percent agreement (TN + TP/n) %	94.9	83.3	96.7	86.1	86.0	90.1 (83.1,95.4)
Register-recorded						
Observer coverage %	3.9 (2.2,6.7)	18.0 (13.8,23.0)	–	19.2 (16.9,21.7)	15.9 (13.3,18.9)	13.6 (7.5,21.1)

Table 3 Individual-level validation in registers and exit surveys of bag-mask-ventilation indicator, EN-BIRTH study (n = 22,752) *(Continued)*

	Bangladesh Azimpur Tertiary	Bangladesh Kushtia District	Nepal Pokhara Regional	Tanzania Temeke Regional	Tanzania Muhimbili National	All sites Pooled (random effects)
Register recorded coverage %	2.9 (1.3,6.1)	8.2 (5.1,12.7)	–	14.9 (12.8,17.3)	7.1 (5.4,9.4)	8.0 (3.7,13.7)
Not recorded %	97.1 (93.9,98.7)	90.5 (85.8,93.8)	–	9.7 (8.0,11.8)	46.6 (42.8,50.4)	64.2 (19.3,97.2)
Not readable %	0.0 (0.0,1.9)	1.3 (0.3,4.0)	–	0.2 (0.0,0.8)	0.0 (0.0,0.7)	0.2 (0.0,0.7)
Sensitivity % (95% CI)	25.0 (5.5,57.2)	11.1 (3.1,26.1)	–	51.3 (44,58.5)	15.7 (9.4,24)	25.6 (7.2,50.2)
Specificity % (95% CI)	98.3 (95.6,99.5)	92.3 (87.7,95.7)	–	93.8 (92,95.4)	94.5 (92.3,96.2)	94.8 (92.5,96.7)
Percent agreement (TN + TP/n) %	94.6	79.7	–	85.6	82.1	85.9 (80.2,90.8)

Full denominator details presented in Additional file 14

blank for 'no'). The Tanzanian register captured resuscitation steps by numerical code in a column headed "Helping Babies Breathe" (suction = 1, stimulation = 2, BMV = 3) or "no", and blanks are treated as not recorded. There was no specific column in the Nepal register for resuscitation.

Documentation practices in registers

Resuscitation practices were recorded in varying order into multiple documents (Additional file 11). Reported time between care and documentation ranged from 2.5 min in Pokhara NP to 22.5 min in Temeke TZ.

Register design Register design largely acted as a barrier to recording in Pokhara NP:

> *"Drying, stimulation, and bag-mask ventilation are written [in the patient's chart], but in the main register it is not present… we do not have routine care of the newborn in the register, only in the patient's chart."*
> -Data collector, Pokhara NP

In the other hospitals health workers duplicated documentation in registers with multiple other documents (e.g. partographs, patient case notes) (Additional file 12).

Register filling Aspects of register filling acted as both barriers and enablers. Training and support from senior nurses enabled improved accuracy of documentation, while limited time acted as a barrier. Health workers across the hospitals discussed the lack of time to document, particularly for complicated cases and resuscitation when they are focused on delivering care:

> *"Just after finishing [resuscitation], you must keep everything clear… time is a problem… you must estimate, there are times it is difficult and other times you ask the [senior nurse]… because in an emergency you all work together; thus, you remind each other."*
> – Health worker, Temeke TZ

Health workers in Pokhara NP received specific support for documentation in neonatal resuscitation:

Bag mask ventilation

		Bangladesh Azimpur Tertiary	Bangladesh Kushtia District	Nepal Pokhara Regional	Tanzania Temeke Regional	Tanzania Muhimbili National	Pooled (random effects)
Denominator D1: All live births and fresh stillbirths	Ratio: Survey: Observed	0.71	0.32	0.14	0.11	0.20	0.23
Denominator D2: Non crying	Ratio: Survey: Observed	0.56	0.32	0.18	0.14	0.14	0.21
Denominator D3: Non breathing	Ratio: Survey: Observed	0.49	0.32	0.20	0.14	0.14	0.21
Denominator D1: All live births and fresh stillbirths	Ratio: Register: Observed	1.29	1.22		0.84	0.70	0.94
Denominator D2: Non crying	Ratio: Register: Observed	0.92	0.71		0.75	0.40	0.63
Denominator D3: Non breathing	Ratio: Register: Observed	0.74	0.46		0.78	0.45	0.58

< 0.80	OR	>1.20	Poor
0.80 to 0.84	OR	1.16 to 1.20	Moderate
0.85 to 0.89	OR	1.11 to 1.15	Good
0.90 to 0.94	OR	1.06 to 1.10	Very Good
0.95 to 0.99	OR	1.00 to 1.05	Excellent
Data not captured in register			

Fig. 7 Validity ratios for exit survey-reported and register-recorded coverage of bag-mask-ventilation, EN-BIRTH study (*n* = 22,752). Full denominator details presented in Additional file 14

Table 4 Individual-level validation in exit survey of crying at birth indicator, EN-BIRTH study (n = 22,752 births)

	Bangladesh Azimpur Tertiary	Bangladesh Kushtia District	Nepal Pokhara Regional	Tanzania Temeke Regional	Tanzania Muhimbili National	All sites Pooled Random effects
Cry at birth - Survey reported – live births + fresh stillbirths						
Observer prevalence %	86.7 (85.3,87.9)	75.7 (73.8,77.4)	77.4 (76.4,78.3)	72.0 (70.9,73)	72.6 (71.1,74.1)	77.1 (72.0,81.8)
Survey reported prevalence %	94.4 (93.4,95.2)	94.3 (93.3,95.2)	95.8 (95.3,96.3)	93.0 (92.3,93.6)	90.5 (89.3,91.6)	93.7 (91.9,95.3)
"Don't know" responses %	0.8 (0.5,1.2)	1.2 (0.8,1.7)	1.3 (1.0,1.6)	0.8 (0.6,1.0)	2.6 (2.0,3.3)	1.3 (0.8,1.8)
Sensitivity % (95% CI)	96.0 (95.1,96.7)	98.5 (97.8,99.0)	97.3 (96.8,97.7)	97.7 (97.2,98.1)	95.4 (94.3,96.3)	97.1 (96.1,97.9)
Specificity % (95% CI)	15.7 (12.2,19.8)	15.2 (12.3,18.5)	8.7 (7.3,10.2)	18.1 (16.2,20.2)	21.4 (18.3,24.8)	15.6 (10.8,21.0)
Percent agreement (TN + TP/n) %	85.3	78.4	77.0	76.6	76.3	78.8 (75.6,81.9)

Full denominator details presented in Additional file 14

"We have received training on HBB and we were trained for documentation in that. We were doing documentation before, but we received direction for improving it."
-Health worker, Pokhara NP

However, while health workers in Pokhara NP record resuscitation in other documents, it is not recorded in routine hospital registers.

Register use While improved patient care and use of data by managers motivated documentation and was

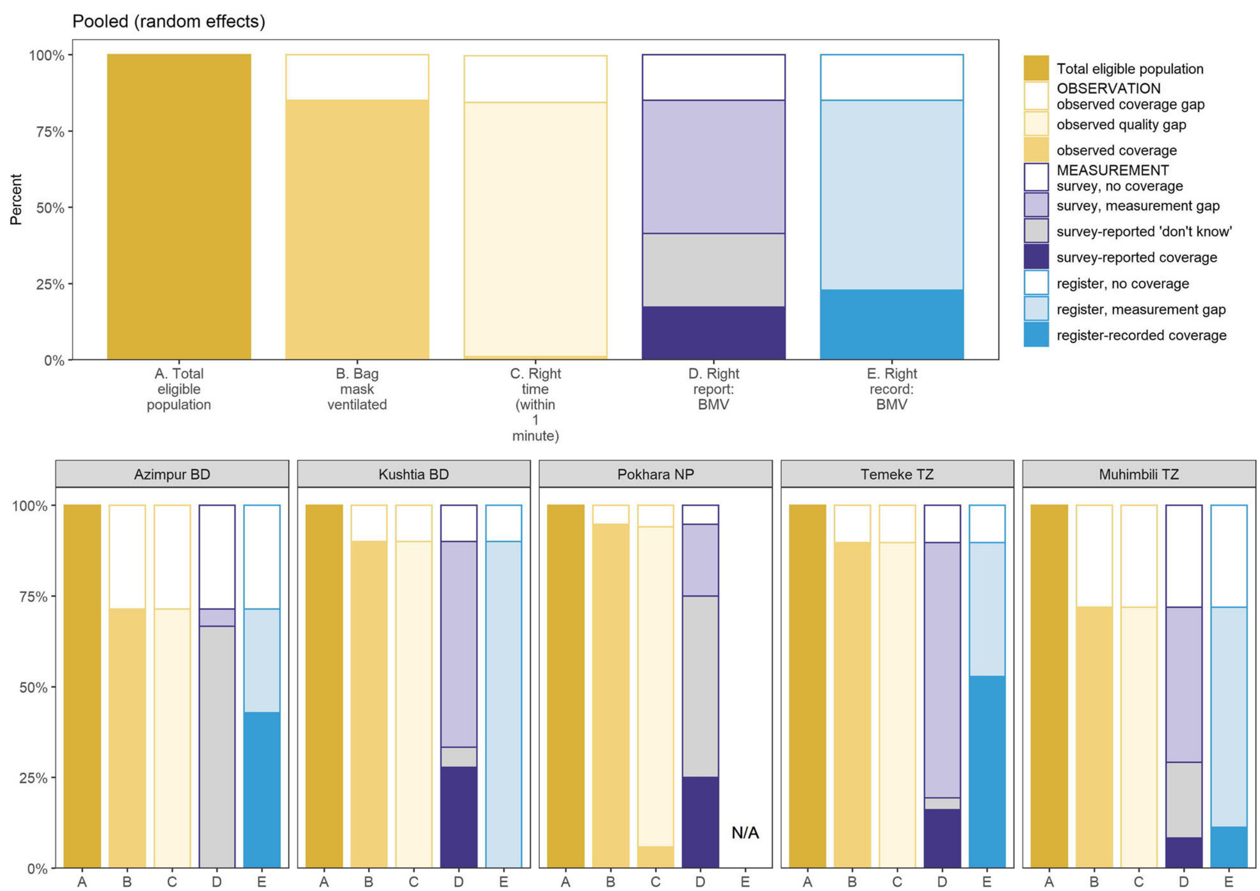


Fig. 8 Gap analysis for coverage and quality among newborns non-crying/not responding to stimulation/suction, EN-BIRTH study (*n* = 200). BD = Bangladesh, NP = Nepal, TZ = Tanzania; BMV = Bag-mask-ventilation; Full denominator details presented in Additional file 14

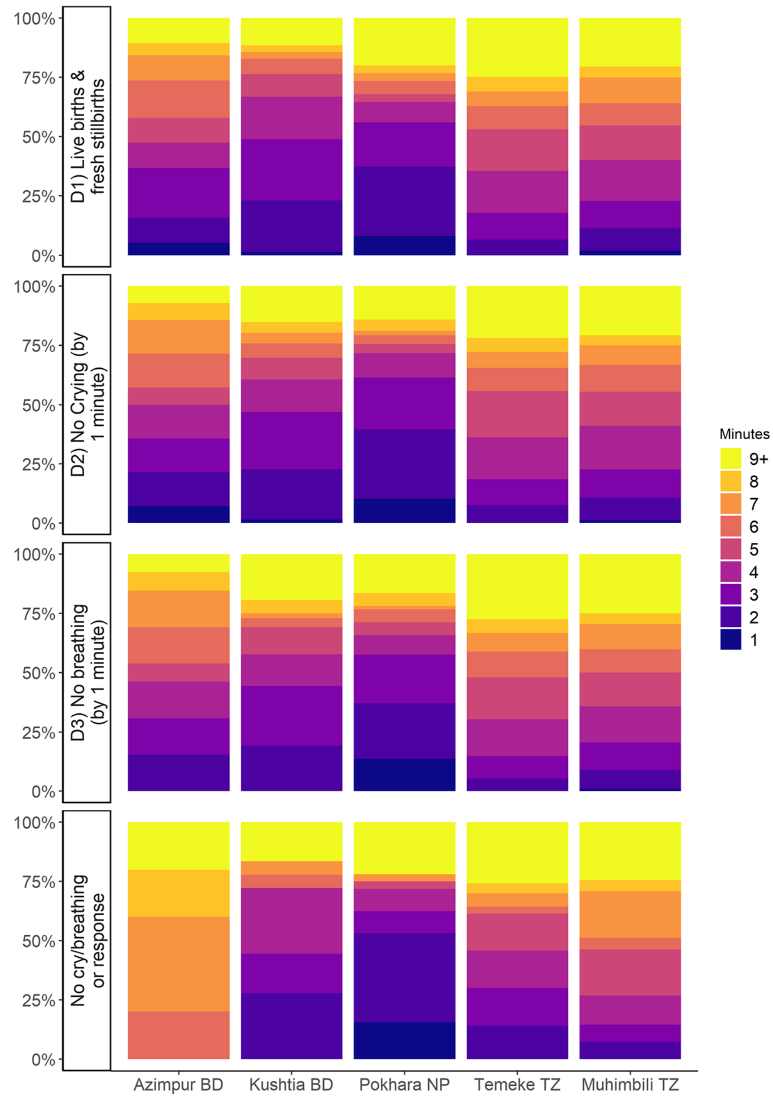


Fig. 9 Time to bag-mask-ventilation by denominator, EN-BIRTH study ($n = 991$). BD = Bangladesh, NP = Nepal, TZ = Tanzania; D1: n = 991, D2: $n = 672$, D3: $n = 454$, No cry/breath/response: $n = 142$

affirmation of the care health workers were giving, not all respondents could identify the use for resuscitation data in routine registers.

Feedback was lacking where documentation didn't line up with clinical need:

"Sometimes when you look at the [APGAR] score of the baby, maybe it's 5, you wonder why they didn't perform resuscitation, there's a possibility they [did] but they haven't documented that... There's no one to follow up on that... The person responsible for data comes and copies what's written in the register, be it a low score... but they never ask them why they didn't perform resuscitation if the baby had a low score"
– Health worker, Temeke TZ

Conversely, in Bangladesh, health workers were not sure what happened with resuscitation data:

"Resuscitation is an emergency subject. There remains a referral slip while resuscitating a baby on emergency that indicates the baby went to operating theatre... We write down the procedures of resuscitation in that slip... I am not sure whether this actually goes in the monthly report or not."
-Health worker, Azimpur BD

Data culture Data culture was both an enabler and barrier to routine documentation of resuscitation. It acted as a barrier where minor interventions were not seen as worth recording:

> *"Minor things like suctioning were not recorded and they only documented on a resuscitation case that took more than ten minutes."*
> -Data collector, Muhimbili TZ

However, the importance of documentation was noted for organizational and personal protection:

> *"For instance, if a child has been born but unfortunately, let us say she had a problem, you have resuscitated her, but you did not document… and the mother/parent has become very angry and start complaining, or the whole management has become angry with you why the child had this situation, but you did not record what you have done ... You will not defend yourself, but documentation defends you."*
> -Health worker, Muhimbili TZ

Discussion

EN-BIRTH study's large sample size (22,752 live births and fresh stillbirths) allowed the first validity assessment of measurement for neonatal resuscitation coverage in routine hospital registers and surveys, against a gold standard of clinical observation. We found that survey report poorly captured resuscitation indicators. Routine labour ward registers performed better, but variably, and have potential, especially with data quality improvement.

Survey-reported coverage was challenging, which is not surprising. We found most women who reported their baby had trouble breathing after birth did not know if their baby had been stimulated or received BMV. We recommend resuscitation need or BMV questions should not be added to existing population-based surveys. Furthermore, the sample size required for this relatively low-incidence practice, would be challenging even in DHS surveys with large, nationally representative samples [29].

The numerator for neonatal resuscitation is key. Stimulation by rubbing the baby's back is easily conflated with the similar action of drying every newborn baby and was not recognized at all by mothers (< 3% in survey report). Suction is only necessary if the airway is blocked and a measurement focus on suction may unintentionally encourage this potentially harmful practice which can cause bradycardia. BMV is the most distinguishable option for a clear subset of non-breathing babies and had higher accuracy than stimulation. Though underestimated in surveys, accuracy of BMV was still performed better than stimulation by survey-report. Additionally, BMV is a more suitable intervention for which to assess quality and links to health facility assessments where standard questions include presence and recent use of neonatal bag and masks.

Health facilities are where ~ 80% of women now deliver [10], providing an opportunity to track neonatal resuscitation coverage through routine facility data using BMV as the numerator. Four of the five routine registers assessed were already capturing BMV count data. At the population level, register-recorded coverage of BMV was within 2.1% of observed coverage although individual-level validation metrics suggested low sensitivity. Selective register design is important in capturing what is needed yet avoiding documentation over-burdening. In Tanzania, the register column labelled "HBB" aligns measurement with scale-up programming. The design in Bangladesh instructed health workers to leave the column blank when BMV is not done; thus, calculating completeness and differentiating between truly 'not done' and register 'incomplete' was impossible. Where register instructions in Tanzania state to write "no" if BMV was not done, completeness was moderate to high (54.6–91.0%). Although data collectors rarely indicated data were not readable (< 0.5%), there were low inter-rater kappa results for register-recorded BMV in some sites [23]. Because extraction/aggregation is the first step for data flowing to higher levels in the health system, more research is needed to improve this. Capturing reliable data depends on user-friendly, appropriate recording systems, however, accuracy varied even within the same country using identical register design, highlighting the importance of information culture and supervision. Our qualitative findings suggest differences in understanding of importance and utility of resuscitation data at different hospitals.

Denominators are notably challenging for interventions such as resuscitation which are indicated based on clinical need for only a subset of babies [30]. Current WHO guidance recommends *number of live births in a facility*, with a footnote that this is pragmatic whilst ongoing work to test different denominators, including EN-BIRTH, is completed [31]. Here we have included live births plus fresh stillbirths, for whom resuscitation is recommended. Any newborn without maceration or major malformations, even if they appear completely lifeless, should be given the chance of resuscitation [32]. The reduction in stillbirth rates associated with resuscitation training [33–35] are likely results of reduced misclassification of live births as stillbirths.

Measuring the true denominator for clinical need for resuscitation is complex. Newborns require BMV if non-breathing/gasping after initial drying/stimulation or if they suffer subsequent apnoea at any time. Breathing well may be difficult to measure as the concept excludes gasping, fast breathing and grunting. It is critical to emphasise these breathing patterns during clinical training as BMV is indicated for some (e.g. gasping) but not all of these breathing patterns. EN-BIRTH observers collected breathing or not breathing as a binary variable because formative research suggested other breathing patterns were not feasible to capture. In our study, 2/5 registers captured non-breathing but as a composite non-crying and non-breathing indicator. Consequently, accuracy of this denominator in registers could not be assessed.

Non-crying has potential utility as a denominator as it is simple for health workers to capture and is part of the process in assessing need for resuscitation. Additionally, crying at birth is a single event and thus more straightforward to record as opposed to breathing which is a process and might change over time, particularly for preterm babies. While not all non-crying babies will require further steps of resuscitation, almost all babies who do need BMV are non-crying. One study has shown babies breathing but not crying after birth have an increased risk of death [36]. We found the observed coverage of BMV ranged from 3.6–17.8% among babies not crying in the first minute. Further research is required to assess if non-crying is useful and benchmarking is feasible. However, as considerations turn towards respectful newborn care and minimal handling, further research is needed related to newborn physiological responses after birth and what is appropriate to measure.

Apgar scores are captured in all the routine hospital registers in our study, including in Pokhara NP, which captured no resuscitation interventions. Apgar scores do not capture interventions around the time of birth, rather describe a newborn's physical condition and response to any interventions at 1 and 5 min after birth and are already known to have limitations, notably low inter-rater reliability. The one-minute Apgar score, which includes heart rate, does not fit well with current resuscitation algorithms which recommend checking the baby's heart rate after a minute of ventilation (2 min after birth). As such, the Apgar score is not a useful denominator for neonatal resuscitation and as usually written in individual patient records, we suggest exploring replacing this column in routine labour ward registers with data elements that can be used for coverage measurement e.g. not crying after birth.

Timely resuscitation is essential and even small delays in starting resuscitation can contribute to death or disability [37]. Our assessment of quality of care focused on timeliness of the start of BMV within the first minute after birth. While coverage of BMV was high (85%), only 1% of newborns received the first ventilation within 1 min of birth. In the all newborns denominator, not all will require BMV within 1 min of birth as many were crying/breathing at birth and subsequently became distressed or apnoeic. A coverage gap for BMV of fresh stillbirths is to be expected as it is not appropriate to resuscitate those babies who are diagnosed before birth to have died in utero e.g. confirmed by ultrasound. Measuring timing of BMV is clearly not feasible in surveys and very unlikely to be possible in routine labour ward registers. Given this major quality gap regarding timing of resuscitation initiation, local audit and special studies are important to drive quality improvement.

Strengths and limitations

Strengths of this study include the multi-site and multi-country design and large sample size enabling the capture of multiple decision points on resuscitation algorithms. We evaluated how several possible numerators/denominators performed using clinical observation as a gold standard. We assessed possible bias in the observation data with double observation for a subset of cases. Overall, BMV had good inter-observer agreement. Whilst clinically trained observers provided gold standard data on coverage of interventions, subjectivity remains possible e.g. differentiating stimulation from immediate drying. To limit this, the tablet application was designed to capture stimulation in a specific neonatal resuscitation section separate from the immediate care practices, such as drying. The low coverage of stimulation amongst non-crying/breathing newborns (34–38%) may reflect poor quality of care or difficulty in measurement for stimulation by an observer.

Some other limitations should be noted. Survey-reported coverage was assessed in exit survey, closer in time to the events in question than standard population-based surveys with 2–5-year reference periods. In survey, only women who answered 'yes' to a question asking whether their baby had difficulty breathing at birth were asked further questions about resuscitation, thus some who may have recognised newborn stimulation were not counted towards survey-reported coverage. Additionally, the EN-BIRTH study sample may be healthier than the average in these facilities (women too sick to consent, women with no fetal heart beat heard at admission, etc., were excluded from the study). As the study sites were CEmONC hospitals, case mix, coverage, and measurement may differ at lower-level facilities.

Importantly, the true denominator of babies in need of BMV will not be captured by facility measurement, especially the disadvantaged who are more likely to deliver at home in LMICs. However, home births are less likely to receive BMV in most LMICs, so facility measurement is likely to capture nearly all the numerator in terms of newborns receiving BMV. Hence approaches such as those used in immunisation when the denominator is missing may help to estimate the coverage of the whole population for contexts with many home births.

Conclusion

Neonatal resuscitation is a high impact evidence-based intervention for a leading cause of under-five mortality, preventable stillbirth and disability. Yet the current lack of coverage measurement is impeding global tracking of scale-up in high-burden countries. We found bag-mask-ventilation was the most reliable numerator. Measuring the true denominator for clinical need is complex and further denominator research is required, including respectful care considerations, evaluating non-crying as a potential alternative. Based on these results, we do not recommend tracking this indicator through population-based survey. Register measurement of neonatal resuscitation has potential and if standardised and included in HMIS, could aid in tracking progress towards global targets across countries. An appropriate resuscitation denominator could potentially replace the Apgar score, which was recorded as a column in all five registers. Implementation research is needed regarding how to improve register data quality. Measuring and addressing quality of care gaps, notably for timely provision of resuscitation in the first minute, is crucial for programme improvement and impact, but unlikely to be feasible in routine systems, requiring audits and special studies. Improving data is possible and necessary, informing progress to meet global goals and meet every family's aspiration that their baby will survive and thrive.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12884-020-03422-9.

Additional file 1. Data collection dates by site for the EN-BIRTH study.

Additional file 2. EN-BIRTH metadata definitions of selected indicator options for validity testing.

Additional file 3. STROBE statement—checklist of items that should be included in reports of observational studies.

Additional file 4. EN-BIRTH data quality assurance for Gold standard – double observation and data entry.

Additional file 5. Validity measures for stimulation (AUC, IF, PPV, NPV) by hospital for EN-BIRTH study.

Additional file 6. EN-BIRTH study – testing validity – stimulation coverage indicator 2 way tables.

Additional file 7. Validity measures for BMV (AUC, IF, PPV, NPV) by hospital and denominator option for EN-BIRTH study.

Additional file 8. EN-BIRTH study – testing validity – BMV coverage indicator 2 way tables.

Additional file 9. Validity measures for survey-reported crying at birth by hospital for EN-BIRTH study.

Additional file 10. EN-BIRTH study – testing validity – cry at birth 2 way tables.

Additional file 11. Recording order of neonatal resuscitation in hospital documents according to EN-BIRTH data collectors.

Additional file 12. Barriers and enablers to routine recording of neonatal resuscitation.

Additional file 13. Ethical approval of local institutional review boards, EN-BIRTH study.

Additional file 14. Numerators and denominators for observed, survey-reported and register-recorded indicators.

Abbreviations

EN-BIRTH: Every Newborn–Birth Indicators Research Tracking in Hospitals; BD: Bangladesh; BMV: Bag-and-mask ventilation; CEmONC: Comprehensive Emergency Obstetric and Neonatal Care; DHS: Demographic and Health Survey Program; HBB: Helping Babies Breathe; HMIS: Health Management Information Systems; icddr,b: International Centre for Diarrheal Disease Research, Bangladesh; IHI: Ifakara Health Institute; LMIC: Low- and middle-income country; LSHTM: London School of Hygiene & Tropical Medicine; MCHTI: Maternal and Child Health Training Institute, Azimpur, Bangladesh; MICS: Multiple Indicator Cluster Survey; MUHAS: Muhimbili University of Health and Allied Sciences; NP: Nepal; TZ: Tanzania; WHO: World Health Organization

Acknowledgements

Firstly, and most importantly, we thank the women, their families, the health workers, and data collectors. We credit the inspiration of the late Godfrey Mbaruku. We thank Claudia DaSilva, Veronica Ulaya, Mohammod Raisul Islam, Sudip Karki and Rabina Sarki for their administrative support and Sabrina Jabeen, Goutom Banik, Md. Shahidul Alam, Tamatun Islam Tanha and Md. Mohsiur Rahman for support during data collectors training.
We acknowledge the following groups for their guidance and support:
National Advisory Groups:
Bangladesh: Mohammod Shahidullah, Khaleda Islam, Md Jahurul Islam.
Nepal: Naresh P KC, Parashu Ram Shrestha.
Tanzania: Muhammad Bakari Kambi, Georgina Msemo, Asia Hussein, Talhiya Yahya, Claud Kumalija, Eliudi Eliakimu, Mary Azayo, Mary Drake, Honest Kimaro.
EN-BIRTH validation collaborative group:
Bangladesh: Md. Ayub Ali, Bilkish Biswas, Rajib Haider, Md. Abu Hasanuzzaman, Md. Amir Hossain, Ishrat Jahan, Rowshan Hosne Jahan, Jasmin Khan, M A Mannan, Tapas Mazumder, Md. Hafizur Rahman, Md. Ziaul Haque Shaikh, Aysha Siddika, Taslima Akter Sumi, Md. Taqbir Us Samad Talha.
Tanzania: Evelyne Assenga, Claudia Hanson, Edward Kija, Rodrick Kisenge, Karim Manji, Fatuma Manzi, Namala Mkopi, Mwifadhi Mrisho, Andrea Pembe.
Nepal: Jagat Jeevan Ghimire, Rejina Gurung, Elisha Joshi, Avinash K Sunny, Naresh P. KC, Nisha Rana, Shree Krishna Shrestha, Dela Singh, Parashu Ram Shrestha, Nishant Thakur.
LSHTM: Hannah Blencowe, Sarah G Moxon.
EN-BIRTH Expert Advisory Group: Agbessi Amouzou, Tariq Azim, Debra Jackson, Theopista John Kabuteni, Matthews Mathai, Jean-Pierre Monet, Allisyn C. Moran, Pavani K. Ram, Barbara Rawlins, Jennifer Requejo, Johan Ivar Sæbø, Florina Serbanescu, Lara Vaz.
We are also very grateful to fellow researchers who peer-reviewed this paper.

About this supplement

This article has been published as part of BMC Pregnancy and Childbirth Volume 21 Supplement 1, 2021: *Every Newborn* BIRTH multi-country validation study: informing measurement of coverage and quality of maternal and newborn care. The full contents of the supplement are available online at https://bmcpregnancychildbirth.biomedcentral.com/articles/supplements/volume-21-supplement-1.

Authors' contributions

The EN-BIRTH study was conceived by JEL, who acquired the funding and led the overall design with support from HR. Each of the three country research teams input to design of data collection tools and review processes, data collection and quality management with technical coordination from HR, GRGL, and DB. The icddr,b team (notably AER, TT, TH, QSR, SA, and SBZ) led the development of the software application, data dashboards, and database development with VG and the LSHTM team. IHI (notably DS) coordinated work on barriers and enablers for data collection and use, working closely with LTD. QSR was the main lead for data management working closely with OB, KS and LTD. For this paper, KP led the analyses and first draft of the manuscript with AKC, working closely with LTD and JEL. Authors made substantial contributions to the conception, design, data collection or analysis or interpretation of data for the work including: icddr,b Bangladesh: SA, QSR, SBZ, SEA. Golden Community, Nepal: AKC, OB, AKS, HM. Ifakara Health Institute, Tanzania: MM, JS, EK, GM. LSHTM: KP, LTD, SK, HR, JEL, as well as others: RCB, EA, NPKC, NS, SN, OL. All authors revised the manuscript and gave final approval of the version to be published and agree to be accountable for the work. The EN-BIRTH study group authors made contributions to the conception, design, data collection or analysis, or interpretation of data. This paper is published with permission from the Directors of Ifakara Health Institute, Muhimbili University of Health and Allied Sciences, icddr,b, and Golden Community. The authors' views are their own,

and not necessarily from any of the institutions they represent, including WHO.
EN-BIRTH Study Group.
Bangladesh: Qazi Sadeq-ur Rahman, Ahmed Ehsanur Rahman, Tazeen Tahsina, Sojib Bin Zaman, Shafiqul Ameen, Tanvir Hossain, Abu Bakkar Siddique, Aniqa Tasnim Hossain, Tapas Mazumder, Jasmin Khan, Taqbir Us Samad Talha, Rajib Haider, Md. Hafizur Rahman, Anisuddin Ahmed, Shams El Arifeen.
Nepal: Omkar Basnet, Avinash K Sunny, Nishant Thakur, Rejina Gurung, Anjani Kumar Jha, Bijay Jha, Ram Chandra Bastola, Rajendra Paudel, Asmita Paudel, Ashish KC.
Tanzania: Nahya Salim, Donat Shamba, Josephine Shabani, Kizito Shirima, Menna Narcis Tarimo, Godfrey Mbaruku (deceased), Honorati Masanja.
LSHTM: Louise T Day, Harriet Ruysen, Kimberly Peven, Vladimir Sergeevich Gordeev, Georgia R Gore-Langton, Dorothy Boggs, Stefanie Kong, Angela Baschieri, Simon Cousens, Joy E Lawn.

Funding
The Children's Investment Fund Foundation (CIFF) are the main funder of the EN-BIRTH Study and funding is administered via The London School of Hygiene & Tropical Medicine. The Swedish Research Council specifically funded the Nepal site through Lifeline Nepal and Golden Community. We acknowledge the core funders for all the partner institutions. Publication of this manuscript has been funded by CIFF. CIFF attended the study design workshop but had no role in data collection, analysis, data interpretation, report writing or decision to submit for publication. The corresponding author had full access to study data and final responsibility for publication submission decision.

Availability of data and materials
The datasets generated during and/or analysed during the current study are available on LSHTM Data Compass repository, https://datacompass.lshtm.ac.uk/955/.

Ethics approval and consent to participate
This study was granted ethical approval by institutional review boards in all operating counties in addition to the London School of Hygiene & Tropical Medicine (Additional file 13). Voluntary informed written consent was obtained from all observed participants, their families for newborns, and respondents for the qualitative interviews. Participants were assured of anonymity and confidentiality. All women were provided with a description of the study procedures in their preferred language at admission, and offered the right to refuse, or withdraw consent at any time during the study. Facility staff were identified before data collection began and no health worker refused to be observed whilst providing care. EN-BIRTH is study number 4833, registered at https://www.researchregistry.com.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]International Maternal and Child Health, Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden. [2]Maternal, Adolescent, Reproductive & Child Health (MARCH), London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK. [3]Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, King's College London, London, UK. [4]Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh. [5]Department of Health Systems, Impact Evaluation and Policy, Ifakara Health Institute, Dar es Salaam, Tanzania. [6]Ministry of Health and Social Welfare, Dar es Salaam, Tanzania. [7]Research Division, Golden Community, Lalitpur, Nepal. [8]Pokhara Academy of Health Sciences, Pokhara, Nepal. [9]Ministry of Health and Population, Kathmandu, Nepal. [10]Muhimbili University of Health and Allied Sciences (MUHAS), Dar Es Salaam, Tanzania. [11]Society of Public Health Physicians Nepal, Kathmandu, Nepal. [12]Department of Paediatrics, University of Calgary, Calgary, Canada. [13]University of Colorado School of Medicine, Colorado School of Public Health, Aurora, CO, USA. [14]Department of Maternal, Newborn, Child and Adolescent Health and Ageing, WHO, Geneva, Switzerland.

Published: 26 March 2021

References

1. Lee AC, Cousens S, Wall SN, Niermeyer S, Darmstadt GL, Carlo WA, et al. Neonatal resuscitation and immediate newborn assessment and stimulation for the prevention of neonatal deaths: a systematic review, meta-analysis and Delphi estimation of mortality effect. BMC Public Health. 2011;11:S12.
2. Lee ACC, Kozuki N, Blencowe H, Vos T, Bahalim A, Darmstadt GL, et al. Intrapartum-related neonatal encephalopathy incidence and impairment at regional and global levels for 2010 with trends from 1990. Pediatr Res. 2013;74(Suppl 1):50–72.
3. Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, et al. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the sustainable development goals. Lancet. 2016;388:3027–35.
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.
5. United Nations Inter-agency Group for Child Mortality Estimation (UN IGME). A Neglected Tragedy: The global burden of stillbirths. New York: United Nations Children's Fund; 2020. https://data.unicef.org/resources/a-neglectedtragedy-stillbirth-estimates-report/. Accessed 12 Oct 2020.
6. Gurung R, Litorp H, Berkelhamer S, Zhou H, Tinkari BS, Paudel P, et al. The burden of misclassification of antepartum stillbirth in Nepal. BMJ Glob Health. 2019;4:e001936.
7. KC A, Berkelhamer S, Gurung R, Hong Z, Wang H, Sunny AK, et al. The burden of and factors associated with misclassification of intrapartum stillbirth: Evidence from a large scale multicentric observational study. Acta Obstet Gynecol Scand. 2019;aogs:13746.
8. Bhutta ZA, Das JK, Bahl R, Lawn JE, Salam RA, Paul VK, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? Lancet. 2014;384:347–70.
9. World Health Organization. Every newborn: an action plan to end preventable deaths. Geneva: World Health Organization; 2014. https://www.who.int/maternal_child_adolescent/documents/every-newborn-action-plan/en/. Accessed 16 Jan 2020.
10. UNICEF. The State of the World's Children. Children, food and nutrition: growing well in a changing world. New York: UNICEF; 2019. https://www.unicef.org/reports/state-of-worlds-children-2019. Accessed 14 Jan 2020.
11. Budhathoki SS, Gurung R, Ewald U, Thapa J, KC A. Does the Helping Babies Breathe Programme impact on neonatal resuscitation care practices? Results from systematic review and meta-analysis. Acta Paediatr. 2019;108:806–13.
12. Niermeyer S. Global gains after Helping Babies Breathe. Acta Paediatr. 2017;106:1550–1.
13. Gurung R, Jha AK, Pyakurel S, Gurung A, Litorp H, Wrammert J, et al. Scaling up safer birth bundle through quality improvement in Nepal (SUSTAIN)—a stepped wedge cluster randomized controlled trial in public hospitals. Implement Sci. 2019;14:65.
14. Enweronu-Laryea C, Dickson KE, Moxon SG, Simen-Kapeu A, Nyange C, Niermeyer S, et al. Basic newborn care and neonatal resuscitation: a multi-country analysis of health system bottlenecks and potential solutions. BMC Pregnancy Childbirth. 2015;15:S4.
15. Perlman JM, Wyllie J, Kattwinkel J, Wyckoff MH, Aziz K, Guinsburg R, et al. Part 7: neonatal resuscitation: 2015 international consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations (reprint). Pediatrics. 2015;136(Suppl 2):S120–66.
16. World Health Organization. Guidelines on basic newborn resuscitation. Geneva: WHO; 2012. https://www.who.int/maternal_child_adolescent/documents/basic_newborn_resuscitation/en/. Accessed 16 Jun 2020.
17. Kamath-Rayne BD, Thukral A, Visick MK, Schoen E, Amick E, Deorari A, et al. Helping Babies Breathe, second edition: A model for strengthening educational programs to increase global newborn survival. Glob Health Sci Pract. 2018;6:538–51.
18. Day LT, Gore-Langton GR, Rahman AE, Basnet O, Shabani J, Tahsina T, et al. Labour and delivery ward register data availability, quality, and utility - every newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries. BMC Health Serv Res. 2020;20:737.
19. Nolen LB, Braveman P, Dachs JNW, Delgado I, Gakidou E, Moser K, et al. Strengthening health information systems to address health equity challenges. Bull World Health Organ. 2005;83:597–603.

20. Maternal and Child Survival Program. Health Management Information Systems Review - Survey on Data Availability in Electronic Systems for Maternal and Newborn Health Indicators in 24 USAID Priority Countries. 2016. https://www.mcsprogram.org/wp-content/uploads/2016/09/Health-Management-Information-Systems-Review.pdf. Accessed 16 Jun 2020.
21. MEASURE Evaluation. Barriers to use of health data in low- and middle-income countries. A review of the literature. North Carolina: Carolina Population Center, University of North Carolina at Chapel Hill; 2018.
22. Day LT, Ruysen H, Gordeev VS, Gore-Langton GR, Boggs D, Cousens S, et al. "Every newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. J Global Health. 2019;9:010902.
23. Day LT, Rahman QS, Rahman AE, Salim N, KC A, Ruysen H, et al. Assessment of the validity of the measurement of newborn and maternal health-care coverage in hospitals (EN-BIRTH): an observational study. Lancet Global Health. 2020. https://doi.org/10.1016/S2214-109X(20)30504-0b.
24. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for statistical Computing; 2018. http://www.R-project.org/. Accessed 10 Jul 2018.
25. World Health Organization. Data quality review: module 2: desk review of data quality. Geneva: World Health Organization; 2017. https://apps.who.int/iris/handle/10665/259225. Accessed 7 Jan 2020.
26. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009;24:217–28.
27. Shamba D, Day LT, Zaman SB, Sunny AK, Tarimo MN, Peven K, et al. Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03517-3.
28. Peven K, Day LT, Ruysen H, Tahsina T, KC A, Shabani J. Stillbirths including intrapartum timing: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03238-7.
29. Ameen S, Siddique AB, Rahman AE, KC A, Day LT, Peven K, et al. Survey of women's report for 33 maternal & newborn indicators: EN-BIRTH multi-country validation study. BMC Pregnancy Childbirth. 2021. https://doi.org/10.1186/s12884-020-03425-6.
30. 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:S8.
31. World Health Organization. Analysis and use of health facility data: guidance for RMNCAH programme managers. Geneva: WHO; 2019. https://www.who.int/publications/m/item/analysis-and-use-of-health-facility-data-guidance-for-rmncah-programme-managers. Accessed 7 Jan 2020.
32. Wall SN, Lee AC, Niermeyer S, English M, Keenan WJ, Carlo W, et al. Neonatal resuscitation in low-resource settings: what, who, and how to overcome challenges to scale up? Int J Gynaecol Obstet. 2009; 107(Suppl 1):S47–64.
33. KC A, Wrammert J, Clark RB, Ewald U, Vitrakoti R, Chaudhary P, et al. Reducing perinatal mortality in Nepal using Helping Babies Breathe. Pediatrics. 2016;137. https://doi.org/10.1542/peds.2015-0117.
34. Msemo G, Massawe A, Mmbando D, Rusibamayila N, Manji K, Kidanto HL, et al. Newborn mortality and fresh stillbirth rates in Tanzania after Helping Babies Breathe training. Pediatrics. 2013. https://doi.org/10.1542/peds.2012-1795.
35. KC A, Ewald U, Basnet O, Gurung A, Pyakuryal SN, Jha BK, et al. Effect of a scaled-up neonatal resuscitation quality improvement package on intrapartum-related mortality in Nepal: A stepped-wedge cluster randomized controlled trial. PLoS Med. 2019;16:e1002900.
36. KC A, Lawn JE, Zhou H, Ewald U, Gurung R, Gurung A, et al. Not crying after birth as a predictor of not breathing. Pediatrics. 2020;145. https://doi.org/10.1542/peds.2019-2719.
37. Lawn JE, Kinney M, Lee AC, Chopra M, Donnay F, Paul VK, et al. Reducing intrapartum-related deaths and disability: Can the health system deliver? Int J Gynecol Obstet. 2009;107(Supplement):S123–42.

Publisher's Note

Chapter 9 - General Discussion and Conclusion

9.1 Introduction

This chapter summarises the overall aim of this thesis and main findings of the six research chapters in section 9.2. The thesis integrated discussion is in section 9.3, overall strengths and limitations of in section 9.4 and disciplinary implications explored in section 9.5 and final conclusion in section 9.6.[87]

9.2 Summary of the overall aim of the thesis and main findings from research manuscripts

The overall aim of this PhD thesis is to explore hospital labour and delivery routine register data quality and opportunities to improve measurement of newborn indicators in high mortality settings. The research was conducted in five hospitals in three countries – Bangladesh, Nepal and Tanzania as part of the EN-BIRTH study.

The main findings of my PhD are divided into six objectives arranged by two sections:

Section A: Assessing existing labour and delivery routine register data quality for hospital births

Three objectives contribute to section A:

Objective 1 – To describe the protocol for the 'Every Newborn-Birth Indicators Research Tracking in Hospitals' (EN-BIRTH) observational study for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania.[1]
This manuscript provided the rationale for the EN-BIRTH study design aligning with the Every Newborn Action Plan (ENAP) prioritised indicators and measurement improvement roadmap.

Objective 2 – To assess labour and delivery ward baseline routine register data availability, quality, and utility in five EN-BIRTH study hospitals.[24]
This retrospective analysis found data elements used to calculate output, outcome, and impact indicators were more widely available than indicators for intervention coverage. Available data elements were mostly highly complete and plausible, but data quality was mixed e.g., marked birthweight heaping in all hospitals. Data heaping is a measure of the proportion of values that are rounded and fall on specific values: 2000g, 2500g, 3000g and so on.

Objective 3 – To assess validity of coverage indicator measurement of newborn and maternal health care coverage in EN-BIRTH study hospitals.[25]
Observation was used as the gold standard against which register records and women's reports were compared prospectively. Routine register data elements were highly complete, but measures of accuracy varied widely by indicator and between hospitals, even with the same register design. Mode of birth affected both clinical care provided and measures of validity.

Section B: Identifying Opportunities to improve labour and delivery routine register data quality for hospital births

Three objectives contribute to section B:

Objective 4 – To explore barriers and enablers for health professionals to record high quality data for newborn and maternal health indicator measurement from labour ward routine registers in five EN-BIRTH study hospitals.[26]
This manuscript used mixed methods to explore barriers and enablers for health professionals to

record high quality data in the labour and delivery routine registers that were assessed for validity of indicator measurement in objective 3. Cross-cutting common themes identified included: register design (complexity, standardisation and paper/electronic systems), register filling (health professional responsibility, time available, training, logistics, supervision and motivation) and register use (data demand, feedback and trust in quality). Depending on the hospital setting, resources and management, these themes acted as either barriers or enablers.

The final two objectives explored opportunities for two specific clinical interventions:

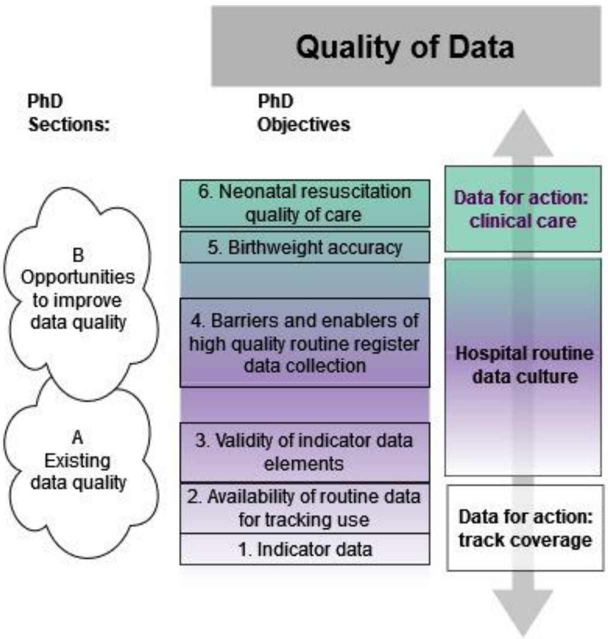
Objective 5 – To assess routine birthweight in EN-BIRTH study hospitals: accuracy, gaps and opportunities to measure coverage and quality of care.[27]
Birthweight measurement findings across all five hospitals in this analysis found most babies were weighed according to recommendations within one hour of birth. The register documentation of birthweight was accurate compared to observation and measurement for low birthweight newborns outperformed non-low birthweight newborns. Opportunities were identified to improve data quality, including reducing birthweight heaping by using digital instead of analogue scales. We found diurnal variability with higher birthweight heaping for night births compared to day births, indicating another opportunity for action.

Objective 6 – To assess measurement opportunities for neonatal resuscitation: indicator definitions and quality of care.[88]
This manuscript explored the complexity of indicator measurement for the multi-step algorithm of neonatal resuscitation by comparing data extracted from the diverse register designs found in the five EN-BIRTH hospitals with subsets of the observational data. Numerator accuracy was higher for bag-mask ventilation compared to stimulation and suction. Denominator options tested highlighted the challenges for routine indicator measurement for this subset target group of newborns needing resuscitation. Recommendations were made for further work to determine whether the denominator of total births is more practical to measure in lieu of the true denominator of non-breathing babies.

9.3 Integrated discussion

This integrated discussion will advance the conceptual framework relating data quality with quality of care introduced in the background chapter 2 (Figure 8). In this section I will propose how findings from my six research papers contribute to a new **data quality continuum**, by connecting the two different uses of data for action: for clinical care and to track coverage (Figure 6) using the hospital routine data culture which determines the data quality for both. This data quality continuum runs parallel and interacts with the WHO quality of care framework to form the **'Quality of Care and Quality of Data conceptual framework'** which will be described later in the thesis in section 9.3.5.2.[5]

Figure 11: Data quality continuum – data for action both for clinical care and to track coverage depend on a hospital routine data culture

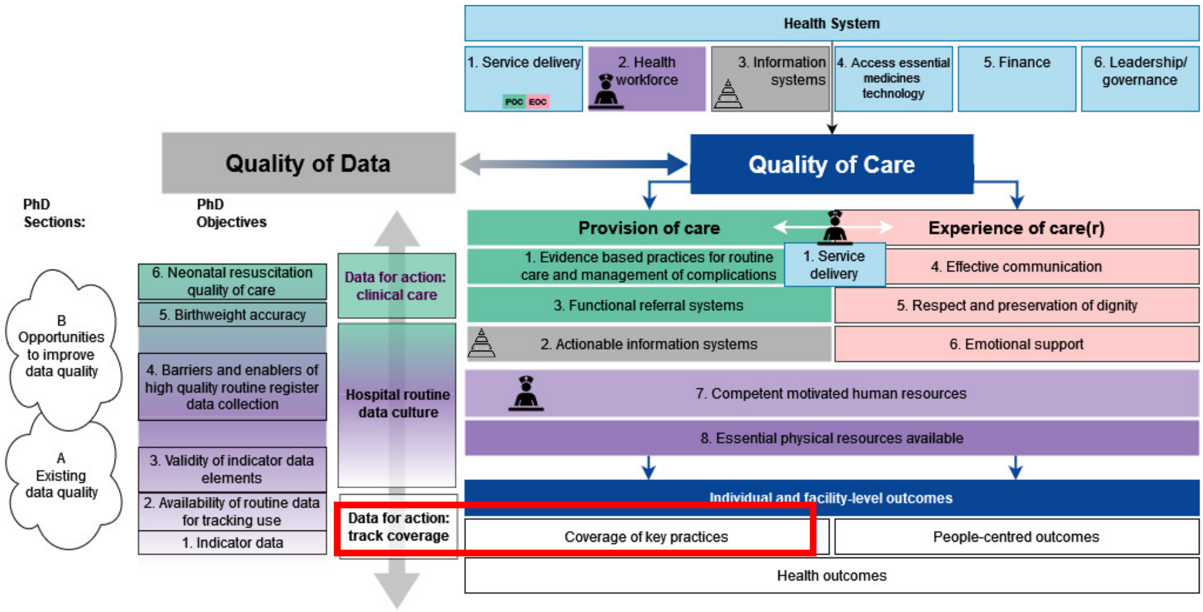


9.3.1 Hospital data for action – what was found and why?

9.3.1.1 Data for action: Track coverage

My PhD has explored the **quality of data for action to track coverage** of core neonatal indicators captured in routine hospital labour and delivery registers. These indicators track *coverage of key practices,* linking to the individual- and facility-level outcomes in the WHO quality of care framework as highlighted in red in Figure 12. Core newborn indicators are designated because they meet "good indicator" criteria: they are action-focused, important, measurable (operational and feasible), simple and valued.[48] But indicator measurement costs time and money, thus only credible data are worth tracking. Data will only be used for action by programme managers and policy makers if it is trusted, and poor data quality hinders data-informed decision-making.[48,89]

Figure 12: 'Quality of Care and Quality of Data conceptual framework' - Data for action: Track coverage



With increasing proportion of births taking place in health facilities, routine data has potential for use at higher levels of the routine health information system. This PhD thesis includes an assessment of the measurement of ENAP core and additional indicators in hospital labour delivery wards. This was a priority identified in the newborn measurement improvement roadmap.[50,51]

Assessment of validity of indicator measurement

The design workshop for facility-based testing of coverage metrics (later called EN-BIRTH study) was held with 64 participants in April 2016, before I joined the research team.[90] The EN-BIRTH design adopted a criterion validation approach similar to studies conducted by the "Improving coverage measurement" researchers.[91] Based on diagnostic validation methodologies, a suite of studies had been conducted to validate measurement of maternal and newborn interventions measured by population-based studies e.g., DHS and MICS. Women's reports were compared against a gold standard (observation or record) using cut-offs to determine 'valid' measurement and showed poor accuracy of women's reports for clinical interventions soon after birth. The EN-BIRTH study was designed to use similar methods to assess the accuracy of data recorded in routine registers and reported by postnatal women in comparison to observation of the clinical intervention/ practice as the gold standard (PhD Thesis Chapter 3, Objective 1).[1]

The baseline retrospective assessment of data availability in routine labour and delivery ward registers in the EN-BIRTH five study hospitals (n=20,075 births) showed many data were available although only the Tanzania design included the coverage indicators of interest (PhD Thesis Chapter 4, Objective 2).[24] Implausible data values were low: for birthweight 0-1.2%, gestational age 0-0.2% % and woman's age 0-0.2%.

Despite register data completeness generally being very high, one or more data elements were much less complete in four of the five hospitals. I explored the register design and found higher completeness for columns "specific" for the data element compared to "non-specific".

The prospective assessment of the validity of measurement of the coverage indicators was the next step in assessing the routine register data quality (PhD Thesis Chapter 5, Objective 3).[25] The EN-BIRTH study protocol paper (PhD Thesis Chapter 3, objective 1) defines the terms: validation as the process whereby the ability of health indicators to measure what they are supposed to measure is determined; accuracy as the closeness of a measured value to a standard value.[1] Whilst the EN-BIRTH analysis was ongoing, Benova et al published: "What gets measured better gets done better": The landscape of validation of global maternal and newborn health indicators through key informant interviews". This landmark paper describes different conceptualisations of indicator validity based around four measurement concepts among which the EN-BIRTH study was assessing the measurement concept Table 2.[92]

Table 2: Three most common definitions of indicator validity[92]

Indicator Concepts	**Definition**
Meaning	Conceptual clarity of intent and construct – what do we want the indicator to measure?
Meaningfulness	Usefulness and use of indicator – what can be achieved through its use?
Measurement	Method of obtaining an estimate – what measure is good enough?
Measurability	Feasibility, cost, acceptability – is it practical to measure this indicator?

The validity methodology agreed by the EN-BIRTH study team adapted the published guidance from previous validation research of indicator measurement from surveys for calculations of validity (sensitivity, specificity and percent agreement).[16] However, under the statistical guidance of Professor Simon Cousens, the EN-BIRTH study team decided not to use the recommended discreet area-under-the-curve (AUC) cut-offs designed for continuous variable measurement. The EN-BIRTH team agreed cut-offs were not appropriate in the context of capturing binary measures of an intervention done/not done compared to documented/not documented.

Without AUC "cut-offs", different options were needed to communicate the EN-BIRTH study results. Although mode of delivery was not originally considered as a stratifier in the protocol design, I drew on my experience of running a hospital HIS in a similar setting to the EN-BIRTH study hospitals and proposed we explore how accuracy varied by mode of delivery.[23]

Individual validation metrics of sensitivity and specificity were valued by some academic and programme audiences, (Objective 3, Table 2[25]) others found the figure I developed depicting coverage measured by observation, register and survey, including confidence intervals, more helpful (Objective 3, Figure 2[25]). I also designed indicator 'validity ratios' of measured: observed coverage

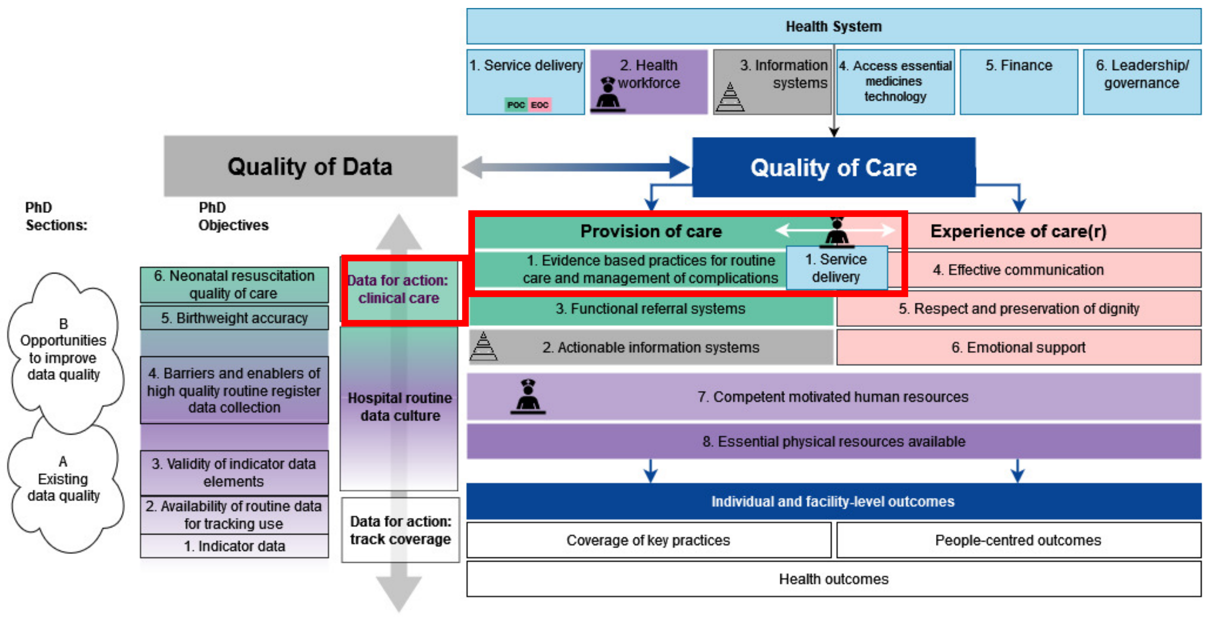
(Objective 3, Figure 2[25]). These are based on 'verification ratios' in the standard data quality assessment methodology.[93] The validity ratios enabled us to categorise accuracy: ratios > 1 indicating overestimation of register-recorded coverage compared to observed, while ratios < 1 indicated underestimation. These categories facilitated heat-mapping in a summary figure which resonated with many audiences. Among the indicators assessed for validity of measurement in routine labour and delivery registers, only two met validity ratio cut offs of good or more: birthweight (1.0) and bag-mask ventilation (0.85). Both uterotonics (0.61) and breastfeeding (4.29) were categorised as poor.

9.3.1.2 Data for action: Clinical Care

In this integrated discussion of my PhD, I will explore **quality of data for action for clinical care** and how this interacts with data for action to track coverage.

The clinical information captured by health workers regarding the care provided is highlighted in red in Figure 13 at the top of the data quality continuum and connecting to quality domain 1, evidence-based practices. In this next section I will discuss the quality of data to enable this process of care.

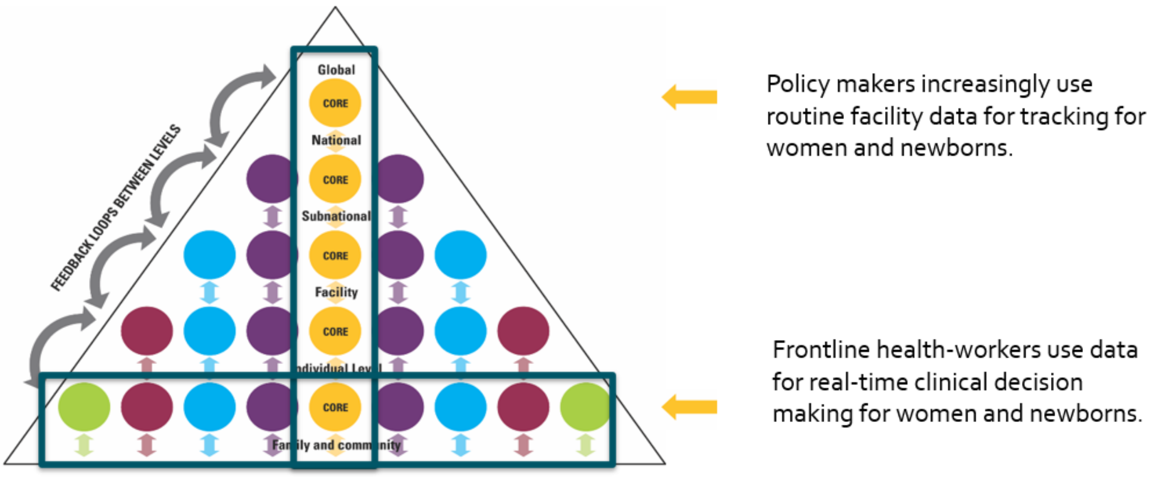
Figure 13: 'Quality of Care and Quality of Data conceptual framework' – Data for action: clinical care



Health professionals use clinical information to provide care to women and newborns. Clinical history, examination and investigation are data points combined minute by minute to guide clinical assessment and action. These clinical information or data are documented in medical records, typically case notes, as a record of care provided, and read by the clinical team providing ongoing care. These data are also used for health facility quality improvement processes including clinical audit.

Core indicators such as the prioritised ENAP indicators track a small subset of all the care provided for newborns in health facilities, selected to pass up the data-information pyramid. These health facility data used to measure core newborn indicators are the same data used for action for clinical care (Figure 14).

Figure 14: Relationship of data reported up the data pyramid to clinical data used by health professionals.



Two clinical practices captured in routine hospital labour and delivery registers will be used to illustrate how the data to track coverage is the same of data used for clinical care: measurement of birthweight (PhD thesis Chapter 7, Objective 5) and neonatal resuscitation (PhD thesis Chapter 8, Objective 6).

Measurement of birthweight

Measurement of birthweight is an evidence-based practice for all newborns. Birthweight measurement is important for clinical care and the low birthweight rate is tracked as a global nutrition goal. Among all the indicators assessed during the EN-BIRTH study for the validity of measurement, birthweight measurement was a high performer, a consistent finding across all five hospitals.[94]

In the prospective register review (PhD Chapter 7, Objective 5) routine labour and delivery register recording of birthweight only slightly underestimated observed birthweight by 0.7% for low birthweight newborns and by 2.5% for non-low birthweight newborns.

The observed clinical practice of weighing in these hospitals was mostly performed within the recommended one hour after birth. However, substantial birthweight heaping was found including in the observational dataset which reflects the quality of clinical birthweight measurement. The type of scale used was associated with different levels of heaping: 36.0% of weighs were heaped for analogue scales compared to 15.7% for digital scales.

The retrospective register assessment conducted in 2016-17 (PhD Thesis Chapter 4, Objective 2) showed substantial birthweight heaping. In the prospective study conducted in 2017-2018 (PhD Thesis Chapter 5, Objective 3) compared to observation, the birthweight extracted from the routine register further increased this birthweight heaping. Insights from the qualitative interviews suggested that health professionals, sometimes noted the birthweight on a scrap of paper, or the mother was asked to remember the birthweight prior to documentation in the register.

Measurement of neonatal resuscitation

In contrast to the single measurement of birthweight, the measurement of neonatal resuscitation (PhD Thesis Chapter 8, Objective 6) demonstrates the complexity of measuring a process of care that begins when clinically indicated which could be immediately after birth or minutes or hours later.

I based the design of the EN-BIRTH neonatal resuscitation analysis as closely as possible to the 'Helping Babies Breathe' (HBB) clinical algorithm to align measurement with clinical care. The HBB programme began to be scaled up over ten years ago in high mortality settings.[56] Available observational data were used to assess numerator and denominator options which represented some of the main resuscitation algorithm decision points. Among two numerator options tested, bag-mask ventilation outperformed stimulation. However, the three denominator options (all newborns excluding macerated stillbirths, newborns not-crying within one minute or not-breathing within one minute) showed mixed results.

Quality gap analyses

The gap analyses for these two clinical interventions, birthweight and neonatal resuscitation measurement were notably different. Birthweight measurement had a small coverage gap, (PhD Chapter 7, Objective 5, Figure 5, column B[27]) and was done at the "right time" (Figure 5, column C[27]) but with a large gap for "right device": digital scales (Figure 5, column D[27]). By comparison the neonatal resuscitation showed a large coverage gap (PhD Chapter 8, Objective 6, Figure 8, column B[28]) and the very concerning finding that almost no newborns received bag-mask ventilation as per algorithm within 1 minute (Figure 8, column C[28]).

The EN-BIRTH gap analyses are similar to effective coverage cascades, a concept that has continued to evolve during the EN-BIRTH study. In 2020, WHO defined effective coverage as the proportion of a population in need of a service that had a positive health outcome from the service.[95] Since this publication of this definition, the observation columns in the EN-BIRTH study gap analyses can be considered to show intervention-coverage and quality-adjusted coverage within a health facility population.[95] In comparison to the proposed standardised cascade for measuring effective coverage, the EN-BIRTH gap analyses figures also show measurement columns for register record and survey-report of the intervention.[95] For neonatal resuscitation, the register-record highlights the large gap between "crude" coverage measurement with observation of any bag-mask ventilation. Quality-adjusted coverage of bag-mask ventilation within the recommended 1 minute of birth was observed for only <1% of babies in the hospitals assessed (Figure 8, column C[28]). The cascade shows how register-record underestimates the "crude coverage" and overestimates the quality-adjusted coverage.

Purpose of routine registers and case notes

Health professionals document clinical care data (document history, examination, investigations requests and results, clinical impression, problem lists, diagnoses, and outcomes) in individual patient case notes. The case notes contain data and information for each admission episode, forming a source of longitudinal data. In high-income settings this is increasingly achieved with electronic medical records, contributing to "big data" that are used for research. In high mortality settings, paper-based individual case notes are still the norm. I will explore case notes data more in the later section: opportunities to improve the quality of data.

My PhD explores data quality in the routine register, a large paper-based medical record book located in the labour and delivery wards of the five study hospitals. Routine registers are a parallel and duplicate subset of the data that overlap with the clinical information in the neonatal case notes. The register data exists for monthly aggregation to send summary reports for use by managers at the facility or higher up the data-information pyramid. Health professionals may occasionally check a data point in the register if, for example, the case notes details were illegible. But typically, once written in the register, the data would not be used again by the health professional during the newborn's admission.

High quality data for clinical care are vital as neonatal clinical care requires time-sensitive accuracy. Birthweight is used to calculate neonatal intravenous fluid volumes and drug doses, thus an inaccurate birthweight becomes an issue for patient safety and especially for low (< 2500g), very low (< 1500g), and extremely low birthweight (< 1000g) neonates. Although the EN-BIRTH study did not assess individual neonatal case notes, if we assume the birthweight documented in the register is the same as that documented in the case notes, the high rates of birthweight heaping we found is concerning (PhD Thesis Chapter 4, Objective 2 and PhD Thesis Chapter 7, Objective 5).

Neonatal resuscitation is a critical clinical intervention for all babies who do not breathe after birth, including for babies with the final outcome of stillbirth. Improving measurement for stillbirths is an important focus within ENAP strategic objective 5. Indeed, the use of the word "newborn" in the Every Newborn Action Plan (ENAP) has a dual meaning – for live births and stillbirths.

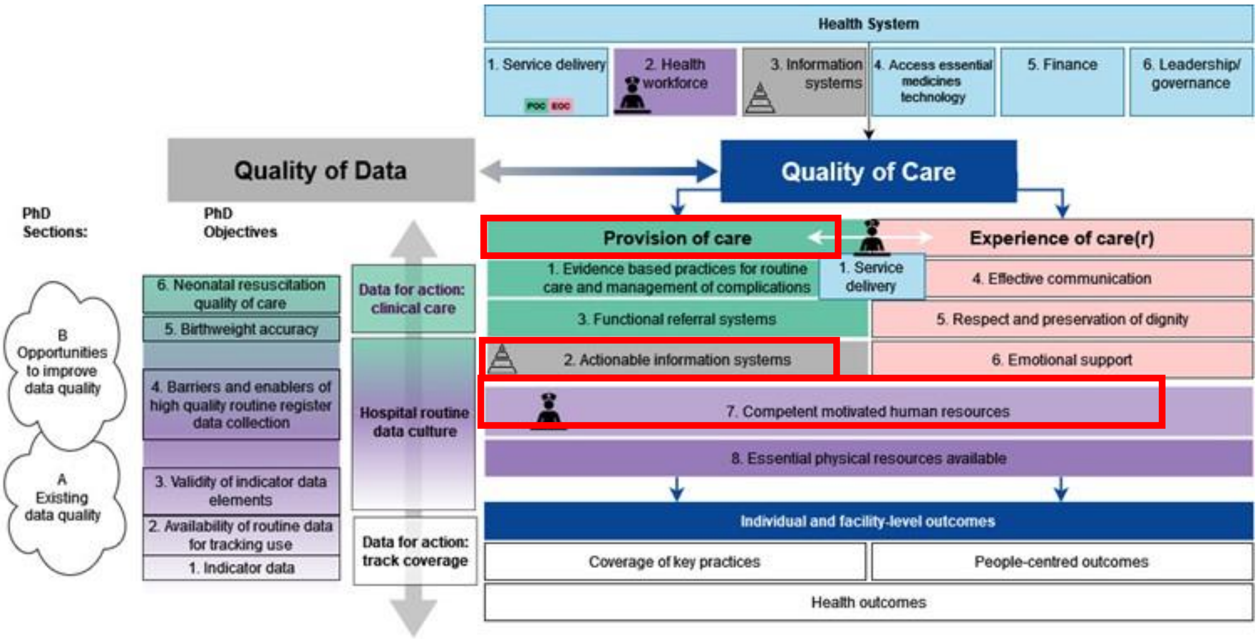
Unless a baby is known to have died in utero following ultrasound diagnosis, neonatal resuscitation is initiated for all babies who do not breathe. If signs of life are identified before, during or after resuscitation, the baby is a livebirth and if not, a stillbirth. All babies, including stillbirths, should be weighed, and examined for congenital abnormalities. Depending on hospital policy, a stillborn baby may or may not be assigned an individual case notes, thus the documentation of this care a stillborn baby receives (weight, resuscitation) may only be written in the labour and delivery register and sometimes in their mother's case notes. Improving documentation of resuscitation efforts will help improve stillbirth measurement and could contribute to reducing misclassification between neonatal death and stillbirth. These data can be used to improving quality of care for stillbirths linked to maternal and perinatal death audit/ surveillance and response.

9.3.2 Hospital data for action – how can data quality and use be improved?

The studies in this PhD found data quality varied substantially by indicator and by hospital. Similar findings for validation research for population-based surveys would result in the question being removed from the questionnaire due to lack of diagnostic validity. For RHIS core indicator measurement, the task is to improve data quality to enable the data to be trusted and used.

In this subsection of the thesis discussion, I use the research studies from chapters 3-8 to explore opportunities to improve routine health facility data quality for tracking coverage and clinical care. Specifically, I will discuss the relationship among the three domains of interest in the 'Quality of Care and Quality of Data Conceptual Framework': actionable information system, health professionals and provision of care (Figure 15).

Figure 15: 'Quality of Care and Quality of Data conceptual framework' - Data for action: clinical care



I will illustrate with findings from my clinical and HIS experience practicing in high mortality settings and from follow-on research **Advancing Routine Health Information Systems (RHIS) to Deliver for Every Newborn - EN-BIRTH 2 study** which I co-created as the principal investigator.

9.3.2.1 Routine newborn data - for whom and for what purpose?

This subsection considers how improving data quality is dependent on understanding why the data are being collected – for whom and what purpose. First, I will discuss the users of data to track coverage and outcomes for four indicators (neonatal resuscitation, low birthweight, early initiation of breastfeeding and uterotonics to prevent postpartum haemorrhage). Second, I will discuss data for clinical care and quality improvement.

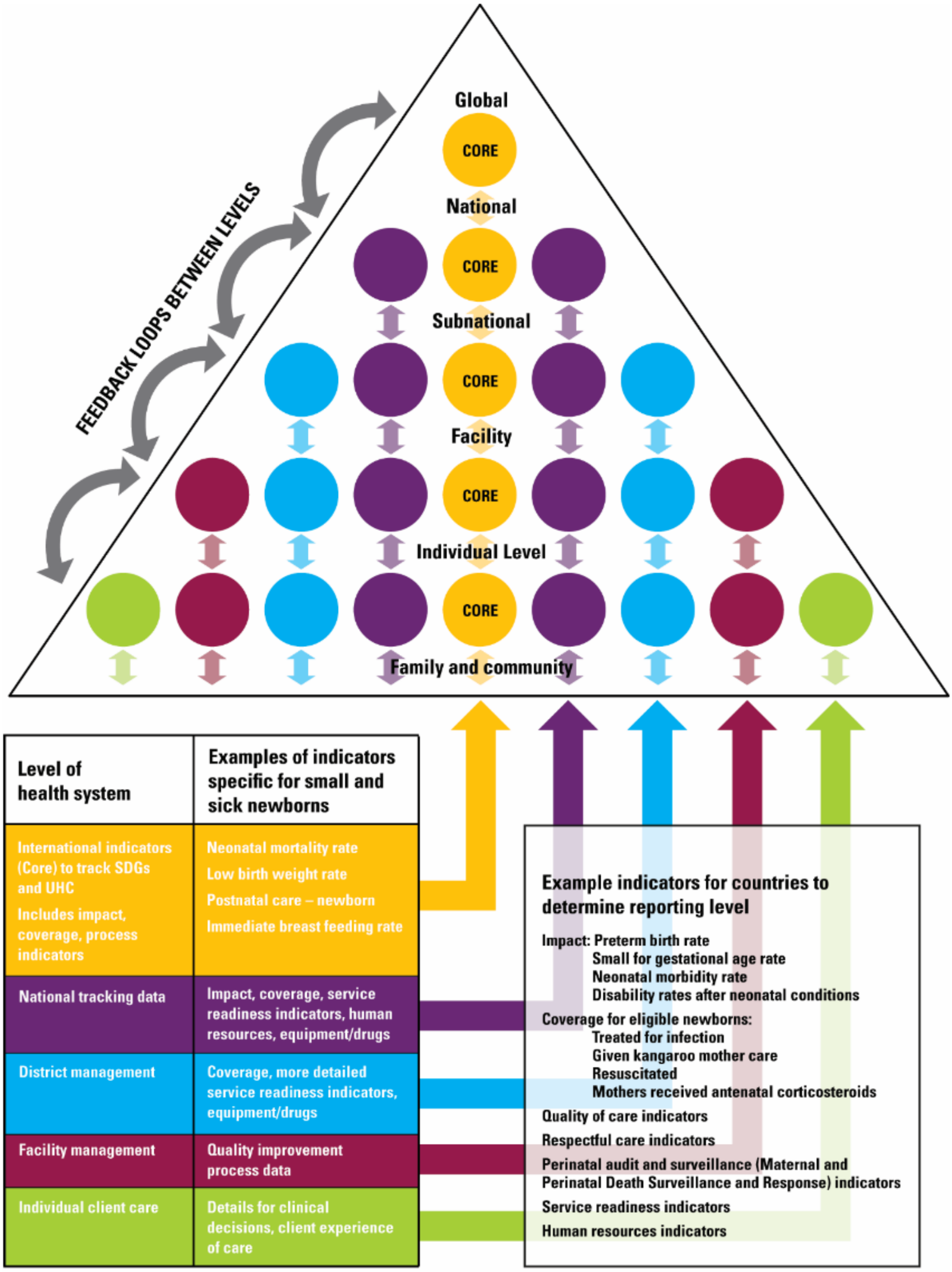
9.3.2.1.1 Data to track coverage and outcomes.

The ENAP core indicators prioritised in 2014 (Figure 4) were intended to be used for action during the SDG era.[50,96] A recent scoping review regarding data in LMIC indicated that better data might improve health service delivery but the evidence remains limited regarding link to health outcomes.[89]

In the 2023 WHO and UNICEF 'Improving maternal and newborn health and survival and reducing stillbirth: progress report', among 106 countries reporting, only 44% track neonatal resuscitation in national RHIS.[37] By contrast, low birthweight, also listed as a 100 Global Core Health Indicator, is reported in national RHIS by 88% of these same countries.[37,49] In the 2020 WHO 'Survey, Count, Optimize, Review, and Enable (SCORE) Global report on health data systems and capacity' report, among 133 countries, only 51% track low birthweight in RHIS sub-nationally and only 25% of countries disaggregate by sex.[7]

The presence of newborn indicators in national RHIS reflect the reality that countries determine reporting level for indicators, as indicated in Figure 16**,** which I designed for my first author chapter 5, 'Data for Action' in the 2019 WHO/UNICEF 'Survive and Thrive Transforming care for every small and sick newborn' report.[97]

Figure 16: Data pyramid adapted for routine newborn data at different levels of the health information system.



Adapted from: Heywood and Rohde, 2000.

Tracking neonatal resuscitation

The EN-BIRTH assessment of routine register data quality for neonatal resuscitation was one of the higher performing indicators, recording a bag-mask ventilation rate of 4.3% of total births compared to 5.1% observed to receive bag-mask ventilation (PhD Thesis Chapter 8, Objective 6). For national tracking this is surely a "good-enough" measurement, but the question remains, for what purpose is a crude measure of bag-mask ventilation being tracked in registers?

The bag-mask ventilation indicator is currently listed as a coverage indicator, implying the true denominator is babies requiring bag-mask ventilation are babies not breathing after birth and after initial steps of stimulation have been tried. As reported in the manuscript (PhD Thesis Chapter 8, Objective 6), trying to capture the true denominator for that target group "non-breathing babies" in routine registers may not be practical. Frontline health professionals are unlikely to record in a routine register for a non-macerated baby "not breathing" without also documenting "bag-mask ventilation given". Thus, tracking bag-mask ventilation as a coverage indicator would likely yield close to 100% coverage and not be meaningful.

However, using the bag-mask ventilation indicator as an outcome measure of quality of intrapartum care instead of a process measure of care may be worth considering. Currently the crude measure for quality of intrapartum care is the outcome intrapartum stillbirth, which has previously been shown as challenging to measure.[98]. An intrapartum stillbirth is defined as intrauterine fetal death diagnosed prior to the onset of labour. Latest guidance states that if fetal heart sounds are absent on admission in early labour, this baby should be classified as an antepartum death.[99] Thus an in-facility intrapartum stillbirth is defined as a fetus is alive at the time of admission and dies during the provision of care in the health facility.

The concept of near-miss death is well established for maternal health audit, but less so for newborns/ stillbirths.[100] While neonatal near-miss death audit aims to identify cases to target for clinical audit to explore quality improvement opportunities, currently there is no consensus on criteria to define neonatal near-miss and multiple definitions are in use.[101] In high mortality settings, definitions currently include clinical and management criteria including: 5 minute Apgar score less than 7, a birthweight less than 1750g, gestational age less than 33 weeks, use of therapeutic intravenous antibiotics, intubation and phototherapy.[102]

Measuring bag-mask ventilation rates among total births from labour and delivery registers and tracking as a crude "near-miss stillbirth" outcome indicator may complement measuring intrapartum stillbirth rate. When the 'Helping Babies Breathe' algorithm is implemented, the one minute Apgar cannot be measured, as the first assessment of the neonatal heart rate is when the newborn is two minutes old.[56] For babies born in primary or secondary apnoea, their outcome as a live birth or stillbirth is only determined after the resuscitation including bag-mask ventilation. Disaggregating the bag-mask ventilation rate by livebirth and stillbirth will enable checking that babies with no signs of life are resuscitated.

Including near-miss deaths alongside deaths in audits was very helpful in my experience of leading hospital perinatal death audit in the high mortality setting of south Asia. Specifically, for the interprofessional team to be able to celebrate together a life saved as well as identify any missed opportunities was important for team morale. The outcome for babies receiving bag-mask ventilation will be include both good and poor outcomes. Thus, reviewing these babies and mothers care where the positive action of neonatal bag-mask ventilation had been provided by a health professional can be productive, alongside reviewing stillbirths and women and newborns who had died.

As bag-mask ventilation is an intervention relevant for both live births and stillbirths, improving indicator measurement will likely require strengthening routine labour and delivery registers. This data source is critical as stillborn babies in high mortality settings are typically not assigned their own individual case notes, although sometimes care provided is written in their mother's case notes. More details regarding the observed care stillborn babies received is published in one of the EN-BIRTH supplement papers.[88]

Tracking low birthweight

Low birthweight rate is a long-standing core health and nutrition indicator.[49] The indicator is also listed for the broader monitoring context for SDG target 2.2 "By 2030 end all forms of malnutrition, including achieving by 2025 the internationally agreed targets on stunting and wasting in children under five years of age, and address the nutritional needs of adolescent girls, pregnant and lactating women, and older persons".

The research presented in this thesis Chapter 7, Objective 5, showed the contribution of digital scales to improve birthweight measurement. Further exploration as to why birthweights recorded at night were more heaped, possibly reflecting suboptimal staffing levels. A linked EN-BIRTH co-author publication identifies more opportunities to improve this clinical practice including weighing scales calibration.[103]

Tracking early initiation of breastfeeding

The early initiation of breastfeeding is another cross-cutting core nutrition and health indicator and similarly to low birthweight rate cited in the broader monitoring context for SDG 2.2. Unlike low birthweight rate, the early initiation for breastfeeding performed extremely poorly when assessed for measurement validity as discussed at length in the PhD Thesis Chapter 5, Objective 3.

To improve measurement of this indicator, ensuring the correct data are captured is critical. In the protocol paper (PhD Chapter 3, Objective 1) the indicator definition was based on definitions from the ENAP measurement documents as the "proportion of babies whose breastfeeding was initiated within one hour of birth". Data collectors captured the early initiation of breastfeeding observation using the EN-BIRTH data collection tablet app, by pressing the screen over the word "breastfed". The WHO 2018 document *Global reference list of 100 core health indicators* uses the definition "percentage of newborns breastfed within one hour of birth". This includes a footnote to the 2008 document *Indicators for assessing infant and young child feeding practices: definitions and measurement* which has the full definition as "put to the breast".[49,104] The differences in the wording of these definitions caused debate among the EN-BIRTH research team leading to discussions with nutrition experts as to the exact meaning of the words "put to the breast". It was agreed that the clarity of the wording in the definition could be improved.

Since the publication of the EN-BIRTH manuscripts, in 2021 WHO published updated *Indicators for assessing infant and young child feeding practices*.[105] This highlights the definition of the early initiation of breast feeding very clearly as "percentage of children born in the last 24 months who were put to the breast within one hour of birth". There is a clear footnote stating "Early initiation of breastfeeding does not require that the infant suckled at the breast or that milk was transferred from breast to infant. It represents the practice of putting the baby to breast within the first hour, which is related to many positive outcomes including reduced mortality and exclusive breastfeeding."

The editorial linked to the validation paper (PhD Chapter 5, Objective 3) notes "And even the gold standard, direct observation, proved hard when the indicator was open to interpretation (e.g., does early initiation of breastfeeding mean that the newborn baby was put to the mother's breast soon after birth or that the baby started sucking within 1 h of birth?). We should take heed of the implications of these findings to support and improve fit-for-purpose data collection mechanisms".[106] This is an important point for both research and routine data collection mechanisms.

To improve the measurement of the early initiation of breastfeeding in routine labour and delivery registers, it is vital that all actors are using the same definition. Health professionals capturing data in labour and delivery wards need to have access to this definition for their clinical practice as well as data capture. Ensuring that labour and delivery register columns are labelled "put to the breast" is clearer than "breastfed" which might be interpreted as attachment or sucking. If the indicator has time in the definition, then the cut-off ideally would be included, if space allows: "put to the breast within one hour of birth". Including these details would support indicator definition dissemination as well as acting as a job aid to ensure tracking the indicator of early initiation of breastfeeding is measuring what it is intended to measure.[92]

Tracking uterotonics to prevent postpartum haemorrhage.

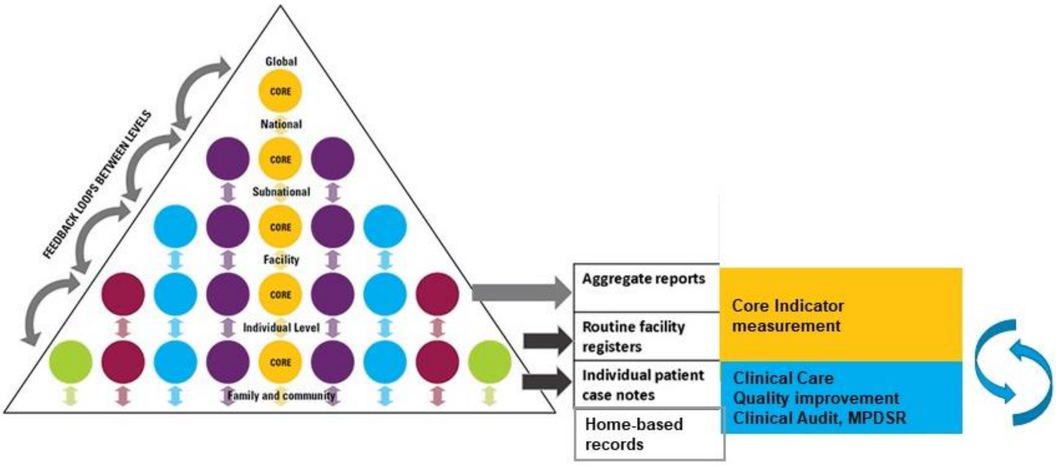
This indicator was proposed by the maternal health community for assessment of validity in the EN-BIRTH study during the design phase. The intervention is critical to prevent PPH, still the leading cause of maternal mortality. As found in the EN-BIRTH study, coverage of the preventative intervention should be near universal, except for a few women requiring or requesting physiological third-stage management.

There are many critical practices and interventions performed during intrapartum and immediate postnatal care, e.g., protecting the perineum and drying the baby. These can be captured in individual case notes, if necessary, but it must be considered whether it is worth the effort to duplicate recording uterotonics to prevent PPH also in routine labour and delivery registers. The appetite for data must respect frontline health professionals' limited time – "is the juice worth the squeeze"? Considering how tracking the indicator uterotonics to prevent PPH may be used, if indicator is tracked and the rate drops to zero, stockouts of drug/ syringe/ needle can be explored. However, logistics information systems would be a more appropriate data source. As the focus of my PhD is routine register data for newborns, I will not discuss this maternal indicator further.

9.3.2.1.2 Data for clinical care and quality improvement

The information required for clinical care decisions is a much larger "dataset" that that required for core indicator measurement. It is important to envisage routine registers in labour and delivery ward (and the registers that exist on postnatal and KMC wards) as interoperable with the individual case notes for the newborn and mother (Figure 17).

Figure 17: Health facility routine register and individual case notes interoperability for the data pyramid



In my thesis thus far, I have focused on core newborn indicators prioritised for ENAP (Figure 4) and in the previous section justified why or why not these data should be captured in routine registers. The interoperability of routine registers and individual case notes is a pressing issue with the focus on improving quality of care to end preventable newborn and maternal death. Global core lists of quality indicators have recently been published by WHO for paediatrics and adolescents and I have been co-leading the parallel prioritisation process for small and sick newborns.[107,108] These quality indicators prioritised from RHIS will need data from individual case notes for efficient measurement.

Registers should contain only the data needed for core indicator measurement. Case notes need to contain all the data needed for clinical care, for use in real-time during admission and for clinical audit, quality improvement and maternal and perinatal death surveillance and response (MPDSR) if the baby dies.

The recent 'Improving maternal and newborn health and survival and reducing stillbirth: progress report 2023' identifies priority actions to improve data quality and use which include 'individual-level data set for neonatal quality of care'. These datasets require efficient individual patient case notes for source data. There is even less evidence regarding neonatal data quality in high mortality settings from case notes than for routine registers.[66] Evidence is needed regarding RHIS determinants for high quality data from individual neonatal case notes data.[66]

In addition to using case notes data for newborn indicator measurement, data element columns in routine registers could be removed, contributing to streamlined systems and less duplicative documentation for health professionals.

9.3.2.2 Routine newborn data – why is the paper and digital connection important?

Recent advances in digital platforms e.g., DHIS, are transforming RHIS in many high mortality settings. However, in our enthusiasm to embrace digital systems we must not overlook that in many high-burden mortality settings, paper-based registers and individual case notes are currently the source data entered to digital data platforms and are likely to remain so for several years.

The "forgotten power of paper" was discussed back in 2009 in relation to electronic medical records.[109] More recently in Kenya, the challenges of paper-based subsystems were highlighted as continuing to affect data for tracking coverage into DHIS2 as well as clinical care.[65] During the time of paper-digital transition, streamlining and organising the paper is an important consideration to improve data quality now and support efficient digitisation in the future.[110] Without paying attention to the paper landscape that underpins powerful digital systems we risk "**garbage in, garbage out"** (**GIGO**) –expressing the concept that flawed or nonsense (garbage) input data produces nonsense output.[111]

Just as digital platforms require system specification and design, it is important to design paper-based health facility medical records, including those which capture neonatal data. All data must be in the individual case notes medical record used clinically, and a subset of the same data in routine registers for easy aggregation, if paper-based summary monthly reports are generated. If the source document for routine measurement currently are routine registers and the data are not captured, e.g., bag-mask ventilation was not in the national Nepal register at the time of the EN-BIRTH data collection, this would result in limited national capacity to track this indicator.

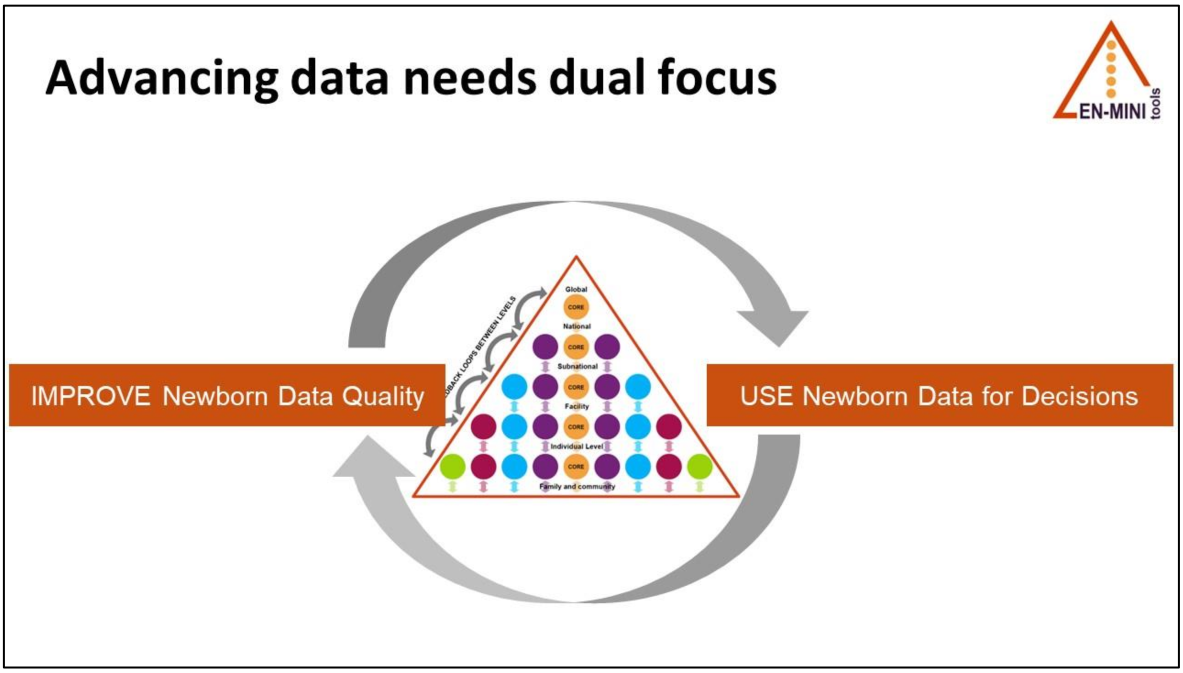
Digitisation can occur in a stepwise fashion, first entering individual data from the routine register in lieu of aggregate data, then adding in additional data points for the individual mother/baby from the case notes. This process enabled high quality data to be collected efficiently in rural south Asia.[112]

9.3.2.3 Routine newborn data – what is the data quality and data use cycle?

One of the aspirations of the EN-BIRTH study was to generate evidence to improve use of RHIS data for action by building confidence in routine register data for use, in response to distrust in data quality. The varied and sub-optimal data quality we found for many of the data elements, reinforced that efforts are needed to improve source data for use.

The relationship between data quality and information use is shown as a bidirectional arrow at the output level of improved RHIS performance in the PRISM framework.[113,114] Data use and data quality can function as a vicious circle. Data that are not trusted will not be used. Data that are not used result in no feedback and no motivation to improve quality. Using "good-enough data" with feedback loops drives a virtuous circle of improving data quality (Figure 18).[115] This concept is similarly expressed in the data demand and use model.[10,114]

Figure 18: The vicious/ virtuous circle of routine data quality and data use



This relationship between data use and improving data quality underpinned the follow-on collaborative EN-BIRTH 2 research leading to the design of the **"Every Newborn-Measurement Improvement for Newborn and Stillbirth Indicators (EN-MINI) tools for Routine Health Information Systems"**.[116] I conceptualised and co-created the EN-MINI Tools as a major output of the EN-BIRTH 2 study. My role included co-designing the EN-MINI-PRISM Analysis Tool and drafting the results report from Tanzania shown in Appendix 4.

The EN-MINI tools guide priority actions to improve availability, quality, and use of newborn indicators in the wider RHIS, beginning with routine register data and moving up the data-information pyramid. The EN-MINI tools measure performance of the wider RHIS including generating quantitative measures of some of the barriers and enablers identified in (PhD Thesis Chapter 6, Objective 4). Details of selected research findings from the pilot EN-MINI Tools assessment in Tanzania, currently published as an open access report are found in Appendix 4.[3]

9.3.3 Enabling a virtuous data quality/ data use cycle

In this section I will describe the potential opportunities identified in my PhD to enable a vicious cycle to become a virtuous cycle for routine newborn data use and data quality.

Evidence to improve routine health data has typically focused on district level and above and suggests interventions to address more than one determinant of the PRISM framework is more successful.[6,117] I will focus in this PhD thesis mainly at the health facility level which has been less studied. The opportunities presented are framed within the rationale established for improving routine paper-based register data quality and described using PRISM framework inputs, determinants and processes.[118] Implementing combinations of interventions addressing both behavioural and technical factors is most likely to be successful in improving data quality and use of RHIS in LMIC according to a recent scoping review.[119]

9.3.3.1 Increase routine register data use

Identifying ways to increase use of routine register newborn data in the in the health facility in addition to passing aggregate data up the data-information pyramid is an important step to enable a virtuous circle of data use/quality.

Increased use of register data, in theory would create ownership, develops data and research competencies, and increase the relevance of the data capture process – an RHIS output.[120] The EN-MINI-PRISM Tools pilot assessment in Tanzania found evidence of data use in the district/sub-national data office was higher than in the health facility (Appendix 4.1 Figure 28).[3] Register data could be used for similar purposes at the health facility as at the district/ sub-national data office: for planning, problem-solving and improving data quality. However, interventions to improve data use for decision-making at health facility level has previously been found to have improved less than at district level by researchers working in Côte d'Ivoire.[121]

In this subsection I suggest specific opportunities to increase routine labour and delivery register use at health facility level, identified from my thesis analyses:

Track birthweight heaping as a marker of data quality

Facility staff, either health or RHIS data professionals, can explore routine register birthweight heaping. Currently, register birthweight data is typically used only to aggregate the number of low birthweight babies for monthly summary reports. Health facility staff could calculate birthweight heaping from their own data using the methodology described in my PhD Thesis Chapter 7, Objective 5. Calculating heaping at 2500g would be ideal to start, as the cut-off weight determining low birthweight or not. For efficiency this should be done entering individual data points into an MS excel worksheet or similar database in the health facility, but if there is no computer access then could be done by manual aggregation to answer the question "what percentage of your livebirths are exactly 2500g".

If birthweight heaping is identified, the health facility team can explore causes and implement solutions to improve birthweight measurement for both clinical care and core indicator tracking. Repeating the birthweight heaping calculation using routine register data should be repeated to identify whether actions taken were useful. Regular reporting of birth weight heaping at 2500g at the health facility level or higher in the data-information pyramid could be an efficient measure of low birthweight data quality and an example of using routine register data to improve the data quality linked to improving the clinical practice of birthweight measurement.

Clinical Audit/ quality improvement

Routine register data can also be used to identify the cohort of total births for clinical audit/ quality improvement including for MPDSR is illustrated by four examples:

Example 1) identify all newborns who received bag-mask ventilation (proposed above to be "near-miss stillbirth") and review the baby and mother's case notes to identify good practice to celebrate that avoided a death as well as opportunities for improvement. The EN-BIRTH gap analysis observation data showed a large quality gap for bag-mask ventilation timeliness – very few babies received their first ventilation breath within the recommended one minute (PhD thesis Chapter 8, Objective 6). It is important to explore ways to link crude coverage with measures of high quality. Observation was the research method used in EN-BIRTH, but in my experience using a low-cost timer/ stop-clock to practice during neonatal resuscitation training in high mortality settings, is effective in refining the first 'golden minute' skills.[56] For fidelity of skills transfer from the classroom to the labour and delivery ward it may be feasible for the clinical team to use the same timer/ stop-clock in the labour and delivery ward to measure time to first inflation breath as well as other time-sensitive care practices.

Example 2) identify all stillbirths and neonatal deaths in the facility and cross-tabulate to calculate what proportion were given bag-mask ventilation. If bag-mask ventilation rates for stillbirths are very low, explore the reasons why and consider prospective tracking.

Example 3) improve birth weight measurement for all stillborn babies, important for the:
a) Definition of stillbirth is dependent on birthweight in the absence of as gestational age. A stillborn baby is born with no signs of life at 22 or more completed weeks of gestation: early gestation stillbirth (at 22 to 27 completed weeks of gestation) and late gestation stillbirth (at 28 or more completed weeks of gestation). For purposes of international comparison, stillbirth estimates represent late gestation stillbirths.[36] If gestation is not available a birthweight of >1000g is used.

b) Assessment of whether the stillborn baby's in utero growth was appropriate or small or large- for gestational age. This requires accurate gestational age assessment and standardised charts.[122]. Linkage to antenatal care assessment of fetal size and growth and whether there were missed opportunities to prevent stillbirth for small or large-for-gestational age babies is an important antenatal/ intrapartum quality improvement linked initiative.

Example 4) identify stillbirths born by caesarean section. Systematic review evidence has reported up to one in ten babies born by caesarean section in the sub-Saharan African region are stillborn.[123] Typically maternal and newborn health outcomes from LMIC have been tracked using population-based studies which, until recently, have used livebirth denominators. Thus, stillbirth as an outcome after caesarean section may be obscured. A recent multi-site district study found, compared to vaginal delivery, caesarean section was associated with lower stillbirth rate (SBR) in South American (Guatemala) and south Asian (India, Pakistan) but higher SBR in Africa (Kenya, Zambia and Democratic Republic of Kenya) and called for studies to understand these differences in SBR.[124,125]

I propose using routine labour and delivery registers to identify stillbirths born by caesarean section to increase routine register data use would be beneficial alongside the intrapartum stillbirth rate which is often used as an overall measure for quality of provision of intrapartum care.[126,127] Clinical audit of these women and baby's case notes to explore any missed opportunities for provision of routine care effective to prevent stillbirth, include monitoring fetal heart rate for timely intervention and neonatal resuscitation when needed.[128,129]

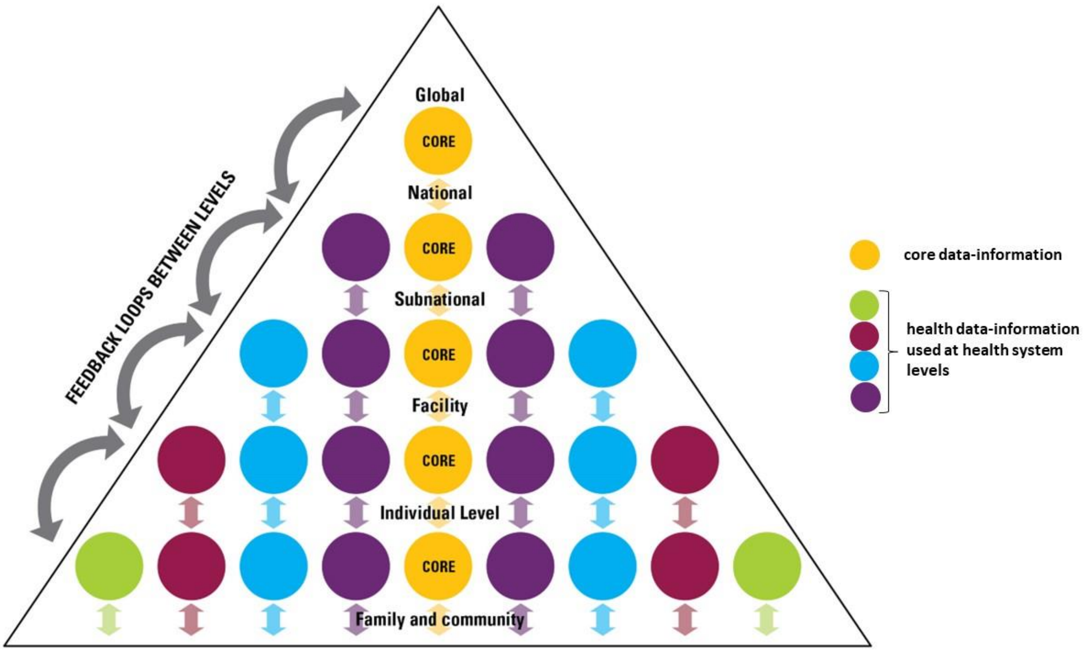
9.3.3.2 Strengthen RHIS feedback

Feedback is an important enabler for a virtuous cycle of data quality/ data use, yet this was lacking between data users higher up the data-information pyramid to those capturing the data in the labour and delivery ward (PhD Thesis Chapter 6, Objective 4). Health professionals interviewed described the lack of feedback and not feeling valued for RHIS tasks as demotivating to capture high quality data. Frontline health professionals deserve to know for whom and for what purpose they are capturing such a large amount of data in registers.

Feedback on RHIS data ideally would include acknowledgement of completeness and timeliness of reports, and importantly visualisation of analysed results from the data sent up the data-information pyramid.

Strengthening the feedback loops between levels in the data-information pyramid is therefore needed to improve data quality for use. To illustrate this I have added bidirectional feedback loops to later data-information pyramid graphics as shown in Figure 19 after the use of this figure in EN-BIRTH protocol paper (PhD Thesis Chapter 3, Objective 1) and in Figure 3 in this thesis. These feedback arrows are can also be seen in Figure 14, Figure 16, Figure 17.

Figure 19: Data-information pyramid showing feedback arrows between levels.



The nature of feedback at minimum should include the neonatal core indicator report generated for the health facility from the data "sent up" the data pyramid for outcomes and coverage indicators. If facilities are entering their data directly on the electronic RHIS in the health facility (e.g., DHIS2), ensure the software is set up at minimum to visualize their own health facility analysis, ideally aggregated with sub-national and national data for comparison to promote use of data.

This need to strengthen feedback was found during the EN-MINI-PRISM Tools Pilot assessment in Tanzania in 2021 (Appendix 4.2 Figure 29).[3]

9.3.3.3 Health facility data quality check

EN-BIRTH was a research study that used observation as an external gold standard to assess validity measurement of newborn core indicators data element in five hospitals at one point in time. Identifying ways to check data quality as part of regular hospital activities can be enabling for the team of health professionals to "pay attention" to data. One possibility for regular data quality checking is health facility-level MPDSR. During the process of reviewing each perinatal deathand near-miss death the data elements captured in the routine labour and delivery ward register can be cross-checked against details documented in the maternal and newborn individual case notes. Whilst this may be a feasible process to pilot, one limitation is that the same health professional is likely to have documented the clinical data regarding intrapartum and immediate postnatal care and outcomes in both these medical records, thus the case notes and labour and delivery register are not completely "external" data sources. This research identified that nurse-midwives predominantly take responsibility for routine labour and delivery register documentation (PhD Thesis Chapter 6, Objective 4). When doctors have conducted the delivery including after caesarean section in the operating room, they should take responsibility for checking the register to ensure that good quality data capture is the responsibility of the whole team. Using the routine register as the source for clinical handover between shifts, reviewing the previous day's births at the start of the day or weekly review meetings are other examples to institutionalise register data use that someone in the team might notice incomplete or inaccurate data.

Designing the hospital electronic RHIS intentionally for data entry from more than one source can act as an ongoing data quality check. For example, in the system I designed in rural south Asia, each day the labour and delivery register was physically taken to the MIS office and each data element captured in the register was entered for each woman by her unique identifier. Meanwhile, the clinical care on the postnatal ward was continuing to be written in the individual case notes. After discharge the case notes were taken to the MIS office and the additional data elements added to the initial dataset from the routine register. In the process, any discrepancies between the register data elements and case notes data elements were spotted and discussed with the clinical team for resolution.

9.3.3.4 Standardise register design

One of the major themes identified in PhD Thesis Chapter 6, Objective 4[26] was design of the paper register is an opportunity to improve data quality. Register design and complexity are both technical factors of the RHIS input determinants in the PRISM framework.

The research was conducted in hospitals with many data elements captured in the national standard printed labour and delivery registers: there were 35 in Nepal, 48 in Tanzania and 58 in Bangladesh. Additionally, multiple hand-written “informal” registers were operationalised alongside the “formal” printed registers described as necessary to collect data elements needed for reporting. In Nepal, none of the coverage indicator of interest were captured. In Bangladesh and Tanzania, health professionals described the large amount of time they spend during their shift on documentation and how this detracts from time to provide high quality care.[26]

As described earlier, considering that registers are duplicate data in parallel to individual case notes, the burden of reporting by busy frontline health professionals in these settings is large. These labour and delivery ward registers in high mortality settings collect many more data elements than are typically collected in lower mortality settings. Year on year as expectation from RHIS contribution to data increases, the demand for data fuels “just one more column” syndrome or “just one more register” syndrome linked to more summary forms required to be filled.   As expected, this leads to pressure to the register with the challenges of many columns and health professional burden as described above.

This finding of the differences between register designs led to my co-design of a novel “MAP Newborn Data” EN-MINI tool 0 for all levels of the data-information pyramid: register, tally sheets, summary forms and electronic RHIS.[130] Among the automated report outputs in Section 4: Documentation burden, the RHIS load is presented as a balance of WHO- or nationally recommended core/optional indicators with other newborn data (Appendix 4.3, Figure 30).[131]

Routine register standardisation is proposed to improve routine register data quality, including a standard set of data elements, a design to optimise high quality data capture and efficient digitisation. The EN-BIRTH study identified that register design affected data quality – specific and non-specific columns affected data completeness. The design in Bangladesh meant completeness could not be calculated because “true zero” and missing data could not be distinguished.

An evidence gap exists for register standardisation that is user-friendly to capture high quality data for use. As already described in section 9.3.1.2, routine register and case note are interoperable thus design and standardisation of the whole medical record system is highly preferable (Figure 17).

Health professionals described a large time burden for documentation in general. Research in LMIC outpatients setting found mean time spent on recording was about one-third the total consultation time.[132] Time-motion-studies are a recommended quantitative data collection method where an external observer captures detailed data on the duration and movements required to accomplish a specific task, coupled with an analysis focused on improving efficiency.[133] There is limited use of time-motion-studies in the LMIC inpatient settings, although practical guidance from Tanzania regarding monitoring compliance with hand hygiene guidelines has been published.[134]

Evidence is needed to guide the development of standardised medical records including labour and delivery registers to be user-friendly and capture high quality data. Time-motion-studies in the labour and delivery ward LMIC setting to assess the proportion of time providing and documenting care are needed. Medical records data related activities could be stratified by routine registers and case notes, to identify opportunities to improve efficiencies. Time-motion-studies could be conducted alongside other methods of data quality assessment and health professional experience of proposed standardised registers.

9.3.3.5 Standardise register procedures

The finding of the identical register design performing differently when assessed for validity of measurement of newborn indicators between two hospitals in both Bangladesh and Tanzania, illustrated that other determinants including organisational, behavioural and processes are important. In the PRISM framework the task of filling the register aligns with the RHIS process of data collection and the organisational factors including standard operating procedures (as a component of governance), training and supervision (Figure 5). Notably, evidence for interventions addressing organisational factors were non-existent in the scoping review improving quality and use of RHIS in LMIC, thus conducting implementation research for these proposed interventions could be considered.[119]

Register filling guideline/ standard operating procedures.

Indicators have metadata definitions for the data elements of numerator, denominator etc. These are in WHO and national RHIS documents and embedded in electronic RHIS e.g., DHIS2. Governance of RHIS at facility level tends to be less developed typically listing what needs to be collected and how to use the data to calculate indicators.

Developing locally owned standard operating procedures at the health facility level for register filling and use is important. These can be used during health professional pre-service and in-service training and as a job-aid in daily clinical practice.

For high quality source register data, it is critical to ensure that each health professional capturing data in each column of the routine register can operationalise the definition. Registers in Bangladesh and Tanzania had instructions inside the front cover of the paper-based register books which can be used. Day by day health professionals will read the column label so it is important that this has high fidelity to the data element being captured.

In July 2023, I was an observer at the WHO Mother and Newborn Information for Tracking Outcomes and Results (MoNITOR) meeting where plans to develop guidance on standardisation of medical records was shared. Currently there are no recommendations regarding medical records design, filling, and use.

Training for routine register use

Health professionals reported very little training in the task of register filling in the five EN-BIRTH hospitals. Respondents reported students were also involved in register filling, highlighting the importance of pre-service training. This need for training was also found during the EN-MINI-PRISM Tools Pilot assessment in Tanzania in 2021 (Appendix 4.4, Figure 31).[3]

From my previous experience in south Asia, rotating student nurses and midwives for a one-week pre-service placement in the hospital MIS-Research department I was running to be helpful. The students enjoyed being part of the data team for a week participating in entering register data under supervision as well as medical records logistics and filing. They saw firsthand how the clinical information they and their colleagues documented in both registers and case notes was processed, analysed and reported internally to hospital management and clinicians for audit and quality improvement as well as externally including subnational RHIS.

Systematic review evidence suggests that training alone has moderate effectiveness as a strategy to improve health-care professional practices in LMIC. However, and combining supervision, the topic of the next sub-section, with training has somewhat larger effects than use of either strategy alone.[135]

Supportive supervision for routine data

Register data capture is a skill and just like for clinical skills, supportive supervision and even mentoring is important. Supervision was an emerging theme from the five EN-BIRTH facilities (PhD Thesis Chapter 6, Objective 4). This need to strengthen feedback was found during the EN-MINI-PRISM Tools Pilot assessment in Tanzania in 2021 (Appendix 4.5,Figure 32).[3] To improve supportive supervision, facilities can consider peer-to-peer data team supervision as well as supervision from higher in the data-information pyramid. This could be institutionalised after in-service team training to increase the impact of these interventions. Supervisory checklists need to include components of data accuracy e.g., birth weight heaping in addition to data completeness.

9.3.3.6 Enable RHIS skills of frontline health professionals

The PRISM Framework input determinant lists several factors covering knowledge skills and attitudes of actors involved in RHIS processes entitled "Behavioural" factors.

Quality domain 7 in the WHO quality of care framework is 'competent motivated health professionals'. This motivation and competency must apply to RHIS skills as well as the other quality domains. From a person-centred perspective it is important to consider the frontline health professionals capturing source data in routine registers.

The importance of these behavioural factors including substantial confidence-competence gaps were identified during the EN-MINI-PRISM Tools Pilot assessment in Tanzania in 2021 (Appendix 4.6, Figure 33).[3]

Another opportunity to enable RHIS skills is by building relationships between the health facility professionals responsible for data reporting with the frontline health professionals capturing the data. This can begin during pre-service training as mentioned above. Making data aggregation or entry a daily rather than a monthly activity can help. For example, in my experience of working in rural south Asia, the MIS-Research department team physically took the register each day from the labour and delivery ward for data entry. This enabled daily conversations to express gratitude for complete data capture and opportunities to clarify illegible or missing data. Interdepartmental health professional-data professional teamwork facilitated increased data quality over time. This person-centred aspect of improving data quality is important. On a wider scale similar relationships between health facilities and data offices could be considered data networks similar to networks of care.[136]

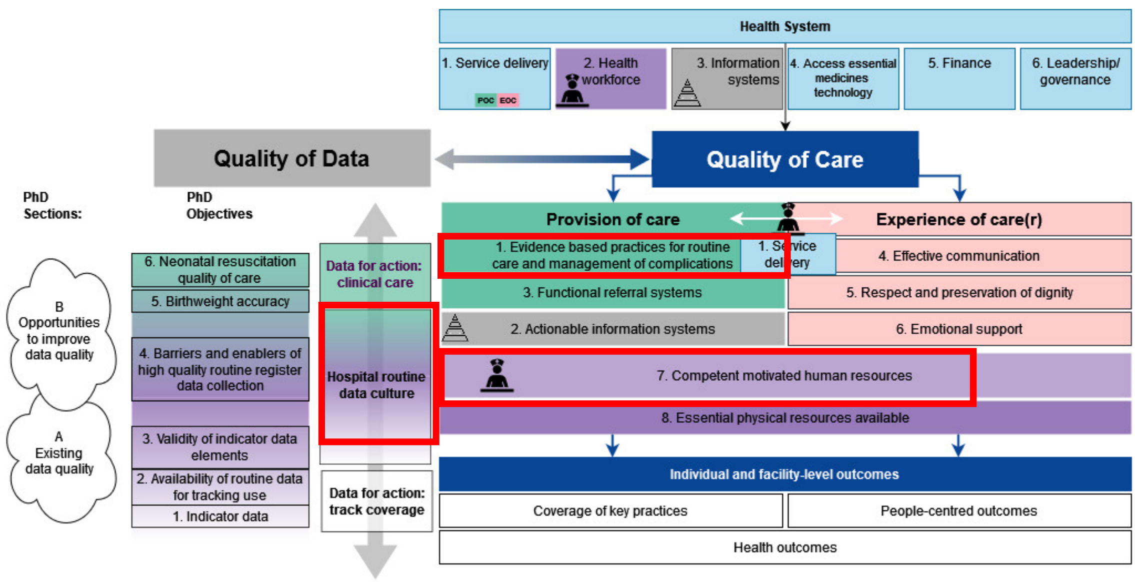
9.3.4 Hospital routine data culture

In this section I explore the importance of the **hospital routine data culture** to improve data quality for both clinical care and tracking coverage as shown at the centre of the data quality continuum (Figure 11).

The culture of information in the RHIS context is a concept defined by Hotchkiss et al as the capacity and control to promote values and beliefs among members of an organization for the collection, analysis, and use of information to achieve an organization's mission and goals.[9] This information culture is an important organisational determinant for RHIS performance.[6] Institutions that promote a culture of information typically promote five components: a) data quality, b) evidence-based decision making and accountability, c) reward mechanisms for good work, d) use of information; and e) efforts and activities to change things for the better[9]. A strong information culture has been implied in instilling a sense of responsibility, accountability and empowerment for RHIS data quality and use.[137]

In my 'Quality of Care and Quality of Data Conceptual Framework' I show the hospital routine data culture connects the people and the data. These are shown as two domains in the WHO quality-of-care framework: competent motivated human resources and the actionable information systems. which they work (Figure 20).

Figure 20: 'Quality of Care and Quality of Data conceptual framework' – Hospital routine data culture



The exploration of the barriers and enablers to documentation by health professionals in registers (PhD Thesis Chapter 6, Objective 4) shows the importance of this hospital routine data culture in the five EN-BIRTH study hospitals.[26] In these large hospitals, the environment described on the labour and delivery ward might represent an information sub-culture within the overall hospital information culture.

In this discussion, I apply the findings in these research chapters to the five components of a hospital culture of information to identify opportunities to improve data quality:

9.3.4.1 Promote data quality

We found supervisors valued *complete* data more than *correct* data. Two indicators assessed for validity of measurement for intervention coverage were for the whole of (or nearly whole) labour and delivery: women (uterotonics to prevent post-partum haemorrhage) and babies (breast-feeding). Although the registers used specific columns in the two countries capturing these two data elements, the design was different. In Bangladesh the register had instructions to tick if the practice/ intervention was done or leave blank if it was not done. In Tanzania by comparison, either "yes" or "no" are written in Swahili in the column, and a blank would indicate incomplete data.

Early initiation of breast-feeding was over-estimated in routine registers more than any other indicator in the EN-BIRTH study. The largest of these over-estimates was in the two Bangladesh hospitals, 89.9% and 87% respectively. If supervisors value completeness, then front-line health professionals would be under pressure to tick rather than leave a column blank, contributing to the over-reporting. By comparison in Tanzania, the column was completed with either a "yes" or "no", and the breast-feeding overestimate was somewhat smaller, by 69.3% and 24.7% respectively.

The results for uterotonics to prevent post-partum haemorrhage (PPH) however showed a different pattern; all four hospitals in both Bangladesh and Tanzania underestimating coverage, ranging between <2% to >30%. One reason for these differences might be that the observed coverage for uterotonics was very high, hence overestimation would not even be possible. Similarly early initiation of breastfeeding was very low, so underestimation would be limited. Pooled register-recorded coverage of uterotonics was 77.9% (95% CI 37.8-99.5) and pooled early initiation of breastfeeding was 85.9% (58.1-99.6%). These are remarkably similar register rates considering uterotonics was observed for 99.2% (98.6-99.7%) of women and early initiation of breastfeeding only 12.5% (4.6-23.6%). For both these interventions for all women/ newborns, coverage was documented in the register for more than 3 out of every birth.

When a hospital routine data culture promotes data quality by completeness rather that correctness, "ticking the column" will lead to over or underestimated coverage depending on the actual intervention/ practice coverage.

9.3.4.2 Promote evidence-based decision making and accountability

This component did not emerge as a theme from the interviews with front-line health professionals. This was a striking finding, that reinforces that there is a need to link RHIS with decision-making effectively communicate for what purpose data are collected in routine labour and delivery registers for use higher up the data-information pyramid. Specifically how is data being used for decision making.

9.3.4.3 Promote reward mechanisms for good work

Many health professionals interviewed noted a lack of appreciation for the data captured in routine registers (PhD Thesis Chapter 6, Objective 4). RHIS source data capture is "work" and acknowledging this task is key to improving the hospital routine data culture. The nature of the "reward" can be considered in each context, but in my experience, any "thank you" to health professionals for data related task goes a long way.

In contrast, health workers interviewed for the barriers and enablers study (PhD Thesis Chapter 5, objective 3) mentioned incomplete records were rewarded with words including "scolded" and "not appreciated". I postulate that this supervisory practice might contribute to the over-estimates found for indicators for every newborn e.g., early initiation of breastfeeding.

Recent evidence from Ethiopia showed evidence of widespread intentional data falsification by frontline health workers.[138] Research is needed to explore the drivers behind data manipulation and how to improve data fidelity.

9.3.4.4 Promote the use of information

Health professionals mainly described use of register data for billing and clinical handover; only one doctor mentioned indicator measurement. The lack of feedback from data users higher up the data-information pyramid to health professionals responsible for capturing the source data, was especially notable and requires action to improve the hospital routine data culture.

Moreover, exploring uses of routine labour and delivery register data can be of use to health professionals and health facility managers for decision making is key and was described in subsection 9.3.3.1.

9.3.4.5 Promote efforts and activities to change things for the better

Health professionals frequently mentioned the burden of documentation, how busy they were, and how documentation detracted from time to care for the woman and newborn. Little was expressed regarding activities to improve the situation. Future digital systems were perceived by health professionals as hopeful to improve the situation, although very few described any experience of such systems.

Overall, in the EN-BIRTH study hospitals, health professionals interviewed described contributing to the RHIS data by filling registers but not being involved in discussions regarding use of RHIS for change. High quality data needs user-friendly systems for health professionals to capture data. This begins with inviting health professionals to the discussion.

9.3.5 Linking quality of Care and Quality of Data

This thesis so far has explored hospital labour and delivery routine register data quality and opportunities to improve measurement of newborn indicators in high mortality setting. In this final section I will discuss the importance of linking data quality with quality of care, how to enable traction between data quality and quality of care and avoid tension between them.

This thesis Chapter 2 – Background and development of conceptual framework first introduced a novel adaptation of the WHO Quality of care framework, the **'Quality of Care and Quality of Data Conceptual Framework'** in subsection 2.7. At the start of the integrated discussion subsection 9.3, the concept was advanced, adding the **data quality continuum** connecting data to track coverage and outcomes with clinical information/ data. In this final subsection of the integrated discussion I will present the connected framework.

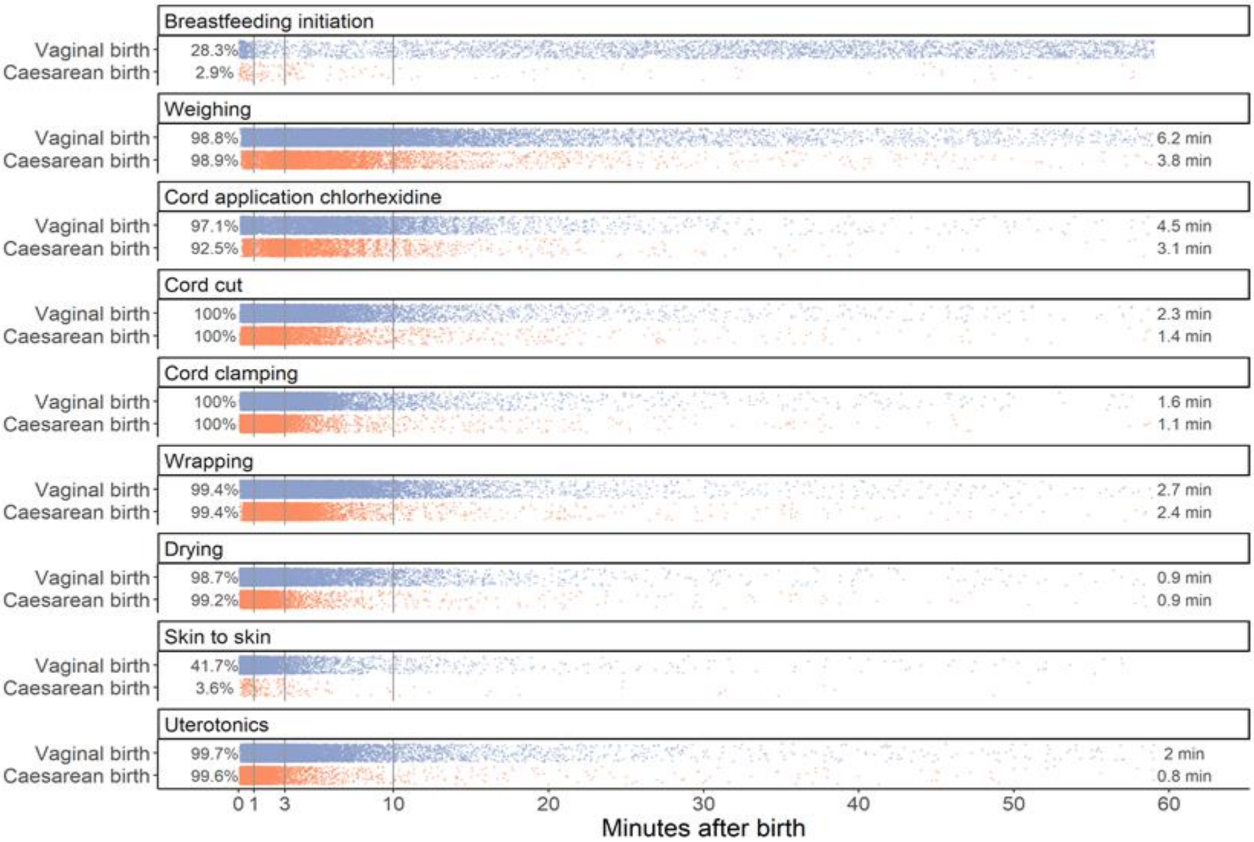
9.3.5.1 Dual responsibilities to provide quality of care and capture high quality data

The methodology used in this PhD thesis chapters 5, 7 and 8 captured observed intrapartum and immediate postnatal care, thus in this last section I will comment briefly on the quality of care provided for which the assessment of measurement validation was conducted.

High quality observed newborn care in the five EN-BIRTH study hospitals included high coverage of birthweight within one hour of birth (95.9% pooled) and for women included administration of uterotonics to prevent PPH (99.2% pooled). In contrast, suboptimal care was observed for early initiation of breastfeeding was only observed for 12.5% of babies overall and for almost no babies born by caesarean section. Observation of bag-mask ventilation for showed a both coverage and quality gaps with <1% of non-breathing babies ventilated within the recommended 1 minute.

Further EN-BIRTH analyses, currently unpublished, explored the caesarean section births in more detail.[139] Figure 21 shows a scatter plot of babies' observed care in their first 60 minutes of life stratified by vaginal and caesarean birth. The figure illustrates the multiple simultaneous care health professionals provide for women and newborns to achieve median times of baby drying at 0.9 minutes, cord-clamping at 1.6 minutes for vaginal births and 1.1 minutes for c-sections births, birthweight 6.2 minutes for vaginal births and 3.8 minutes for c-section births. Supporting immediate skin-to-skin contact in the first hour was only observed for 3.6 % of caesarean births compared to 41.7% of vaginal births.

Figure 21: Timing of newborn and maternal care practices, by mode of birth, among all newborns



Proportion receiving the intervention within first hour of life shown and median time to intervention in minutes (n= 23,182; missing data on mode of birth excluded, missing data on intervention timing excluded from relevant section; Chlorhexidine analysis only includes births in Bangladesh and Nepal as Tanzania hospitals do not have a chlorhexidine policy), EN-BIRTH study

To accelerate progress for ending preventable neonatal mortality and stillbirth, improving quality of care is needed. Improving quality of data for use is a means to an end and not an end in itself.

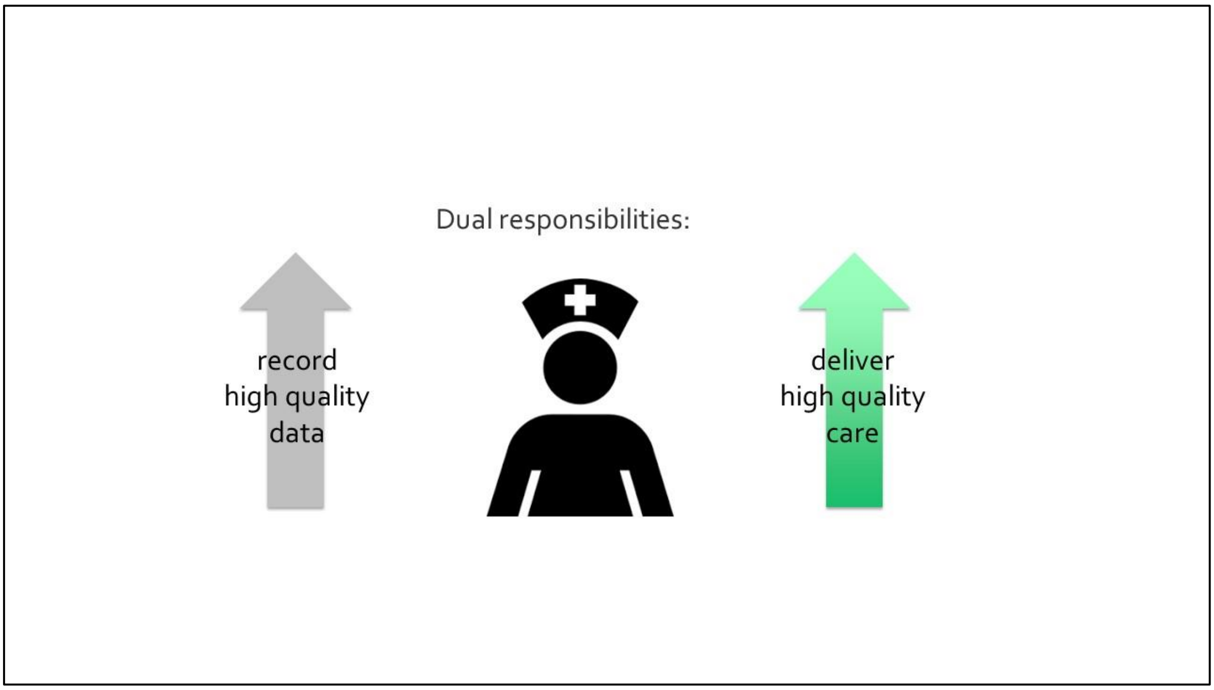
Throughout the integrated discussion I have included a person-centred focus on the health professional capturing data in the labour and delivery routine register - predominantly midwives, and nurse-midwives. The register data they capture in routine labour and delivery registers which becomes the source RHIS data for women and newborns around the time of birth. These data are the foundation of the data-information pyramid, which if over-burdening for health professionals may negatively affect time to provide high quality care.

The WHO quality of care framework describes the ideal of “competent and motivated human resources” as one of eight domains of quality (Figure 7). Competency and motivation applies to RHIS work as well as the providing care enabling a positive experience for the woman and newborn. In this PhD Thesis chapter 6, Objective 4, health workers shared some glimpses of the busy environment they work in with limited scope to do more.

The pressure health professionals face around the world in high mortality settings is well known – shortage of personnel, low salaries, undervalued. The 2021 report ‘The state of the world’s midwifery 2021, emphasise the shortage of nearly 1 million midwives and the high rate leaving the workforce.[140]

Health professionals have dual responsibilities of delivering high quality care and recording high quality data. To provide better quality care AND capture better quality care increases both their dual responsibilities (Figure 22).

Figure 22: Care and data - the dual responsibilities for health professionals



If health professionals are to provide higher quality care as well as capture higher quality data, then other tasks needs to be reduced. Alternatively quality of care and quality of data will compete with each other – *in tension.*

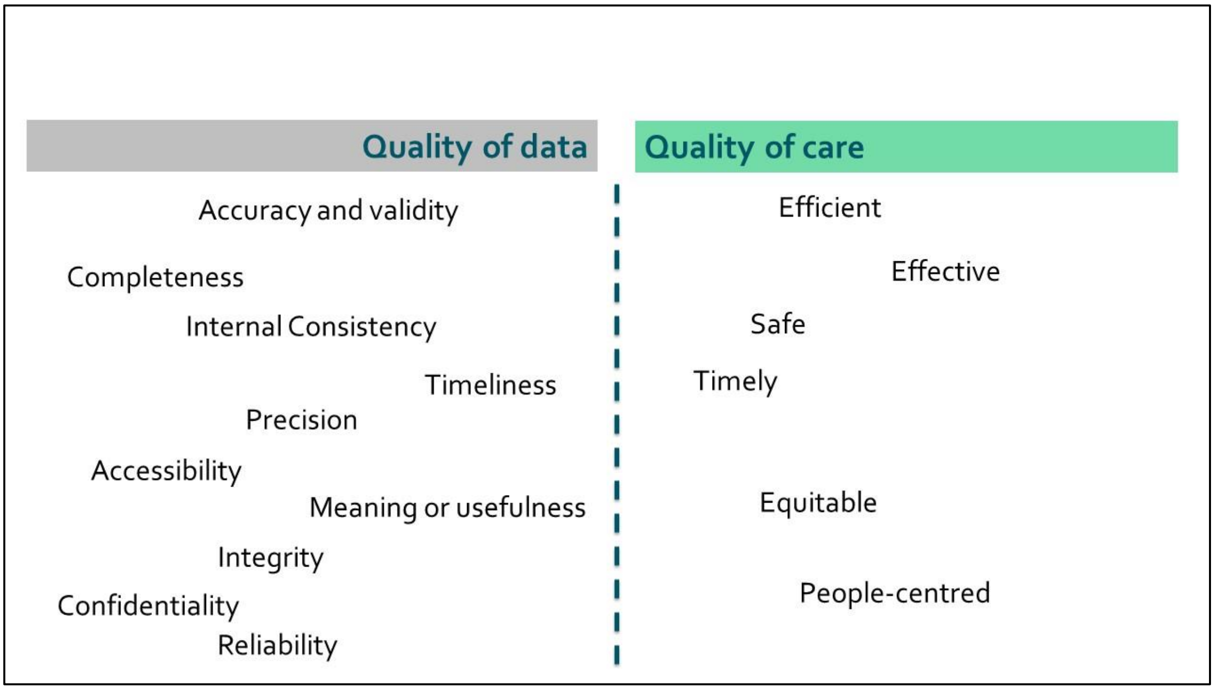
Even though the EN-BIRTH study hospitals were large hospitals, resources were constrained; for example, a lack of digital scales affected the quality of birthweight measurement.

Throughout this PhD thesis I have attempted to keep a person-centred lens – for the woman and baby of course, but also for the health professional as a person. Within the WHO quality of care framework, whilst the dimension *provision of care* is clearly *for* the woman, newborn and family *by* the health professional, the dimension *experience of care* can be considered to include health professionals. Health professionals also deserve the dimensions of experience of care: "6. Emotional support", "5. Respect and preservation of dignity" and "4. Effective communication". Feedback loops from data users across the health system to frontline health professionals who invest time capturing the source data are important and effective to improve data quality.[75]

The research health professionals regarding the barriers and enablers they faced (PhD Thesis Chapter 6, Objective 4). Inviting health professionals to participate in user-centred medical records, facilitating learner-centred RHIS education, demystifying data analysis, feeding back reports, visualisations and listening again to barriers and enablers they face are vital to improve high quality data for use from the foundation of the data-information pyramid – the frontline health professionals who capture the data that is used for public health decision making. Health professionals are not data collectors. Data should not be a major burden to them but useful for effective communication between the clinical team and for the data-information pyramid as a fortuitous spin-off.

In Figure 23, I juxtapose the dimensions of **quality of data** alongside dimensions of **quality of care.** High quality data dimensions are: accuracy and validity, completeness, internal/ external consistency, timeliness, precision, accessibility, meaning or usefulness, integrity, confidentiality, reliability.[17-19] High quality care is efficient, effective, safe, timely, equitable and people-centred.[5,20,21]

Figure 23: Components of quality of data and quality of care



Comparing these terms only timeliness and timely uses the same word. Exploring quality of data using quality of care terms or people-centred, efficient, effective, equitable and safe for health professionals' workload may be helpful to create traction between data and care. High quality data are needed to support high quality care and health professionals are central to both.

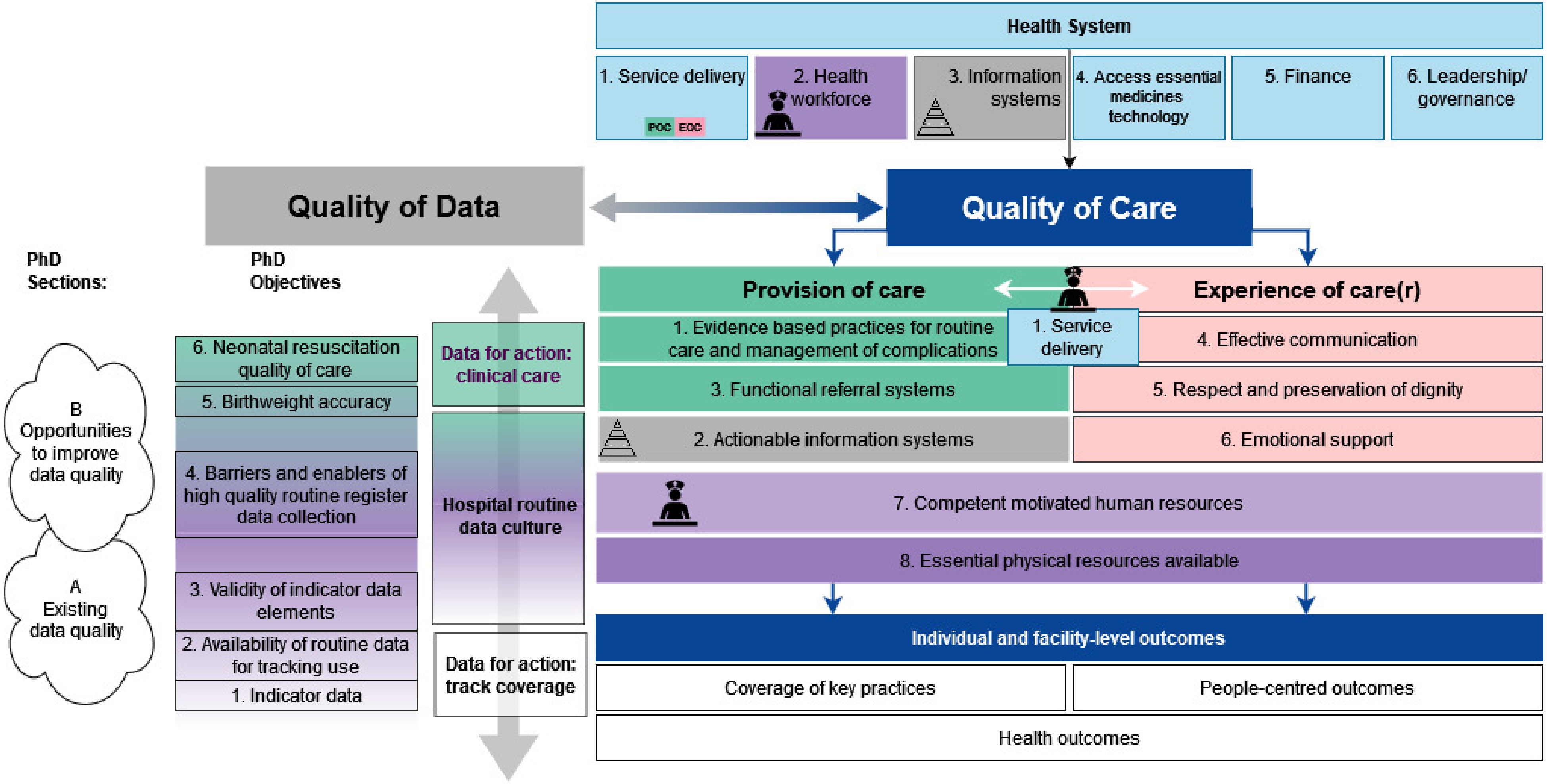
9.3.5.2 Linking data with clinical information

The documentation is both **data for action for clinical care** as well **as data for action to track coverage** for programme and policy makers (Figure 14). For example, the health professional's "data" documentation of a baby's low birthweight in case notes / routine register highlights to the clinical team to provide the extra care needed (e.g., warmth and feeding support) – data for clinical care. That same low birthweight "data" is available for aggregation and monthly reporting of health facility low birthweight rate.

Whilst individual health system building blocks and quality domains are distinct entities, how they interact and work together or against one another – in tension or traction. My novel '**Quality of Care *and Quality of Data* conceptual framework'** (Figure 24) illustrates this relationship of tension or traction between data quality and quality of care. This framework is adapted from the WHO

Quality of care framework to conceptualise the relationship between *quality of data* and *quality of care*.[5] Within my '**Quality of Care and Quality of Data conceptual framework',** all six health system building blocks are shown among which three overlap with the process quality domains. Service delivery (building block 1) is placed between the dimensions of provision and experience of care, health workforce (building block 2) aligns with quality domain 7 competent motivated human resources and information systems (building block 3) with quality domain 2. The relationship of these three components is my topic of interest – i) the quality of *data* captured regarding ii) *care provided* by iii) frontline *health professionals* responsible for both service provision and data at the point of care. Person-centred healthcare focuses experience of care for woman and newborns. The experience of midwives, nurses and doctors as frontline health professionals in the high mortality settings represented in this PhD thesis is important for the agenda of improving data quality and quality of care, hence in the '**Quality of Care and Quality of Data conceptual framework'** I have modified "Experience of care" to "Experience of care**R**".

Figure 24: Quality of Care and Data conceptual framework



Note: My PhD structure links Quality of data to three of the eight domains of the WHO Quality of Care framework (Tunçalp Ö, Were W, MacLennan C, et al. Quality of care for pregnant women and newborns—the WHO vision. Bjog 2015; 122(8): 1045-9): *Domain 1* - evidence based practices for routine care and management of complications, *Domain 2* - actionable information systems, *Domain 7* - competent and motivated human resources.

9.4 Strengths and limitations of the dissertation

The strength of this thesis is the synthesis of evidence it provides on assessment of routine labour and delivery register data quality, and opportunities to improve data quality using data from five EN-BIRTH study hospitals.[87] This work is important because routine labour and delivery ward register data for newborns and women are typically completed on every birth but are currently understudied.

The EN-BIRTH study was an ambitious research study assessing validity of newborn indicator measurement for two data sources: woman report for population-based household studies and health facility routine registers for RHIS. EN-BIRTH captured practices for every newborn including birthweight measurement and for the subset of newborns requiring life-saving interventions including bag-mask ventilation. This required a large number (>23,000) of observations on the labour and delivery ward, interviews with women and register data extraction. The team of investigators was collaborative between research institutions in four countries supported by a wide expert advisory group globally and nationally. Each research institution was responsible for leading specific components of the research. The EN-BIRTH published papers sought for equitable authorship with joint first and last authors from all collaborating investigators in the overall validation results manuscript (PhD Thesis Chapter 5, Objective 3) and thirteen of the fourteen papers in the supplement including (and PhD Thesis Chapter 7 and 8, Objectives 5 and 6).

My PhD thesis goes beyond the EN-BIRTH research design to integrate, synthesis and contextualise further findings. The evidence from this research was used to relate quality of data with quality of care with a focus on the frontline health professional who captures that data leading to the design of the novel '**Quality of Care and Quality of Data conceptual framework'.**

Practical suggestions for how to enable a data quality and data use virtuous cycle are described. A further strength of this work is the dual perspective I can bring from my experience of designing and running a maternal and newborn routine data system.[112]

Limitations of each study are described within the published papers. This thesis also enables further cross-cutting description of limitations which include:

Generalisability:

The EN-BIRTH study utilised a multi-country design to include the regions of the world that contribute most to global neonatal and maternal mortality and stillbirths: Africa and south Asia. The five selected EN-BIRTH hospitals were large CEmONC hospitals. All five hospitals were or had been used as research sites. One was a national referral hospital, another a regional hospital, two were district hospitals and one was a training centre in the capital city with a very high caesarean section rate but referred outpatients with complications. Our findings regarding routine register data quality therefore may not necessarily generalised to other levels of health facilities in the health system and other geographies. It is plausible that routine registers in other settings are more or less accurate than we found. For this PhD thesis I attempted to mitigate the small number of hospitals in EN-BIRTH by referring to some results from the Tanzania EN-MINI-PRISM Tools pilot assessment in 2021 which assessed newborn indicator data quality in 16 health facilities at different levels of the health system.[3]

The variation of register recorded accuracy of the indicators assessed between the study hospitals, indicates we can make a generalised statement that routine labour and delivery register data can be accurate, but we cannot assume it is accurate. As the PRISM framework illustrates, accurate data collection processes are influenced by multiple determinants of input and process.

Methodology: The EN-BIRTH study protocol was designed collaboratively, with multiple investigators in multiple countries. I acknowledge that whilst I joined the team after the data collection design was complete, I designed most aspects of the analysis plan and oversaw the analyses and technical aspect of these papers as well as the EN-BIRTH supplement.

Whilst using observation as an external reference was a strength of the study, there were challenges ensuring consistency across such a large study. Data collector training was country-led, and the standard training materials and data dictionary/standard operating procedure were high-level. The challenges with the observational definition of the early initiation of breast feeding which has already been extensively discussed in the manuscript as well as in section 9.3.3.1.1. Observation for many births in the EN-BIRTH study ceased less than one hour after birth, the sample size for early initiation of breast feeding was reduced, as the indicator definition includes the timing component of *within one hour*.

The register data exaction tool captured generic options (1-Yes, 2-No, 8-Not readable, 9-Not recorded, sometimes included 5-Not possible to record) instead of extracting the "raw" data from the registers e.g., tick/ cross/ blank/ drug name etc. This limited our capacity to do further analyses for register design and data quality.

The low kappa interobserver kappa scores for both observation and especially routine register data extraction (PhD chapter objective 3) were concerning. The manuscript describes the possible statistical explanations identified but future validation studies may consider calculating interobserver scores during data collector training for data quality assurance. In two current register review research projects I am leading, double data extraction during data collector training is proving to be very useful.

Future validation studies may also consider designing the matrices of data elements to be compared between observation and register during study design and interim analysis during tool piloting.

There was a lot of heterogeneity in results between the five EN-BIRTH study groups which we sought to address these through statistical measures and added measures of heterogeneity that had not been anticipated in the protocol. Further post adhoc analyses that had not been planned during the protocol development included comparisons between vaginal and caesarean birth.

The finding of diurnal variability in birth weight measurement with heaping being larger at night was identified towards the end of the study so we were not able to explore this dimension with other selected indicators. From the interviews with health professionals in (PhD Chapter 6, Objective 4) health workers described the effect of workload on documentation. Future validation studies may consider exploring how measures of busyness such as number of births per day and staff ratios were not influence quality of routine documentation.

Terminology: At the time of writing the manuscripts I used the term *low- and middle-income countries (LMIC)* liberally but have subsequently learned from colleagues from these settings to be careful with my phrasing and use more geographically specific language.

I used the terminology *"data collection"* associated with health professional documentation in routine registers throughout the manuscripts e.g., "Data extracted from five CEmONC hospitals show that a large amount of data are being collected in labour ward registers”[24] and “Many data are already being collected by frontline health professionals”.[25] Subsequently I have started to use the term *“data capture”* in relation to health professionals routine documentation tasks to emphasise that health professionals primary task is provision of care and documentation of that care for various uses. The term *data collection* is perhaps best reserved for dedicated staff such as researchers or RHIS staff whose primary role is to gather data.

9.5 Disciplinary implications (e.g., practice, education, leadership, and/or policy, and research)

Disciplinary implications arising from this PhD thesis' assessment of labour and delivery routine register data quality and opportunities identified to improve indicator data quality for use in high mortality settings. These have been described throughout the thesis and are summarised using the PRISM framework in

Table 3.[6]

The research has been disseminated in open-access peer reviewed journals, as oral presentations in conferences and cited in policy documents as listed in Appendix 5.

Research projects that led on from this research included the EN-BIRTH 2 study which developed the EN-MINI Tools. I am currently the Co-PI for the Improving Quality and Use of Newborn Indicators (IMPULSE) phase 1 study which is using the EN-MINI Tools to explore neonatal data quality in regions in Central African Republic, Ethiopia, Tanzania and Uganda. IMPULSE phase 2 seeks to test interventions to improve data quality which may include testing some of the opportunities proposed in this PhD Thesis.

Table 3: PhD Thesis disciplinary implications organised by the PRISM conceptual framework domains.

PRISM framework		Implications from thesis				
Determinants	**Examples of factors identified**	**Practice**	**Education**	**Leadership**	**Policy**	**Research**
Technical Factors	Complexity of reporting forms, procedures, HIS design.	Streamline standardised registers linked to efficient case notes. Label register columns in alignment with indicator definition.	Explore optimal interprofessional education for medical records system education.	Link ministry RHIS and maternal and newborn health departments to co-create updates to HIS design.	Standardised processes for register filling.	What person-centred routine register designs capture high quality data? What proportion of health professional time is spent on data and care?
Organisational Factors	Training. supportive supervision.	Use standardised processes for register filling.	Pre-service, in-service for all health workers capturing routine data.			What educational resources can enable the virtuous cycle of data quality and data use?
Promotion of culture of information		Promote data quality. Promote use of information.	Promote efforts and activities to change things for the better.	Promote evidence-based decision making and accountability.	Promote reward mechanisms for good work.	How to strengthen the hospital culture of information?
Availability of resources			Enable skills of supplied registers.	Digital scales for birthweight measurement.	Test data task-shifting of monthly report aggregation to increase health workers patient contact time.	

PRISM framework		Implications from thesis				
Determinants	**Examples of factors identified**	**Practice**	**Education**	**Leadership**	**Policy**	**Research**
Behavioural factors	Data quality checking skills -	Daily use of routine register for handover, weekly meetings, clinical audit	Health professional educational activities for RHIS linking to quality of care, including standardized birthweight measurement, stillbirth birthweight, stillbirth resuscitation	Value frontline health professionals for data capture in registers. Recognize the data burden on health workers.		What educational resources can enable support health professionals RHIS behavioural factors?
RHIS processes	Data collection, transmission, analysis, feedback	Feedback analytical reports and visualisations to health facilities		Enable the virtual cycle of data use and data quality	Reduce burden on the system of data transmitted and not used.	How to facilitate a data network within networks of care?
Data quality		Data quality assessments implemented in routine practice	Educational activities to strengthen data confidence and competence	Assess data quality	Improve Data quality	Can EN-MINI Tools guide priority actions to improve data quality and use?
Data use		Use RHIS data for decision making at health facility. Consider data quality may be suboptimal when used for RHIS tracking, MPDSR and research studies	Educational activities to use RHIS data for decision making	Use RHIS data for decision making	Use RHIS data for decision making	Can EN-MINI Tools guide priority actions to improve data quality and use?

PRISM framework		Implications from thesis				
Determinants	**Examples of factors identified**	**Practice**	**Education**	**Leadership**	**Policy**	**Research**
Cross-cutting						Can a co-designed standardised multi-purpose maternal and newborn health facility routine medical records system facilitate fidelity to clinical-guideline decisions, effective communication, improve data quality and quality of care?

9.6 Reflection

This PhD thesis has provided an opportunity as a researcher to assess routine labour and delivery register data quality and explore opportunities to improve data quality in high mortality settings. This has advanced my previous frontline experience as a health professional with responsibilities designing and running a hospital management information system, enabling data for clinical use, audit, and research.[23,141-156]

My learning has been across all four domains of the VITAE researcher developer framework[157]:

A. Knowledge and intellectual ability
B. Personal effectiveness skills
C. Knowledge of research and organization
D. Engagement, influence and impact

This PhD thesis has blended my experience as a clinician researcher obstetric-paediatrician with new research skills across the research cycle and specifically co-ordinating research activities, developing analysis plans, overseeing analyses and communicating results. I have seen how these have fed into wider policy and programme discussions and informed the work of EN-BIRTH 2 and the EN-MINI Tools.

9.7 Conclusion

This PhD thesis has explored hospital labour and delivery routine register data quality and opportunities to improve measurement of newborn indicators in high mortality settings using six research chapters conducted as part of the EN-BIRTH study in Bangladesh, Nepal and Tanzania are presented.

There is national and global interest in using high quality data for tracking progress for women, newborns, and stillbirths around the time of birth. Without improving quality of care for women and newborns, country targets of reducing neonatal mortality will not be reached.

Labour and delivery routine registers capture the source data for which are then transmitted up the data-information pyramid to build the RHIS regarding intrapartum care and outcomes for women, stillbirths, and newborns. This thesis presents evidence from five hospitals in Bangladesh, Nepal and Tanzania shows that shows we cannot assume routine labour and delivery register data in high mortality settings are an accurate representation of coverage of care and outcomes for newborns, stillbirths and women within hospitals.

This thesis identified enablers and barriers faced by frontline maternal and newborn health professionals responsible for capturing this hospital register data can be grouped around themes of register design, register filling, and register data use. The existing labour and delivery routine registers captured a large amount of mixed-quality data regarding intervention coverage. Register design varied by country and filling practices were not standardised. Caesarean birth affected newborn indicator measurement. Aggregate data were transmitted up the data-information pyramid but feedback loops regarding data quality and use were not common.

Public health demand for coverage indicator data has resulted in suboptimal systems that are burdensome for frontline health professionals who need to prioritise improving quality of care. Attention in recent years has been establishing data systems higher in the data-information pyramid and the quality of the source data in the registers at health facility level has been less of a priority. Unless the required data are clinically relevant to health professionals there is no incentive for high quality data capture.

The novel '**Quality of Care and Quality of Data conceptual framework'** presented in this thesis shows how labour and delivery register data quality links to the WHO domains of quality of care. Central to this is the missing link expressed as the **'Data Quality Continuum'** – the hospital routine data culture determines quality of neonatal data that are used both for clinical care and to track outcomes.

Labour and delivery register data quality was affected by inputs, processes, and outputs of the PRISM framework, which included register design, filling, and use. Existing hospital routine data culture promotes data quality by completeness rather that correctness leads to overestimations of care practices.   Opportunities to improve labour and delivery register data quality are described in this thesis including technical, organisation and behavioural factors of the PRISM framework. These findings align with other research regarding RHIS data quality at different levels of the data-information pyramid in LMIC settings.

This PhD thesis has generated recommendations for policy, practice, and research. This includes a priority to improve data quality for use is to standardise and streamline medical records, building interoperability between routine registers and clinical care notes. These documents are used every day to capture care provided in every labour and delivery ward. With rapid expansion of electronic RHIS in many contexts improving the source data, typically paper-based in high mortality settings is vital at this stage of the paper-digital transition to ensure we do not generate and use “too much poor-quality data too soon”.

Perfect data are perfectly valid, unobtrusive for data collection and perfectly cheap.[158] Perfect routine data does not exist, but at minimum data needs to be “good-enough” for use both for clinical care and tracking coverage for action to accelerate progress for globally agreed goals.

This PhD thesis proposes that enabling the virtuous cycle of data use and data quality for women and newborns is needed at the health facility level. This needs to include a focus on the person-centredness for the health professional, often caught in the tension between providing care or capturing data. The WHO quality of care framework included the novel dimension of *experience of care* for the woman and newborn. My PhD thesis proposes we need also to improve the *experience of the care**R***, by transforming health facility actionable information systems to be more user-centred and efficient for frontline midwives, nurses and doctors. This will give health professionals time to provide higher quality care and capture higher quality data for action for clinical decisions, quality improvement as well as national and global tracking. Creating traction between data and care in high mortality settings is necessary to enable high quality data for use at all levels of the data-information pyramid for high quality care for women and newborns.

Chapter 10 – References

1. Day LT, Ruysen H, Gordeev VS, et al. "Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. *Journal of Global Health*, 2019. http://jogha.org/documents/issue201901/jogh-09-010902.htm (accessed 18 August 2022).
2. Demographic and Health Surveys. The DHS program. https://dhsprogram.com/ (accessed 18 May 2020).
3. Data4Impact. Every Newborn-Measurement Improvement for Newborn & Stillbirth Indicators EN-MINI-PRISM Tools for Routine Health Information Systems: Tanzania Pilot Study Report – February 2023. 2023. https://www.data4impactproject.org/publications/every-newborn-measurement-improvement-for-newborn-stillbirth-indicators-en-mini-prism-tools-for-routine-health-information-systems-tanzania-pilot-study-report/ (accessed 8 July 2023).
4. UNICEF. Multiple Indicator Cluster Surveys,. http://mics.unicef.org/ (accessed 18 May 2020).
5. Tunçalp Ö, Were W, MacLennan C, et al. Quality of care for pregnant women and newborns—the WHO vision. *Bjog* 2015; **122**(8): 1045-9.
6. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. *Health Policy Plan* 2009; **24**(3): 217-28.
7. World Health Organization. SCORE for health data technical package: global report on health data systems and capacity, 2021 (accessed 6 December 2021).
8. World Health Organization. Data quality review: a toolkit for facility data quality assessment. Module 1: Framework and metrics. Geneva, 2017.
9. Hotchkiss DR, Aqil A, Lippeveld T, Mukooyo E. Evaluation of the performance of routine information system management (PRISM) framework: evidence from Uganda. *BMC health services research* 2010; **10**: 1-17.
10. MEASURE Evaluation. Data Demand and Use. https://www.measureevaluation.org/resources/tools/health-information-systems/prism (accessed 30 August 2023).
11. Heywood Arthur, Rohde Jon. Using information for action - a manual for health workers at facility level. University of Western Cape/HISP/MSH/EQUITY Project.
12. World Health Organization. Guidelines for the development of health management information systems 1993. https://iris.wpro.who.int/bitstream/handle/10665.1/5449/9290611065_eng.pdf (accessed 17 March 2021).
13. World Health Organization. Everybody's business--strengthening health systems to improve health outcomes: WHO's framework for action. 2007.
14. MEASURE Evaluation. Routine Health Information Systems - fact sheet. 2016. https://www.measureevaluation.org/resources/publications/fs-16-187.html (accessed 28 October 2022).
15. Merriam-Webster. https://www.merriam-webster.com/dictionary/traction (accessed 2 July 2023).
16. Munos MK, Blanc AK, Carter ED, et al. Validation studies for population-based intervention coverage indicators: design, analysis, and interpretation. *J Glob Health* 2018; **8**(2): 020804.
17. World Health Organization. Data quality review: a toolkit for facility data quality assessment. Module 2: Desk review of data quality2017. https://apps.who.int/iris/handle/10665/259225 (accessed 18 May 2020).
18. World Health Organization. Improving data quality, 2003.
19. MEASURE Evaluation. Routine Data Quality Assessment Tool - User Manual.2017. https://www.measureevaluation.org/resources/publications/ms-17-117 (accessed 25 June 2020).

20. Committee on Quality of Health Care in America Institute of Medicine Staff. Crossing the quality chasm: a new health system for the 21st century: National Academies Press; 2001.
21. World Health Organization. Quality of care: a process for making strategic choices in health systems: World Health Organization; 2006.
22. Darwin Holmes AG. Researcher Positionality - A Consideration of Its Influence and Place in Qualitative Research - A New Researcher Guide. *Shanlax International Journal of Education* 2020; **8**(4): 1-10.
23. Day LT, Alam S, Mondal N, Mollick B. Implementation of an Integrated Hospital Information System in Limited-resource Setting in Rural Bangladesh. ASCON 13th Annual Scientific Conference, icddr,b, . Dhaka, Bangladesh; 2011.
24. Day LT, Gore-Langton GR, Rahman AE, et al. Labour and delivery ward register data availability, quality, and utility - Every Newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries. *BMC Health Serv Res*, 2020. https://doi.org/10.1186/s12913-020-5028-7 (accessed 18 August 2022).
25. Day LT, Rahman QS, Rahman AE, et al. Assessment of the validity of the measurement of newborn and maternal health-care coverage in hospitals (EN-BIRTH): an observational study. *The Lancet Global Health* 2021; **9**(3): E267-79.
26. Shamba D, Day LT, Zaman SB, et al. Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* 2021; **21**(Suppl 1): 233.
27. Kong S, Day LT, Zaman SB, et al. Birthweight: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth* 2021; **21**(Suppl 1): 240.
28. Kc A, Peven K, Ameen S, et al. Neonatal resuscitation: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth*, 2021. https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03422-9.pdf (accessed 13 April 2021).
29. United Nations Inter-agency Group for Child Mortality Estimation (UNIGME). Levels & Trends in Child Mortality Report 2021, Estimates developed by the UN Inter-agency Group for Child Mortality Estimation. 2021.
30. Lawn JE, Blencowe H, Waiswa P, et al. Stillbirths: rates, risk factors, and acceleration towards 2030. *Lancet* 2016; **387**(10018): 587-603.
31. Hug L, Alexander M, You D, Alkema L. National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. *The Lancet Global Health* 2019; **7**(6): e710-e20.
32. World Health Organization, UNICEF. Every Newborn: an action plan to end preventable deaths2014. https://apps.who.int/iris/handle/10665/127938 (accessed 18 May 2020).
33. United Nations. Transforming our world: The 2030 agenda for sustainable development. 2016.
34. World Health Organization. Global Strategy for Women's, Children's and Adolescents' Health, 2016-2030.2015. https://www.everywomaneverychild.org/wp-content/uploads/2016/11/EWEC_globalstrategyreport_200915_FINAL_WEB.pdf (accessed 18 May 2020).
35. United Nations Inter-agency Group for Child Mortality Estimation (UNIGME). Levels & Trends in Child Mortality Report 2022, Estimates developed by the UN Inter-agency Group for Child Mortality Estimation2022. https://data.unicef.org/resources/levels-and-trends-in-child-mortality/ (accessed.
36. United Nations Inter-agency Group for Child Mortality Estimation (UNIGME). Report of the United Nations Inter-agency Group for Child Mortality Estimation. Never Forgotten - The situation of stillbirth around the globe2022. https://data.unicef.org/resources/never-forgotten-stillbirth-estimates-report/ (accessed 22 January 2023).

37. World Health Organization. Improving maternal and newborn health and survival and reducing stillbirth: progress report 2023.2023. https://www.who.int/publications/i/item/9789240073678 (accessed 13 May 2023).
38. WHO, UNICEF, UNFPA, WORLD BANK GROUP, UNDESA/Population Division. Trends in Maternal Mortality 2000 to 2020: Estimates. 2023.
39. Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? *Lancet* 2005; **365**(9462): 891-900.
40. World Health Organization. Strategies towards ending preventable maternal mortality (EPMM), 2015.
41. United Nations. Sustainable Development Goal 17. https://sdgs.un.org/goals/goal17 (accessed 15 July 2020).
42. Kyu HH, Abate D, Abate KH, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2018; **392**(10159): 1859-922.
43. Olson K. What Are Data? *Qual Health Res* 2021; **31**(9): 1567-9.
44. Ackoff RL. From data to wisdom. *Journal of applied systems analysis* 1989; **16**(1): 3-9.
45. Health Metrics Network, World Health Oganization. Framework and standards for country health information systems. 2nd ed ed. Geneva: World Health Organization; 2008.
46. World Health Organization. Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies 2010. https://www.who.int/healthinfo/systems/WHO_MBHSS_2010_full_web.pdf (accessed 24 March 2021).
47. World Health Organization. Analysing and using routine data to monitor the effects of COVID-19 on essential health services: practical guide for national and subnational decision-makers: interim guidance2021. https://apps.who.int/iris/bitstream/handle/10665/338689/WHO-2019-nCoV-essential_health_services-monitoring-2021.1-eng.pdf (accessed 24 March 2021).
48. UNAIDS. An introduction to indicators 2010. https://www.unaids.org/sites/default/files/sub_landing/files/8_2-Intro-to-IndicatorsFMEF.pdf (accessed 17 May 2021).
49. World Health Organization. Global reference list of 100 core health indicators (plus health-related SDGs)2018. https://apps.who.int/iris/bitstream/handle/10665/259951/WHO-HIS-IER-GPM-2018.1-eng.pdf (accessed 28 October 2022).
50. Mason E, McDougall L, Lawn JE, et al. From evidence to action to deliver a healthy start for the next generation. *The Lancet* 2014; **384**(9941): 455-67.
51. Moxon SG, Ruysen H, Kerber KJ, et al. Count every newborn; a measurement improvement roadmap for coverage data. *BMC Pregnancy Childbirth* 2015; **15 Suppl 2**: S8.
52. Bhutta ZA, Das JK, Bahl R, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? *The Lancet* 2014; **384**(9940): 347-70.
53. Vesel L, Bergh AM, Kerber KJ, et al. Kangaroo mother care: a multi-country analysis of health system bottlenecks and potential solutions. *BMC Pregnancy Childbirth* 2015; **15 Suppl 2**: S5.
54. Medhanyie AA, Alemu H, Asefa A, et al. Kangaroo Mother Care implementation research to develop models for accelerating scale-up in India and Ethiopia: study protocol for an adequacy evaluation. *BMJ Open* 2019; **9**(11): e025879.
55. Stefani G, Skopec M, Battersby C, Harris M. Why is Kangaroo Mother Care not yet scaled in the UK? A systematic review and realist synthesis of a frugal innovation for newborn care. *BMJ Innovations* 2022; **8**(1): 9-20.
56. Niermeyer S, Little GA, Singhal N, Keenan WJ. A Short History of Helping Babies Breathe: Why and How, Then and Now. *Pediatrics* 2020; **146**(Supplement 2): S101-S11.
57. Perlman JM, Velaphi S, Massawe A, Clarke R, Merali HS, Ersdal H. Achieving Country-Wide Scale for Helping Babies Breathe and Helping Babies Survive. *Pediatrics* 2020; **146**(Suppl 2): S194-S207.

58. Maina I, Wanjala P, Soti D, Kipruto H, Droti B, Boerma T. Using health-facility data to assess subnational coverage of maternal and child health indicators, Kenya. *Bull World Health Organ* 2017; **95**(10): 683-94.
59. Doctor HV, Radovich E, Benova L. Time trends in facility-based and private-sector childbirth care: analysis of Demographic and Health Surveys from 25 sub-Saharan African countries from 2000 to 2016. *J Glob Health* 2019; **9**(2): 020406.
60. Dossa NI, Philibert A, Dumont A. Using routine health data and intermittent community surveys to assess the impact of maternal and neonatal health interventions in low-income countries: A systematic review. *Int J Gynaecol Obstet* 2016; **135 Suppl 1**: S64-S71.
61. United Nations Children's Fund. The State of the World's Children 2023: For every child, vaccination, 2023. (accessed 5 July 2023).
62. World Health Organization. Survive and Thrive Transforming care for every small and sick newborn2019. https://apps.who.int/iris/bitstream/handle/10665/326495/9789241515887-eng.pdf (accessed 25 Jan 2021).
63. World Health Organization. Standards for improving the quality of care for small and sick newborns in health facilities 2020. https://apps.who.int/iris/bitstream/handle/10665/334126/9789240010765-eng.pdf (accessed 23 November 2020).
64. Oslo: Health Information Systems Programme. *District Health Information System DHIS2,*, 2020. https://www.dhis2.org/ (accessed 6 December 2021).
65. Hagel C, Paton C, Mbevi G, English M, Clinical Information Network information systems interest g. Data for tracking SDGs: challenges in capturing neonatal data from hospitals in Kenya. *BMJ Glob Health* 2020; **5**(3): e002108.
66. Lundin R, Mariani I, Peven K, Day LT, Lazzerini M. Quality of routine health facility data used for newborn indicators in low- and middle-income countries: A systematic review. *Journal of Global Health* 2022; **12**.
67. Garrib A, Stoops N, McKenzie A, et al. An evaluation of the district health information system in rural South Africa. *S Afr Med J* 2008; **98**(7): 549-52.
68. Maiga A, Jiwani SS, Mutua MK, et al. Generating statistics from health facility data: the state of routine health information systems in Eastern and Southern Africa. *BMJ Glob Health* 2019; **4**(5): e001849.
69. Chiba Y, Oguttu MA, Nakayama T. Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. *Midwifery* 2012; **28**(3): 329-39.
70. Bhattacharya AA, Allen E, Umar N, et al. Monitoring childbirth care in primary health facilities: a validity study in Gombe State, northeastern Nigeria. *J Glob Health* 2019; **9**(2): 020411.
71. Duffy S, Crangle M. Delivery room logbook–fact or fiction? *Tropical doctor* 2009; **39**(3): 145-9.
72. Maternal and Child Survival Program. What Data on Maternal and Newborn Health do National Health Management Information Systems include? A review of data elements for 24 low- and lower middle income countries 2018. https://www.mcsprogram.org/resource/what-data-on-maternal-and-newborn-health-do-national-health-management-information-systems-include/ (accessed 18 May 2020).
73. Faizi N, Kabue MM, Palestra F, Katwan E, Moran AC. Availability of priority maternal and newborn health indicators: Cross-sectional analysis of pregnancy, childbirth and postnatal care registers from 21 countries. *PLOS Global Public Health* 2023; **3**(1).
74. Hotchkiss DR, Aqil A, Lippeveld T, Mukooyo E. Evaluation of the performance of routine information system management (PRISM) framework: evidence from Uganda. *BMC health services research* 2010; **10**(1): 188.
75. Braa J, Heywood A, Sahay S. Improving quality and use of data through data-use workshops: Zanzibar, United Republic of Tanzania. *Bulletin of the World Health Organization* 2012; **90**: 379-84.

76. Sæbø JI, Mesheck Moyo C, Nielsen P. Promoting transparency and accountability with district league tables in Sierra Leone and Malawi. *Health Policy and Technology* 2018; **7**(1): 35-43.
77. Rendell N, Lokuge K, Rosewell A, Field E. Factors that influence data use to improve health service delivery in low-and middle-income countries. *Global Health: Science and Practice* 2020; **8**(3): 566-81.
78. Kumar M, Gotz D, Nutley T, Smith JB. Research gaps in routine health information system design barriers to data quality and use in low- and middle-income countries: A literature review. *Int J Health Plann Manage* 2018; **33**(1): e1-e9.
79. Stanton CK, Rawlins B, Drake M, et al. Measuring coverage in MNCH: Testing the validity of women's self-report of key maternal and newborn health Interventions during the peripartum period in Mozambique. *PLoS One* 2013; **8**(5): e60694.
80. Blanc AK, Warren C, McCarthy KJ, Kimani J, Ndwiga C, RamaRao S. Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya. *Journal of global health* 2016; **6**(1): 010405.
81. Blanc AK, Diaz C, McCarthy KJ, Berdichevsky K. Measuring progress in maternal and newborn health care in Mexico: validating indicators of health system contact and quality of care. *BMC pregnancy and childbirth* 2016; **16**(1): 255.
82. McCarthy KJ, Blanc AK, Warren CE, Kimani J, Mdawida B, Ndwidga C. Can surveys of women accurately track indicators of maternal and newborn care? A validity and reliability study in Kenya. *J Glob Health* 2016; **6**(2): 020502.
83. Goodwin S. Data rich, information poor (DRIP) syndrome: is there a treatment? *Radiology management* 1996; **18**(3): 45-9.
84. Darmstadt GL, Kinney MV, Chopra M, et al. Who has been caring for the baby? *The Lancet* 2014; **384**(9938): 174-88.
85. Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *The Lancet* 2014; **384**(9938): 189-205.
86. Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q* 2005; **83**(4): 691-729.
87. Lewis KB, Graham ID, Boland L, Stacey D. Writing a compelling integrated discussion: a guide for integrated discussions in article-based theses and dissertations. *Int J Nurs Educ Scholarsh* 2021; **18**(1).
88. Peven K, Day LT, Ruysen H, et al. Stillbirths including intrapartum timing: EN-BIRTH multi-country validation study. *BMC Pregnancy Childbirth*, 2021. https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03238-7.pdf (accessed 13 April 2021).
89. O'Neil S, Taylor S, Sivasankaran A. Data Equity to Advance Health and Health Equity in Low- and Middle-Income Countries: A Scoping Review. *Digit Health* 2021; **7**: 20552076211061922.
90. The London School of Hygiene & Tropical Medicine. Every Newborn Action Plan Metrics Design Workshop for Facility-based Testing of Coverage Metrics, Windsor. 2016. https://www.healthynewbornnetwork.org/hnn-content/uploads/ENAP-Metrics-Facility-based-Workshop-Report_April-2016_FINAL.pdf
91. Munos M, Arnold F, Blanc A, et al. Improving Coverage Measurement (ICM) for Maternal, Newborn, and Child Health.pdf; 2018.
92. Benova L, Moller A-B, Moran AC. “What gets measured better gets done better”: The landscape of validation of global maternal and newborn health indicators through key informant interviews. *Plos One*, e0224746, 2019. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0224746 (accessed.
93. World Health Organization. Data quality review: a toolkit for facility data quality assessment. Module 3: Data verification and system assessment. Geneva, 2017.
94. Every Newborn - Birth Indicators Research Tracking in Hospitals (EN-BIRTH) Study Group. Every Newborn BIRTH multi-country validation study; informing measurement of coverage and quality of maternal and newborn care - Supplement2021.

https://bmcpregnancychildbirth.biomedcentral.com/articles/supplements/volume-21-supplement-1 (accessed 2 May 2021).
95. Marsh AD, Muzigaba M, Diaz T, et al. Effective coverage measurement in maternal, newborn, child, and adolescent health and nutrition: progress, future prospects, and implications for quality health systems. *The Lancet Global Health* 2020; **8**(5): e730-e6.
96. Blencowe H, Krasevec J, de Onis M, et al. National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. *Lancet Glob Health* 2019; **7**(7): e849-e60.
97. Day LT, Moran AC, Jackson D, et al. Survive and Thrive Transforming care for every small and sick newborn, Chapter 5 Data for Action.2019. https://apps.who.int/iris/bitstream/handle/10665/326495/9789241515887-eng.pdf (accessed 30 August 2023).
98. Patterson JK, Aziz A, Bauserman MS, McClure EM, Goldenberg RL, Bose CL. Challenges in classification and assignment of causes of stillbirths in low- and lower middle-income countries. *Semin Perinatol* 2019; **43**(5): 308-14.
99. Blencowe H, Okwaraji Y, Hug L, You D. Stillbirth Definition and Data Quality Assessment for Health Management Information Systems (HMIS), a guideline.2022. https://data.unicef.org/resources/stillbirth-definition-and-data-quality-assessment-for-health-management-information-systems/ (accessed 13 May 2023).
100. World Health Organization. Evaluating the quality of care for severe pregnancy complications: the WHO near-miss approach for maternal health.2011. https://apps.who.int/iris/bitstream/handle/10665/44692/9789241502221_eng.pdf;jsessionid=5860A65424A06977E92D34A11CDBBBB4?sequence=1 (accessed 30 August 2023).
101. Medeiros PB, Bailey C, Andrews C, Liley H, Gordon A, Flenady V. Neonatal near miss: A review of current definitions and the need for standardisation. *Aust N Z J Obstet Gynaecol* 2022; **62**(3): 358-63.
102. Ronsmans C, Cresswell JA, Goufodji S, et al. Characteristics of neonatal near miss in hospitals in Benin, Burkina Faso and Morocco in 2012-2013. *Tropical medicine & international health : TM & IH* 2016; **21**(4): 535-45.
103. Gladstone ME, Salim N, Ogillo K, et al. Birthweight measurement processes and perceived value: qualitative research in one EN-BIRTH study hospital in Tanzania. *BMC Pregnancy Childbirth*, 2021. https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03356-2.pdf (accessed 31 December 202).
104. Daelmans B, Darmstadt GL, Lombardi J, et al. Early childhood development: the foundation of sustainable development. *The Lancet* 2017; **389**(10064): 9-11.
105. World Health Organization. Indicators for assessing infant and young child feeding practices: definitions and measurement methods.2021. https://apps.who.int/iris/bitstream/handle/10665/340706/9789240018389-eng.pdf?sequence=1 (accessed 14 May 2023).
106. Brizuela V, Tuncalp O. A road to optimising maternal and newborn quality care measurement for all. *Lancet Glob Health* 2021; **9**(3): e221-e2.
107. Muzigaba M, Chitashvili T, Choudhury A, et al. Global core indicators for measuring WHO's paediatric quality-of-care standards in health facilities: development and expert consensus. *BMC Health Serv Res* 2022; **22**(1): 887.
108. Day Louise Tina, Kathleen Hill, Debra Jackson, et al. Global core indicators for measuring WHO’s small and sick newborn quality-of-care standards in health facilities: development, global consultation, and expert consensus (in submission). 2023.
109. Berg M, Toussaint P. The mantra of modeling and the forgotten powers of paper: A sociotechnical view on the development of process-oriented ICT in health care. *Int J Med Inform* 2003; **69**(2-3): 223-34.
110. Tamrat T, Chandir S, Alland K, et al. Digitalization of routine health information systems: Bangladesh, Indonesia, Pakistan. *Bull World Health Organ* 2022; **100**(10): 590-600.

111. Kilkenny MF, Robinson KM. Data quality: "Garbage in - garbage out". *Health Inf Manag* 2018; **47**(3): 103-5.
112. Day LT, Alam S, Mondal N, Mollick B. Implementation of an Integrated Hospital Information System in Limited-resource Setting in Rural Bangladesh. *ASCON, Annual Scientific Conference - Oral Presenation*, 2011. (accessed.
113. MEASURE Evaluation. PRISM: Performance of Routine Information System Management Series 2019. https://www.measureevaluation.org/resources/tools/health-information-systems/prism (accessed 26 November 2020).
114. Nutley T, Reynolds HW. Improving the use of health data for health system strengthening. *Global health action* 2013; **6**: 20001.
115. Paul O'Connor, Beth Reid, Lee Ridoutt, Cliona O'Donovan, Bin Jalaludin (Jal), Ric Marshall. Moving data collection from a "vicious" cycle to a "virtuous" cycle – a cycle of continuous improvement.2022. https://www.linkedin.com/pulse/moving-data-collection-from-vicious-cycle-virtuous-paul-o-connor/ (accessed 30 August 2023).
116. Data4Impact. Every Newborn-Measurement Improvement for Newborn & Stillbirth Indicators (EN-MINI) Tools for Routine Health Information Systems. 2022. https://www.data4impactproject.org/resources/en-mini-tools/ (accessed 28 October 2022).
117. Lee J, Lynch CA, Hashiguchi LO, et al. Interventions to improve district-level routine health data in low-income and middle-income countries: a systematic review. *BMJ Glob Health* 2021; **6**(6).
118. Hoxha K, Hung YW, Irwin BR, Grepin KA. Understanding the challenges associated with the use of data from routine health information systems in low- and middle-income countries: A systematic review. *Health Inf Manag* 2022; **51**(3): 135-48.
119. Lemma S, Janson A, Persson LA, Wickremasinghe D, Kallestal C. Improving quality and use of routine health information system data in low- and middle-income countries: A scoping review. *PLoS One* 2020; **15**(10): e0239683.
120. Hung YW, Hoxha K, Irwin BR, Law MR, Grepin KA. Using routine health information data for research in low- and middle-income countries: a systematic review. *BMC Health Serv Res* 2020; **20**(1): 790.
121. Nutley T, Gnassou L, Traore M, Bosso AE, Mullen S. Moving data off the shelf and into action: an intervention to improve data-informed decision making in Cote d'Ivoire. *Global health action* 2014; **7**: 25035.
122. Villar J, Cheikh Ismail L, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet* 2014; **384**(9946): 857-68.
123. Sobhy S, Arroyo-Manzano D, Murugesu N, et al. Maternal and perinatal mortality and complications associated with caesarean section in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet* 2019; **393**(10184): 1973-82.
124. McClure EM, Saleem S, Goudar SS, et al. Stillbirth rates in low-middle income countries 2010 - 2013: a population-based, multi-country study from the Global Network. *Reprod Health* 2015; **12 Suppl 2**(Suppl 2): S7.
125. Harrison MS, Pasha O, Saleem S, et al. A prospective study of maternal, fetal and neonatal outcomes in the setting of cesarean section in low- and middle-income countries. *Acta Obstet Gynecol Scand* 2017; **96**(4): 410-20.
126. Fauveau V. New indicator of quality of emergency obstetric and newborn care. *The Lancet* 2007; **370**(9595).
127. United Nations Inter-agency Group for Child Mortality Estimation (UNIGME). A Neglected Tragedy - The global burden of stillbirths. Report of the United Nations Inter-agency Group for Child Mortality Estimation.2022. https://www.unicef.org/media/84851/file/UN-IGME-the-global-burden-of-stillbirths-2020.pdf (accessed 25 October 2020).
128. Bhutta ZA, Yakoob MY, Lawn JE, et al. Stillbirths: what difference can we make and at what cost? *The Lancet* 2011; **377**(9776): 1523-38.

129. Goldenberg RL, Saleem S, Pasha O, Harrison MS, McClure EM. Reducing stillbirths in low-income countries. *Acta Obstet Gynecol Scand* 2016; **95**(2): 135-43.
130. Peven K, Day LT, al e. EN-MINI Mapping tool for Routine Health Information Systems. 2022. https://www.data4impactproject.org/resources/en-mini-tools/map-newborn-data/ (accessed 14 May 2023.
131. Data4Impact. Mapped newborn data availability in routine Health Information Systems EN-MINI mapping tool results: Tanzania – June 2022. 2022. https://www.data4impactproject.org/publications/mapped-newborn-data-availability-in-routine-health-information-systems-en-mini-mapping-tool-results-tanzania-june-2022/ (accessed 15 May 2023).
132. Siyam A, Ir P, York D, et al. The burden of recording and reporting health data in primary health care facilities in five low- and lower-middle income countries. *BMC Health Serv Res* 2021; **21**(Suppl 1): 691.
133. Lopetegui M, Yen PY, Lai A, Jeffries J, Embi P, Payne P. Time Motion Studies in Healthcare: What are we talking about? *Journal of biomedical informatics* 2014; **0**: 292-9.
134. Gon G, Ali SM, Aunger R, et al. A Practical Guide to Using Time-and-Motion Methods to Monitor Compliance With Hand Hygiene Guidelines: Experience From Tanzanian Labor Wards. *Global Health: Science and Practice* 2020; **8**(4): 827-37.
135. Rowe AK, Rowe SY, Peters DH, Holloway KA, Chalker J, Ross-Degnan D. Effectiveness of strategies to improve health-care provider practices in low-income and middle-income countries: a systematic review. *Lancet Glob Health* 2018; **6**(11): e1163-e75.
136. Agyekum EO, Kalaris K, Maliqi B, Moran AC, Ayim A, Roder-DeWan S. Networks of care to strengthen primary healthcare in resource constrained settings. *The BMJ* 2023; **380**: e071833.
137. Belay H, Lippeveld T. Inventory of PRISM framework and tools: application of PRISM tools and interventions for strengthening routine health information system performance. *Chapel Hill*, 2013. https://www.measureevaluation.org/resources/publications/wp-13-138.html (accessed 27 August 2023).
138. Estifanos AS, Gezahegn R, Keraga DW, Kifle A, Procureur F, Hill Z. 'The false reporter will get a praise and the one who reported truth will be discouraged': a qualitative study on intentional data falsification by frontline maternal and newborn healthcare workers in two regions in Ethiopia. *BMJ Glob Health* 2022; **7**(4).
139. Day. LT, Peven. K, Lawn. JE. Birth by Caesarean Section: Who? What? Why? When? Secondary analyses of the Every Newborn –Birth Indicators Research Tracking in Hospitals (EN-BIRTH) Cohort. Report for the Bill and Melinda Gates Foundation, 2021.
140. UNFPA, ICM, WHO. The State of the World's Midwifery. USA, 2014.
141. Mussell F, Ferdous K, Lennox R, Day LT, Edwards CE. Perinatal Death Audit: A useful strategy to Improve Patient Management. 10th ASCODD, icddr,b, Oral Presentation. Dhaka, Bangladesh; 2003.
142. M.Mazumdar, LT.Day. Audit improves clinical practice: the management of neonatal jaundice. ASCON, Annual Scientific Conference, icddr,b, Poster presentation. Dhaka, Bangladesh; 2007.
143. Day LT, Jinks M. Structured Obstetric Clinical Sheets – the design and implementation of a useful clinical tool. Royal College Obstetrics & Gynaecologists / AICC / SAFOG South Asia Day - Poster Presentation. RCOG, London, UK; 2009.
144. Day. LT, Mussell. F, Khatun. H, et al. Sustaining perinatal audit in the high stillbirth rate setting – a 20 year journey in rural Bangladesh…… Poster presentation International Stillbirth Alliance Conference,. Cork, Ireland; 2017.
145. S.Piper, S.Razia, M.Lares, R.Zechariah, L.T.Day. Near-miss Maternal Death in Comprehensive Emergency Obstetric Care Facility in Bangladesh - Oral Presentation ASCON 13th Annual Scientific Conference, icddr,b, . Dhaka, Bangladesh; 2011.

146. L.T. Day, A.Winter, M Pietroni, et al. Improving Coverage of Neonatal Jaundice Management in the Limited-resource Setting - Oral Presentation ASCON 13th Annual Scientific Conference, icddr,b, . Dhaka, Bangladesh; 2011.
147. L.T. Day, S.Hasdak, R.Islam, et al. Programmatic Experience of Facility-based IMCI Implementation in Rural Bangladesh - Oral Presentation. ASCON 13th Annual Scientific Conference, icddr,b, . Dhaka, Bangladesh; 2011.
148. L.T. Day, F.Mussell, S.Alam, B.Mohanto, G.Singh. Mode of Delivery by Socioeconomic Status at a Comprehensive Emergency Obstetric Care Facility - Poster Presentation ASCON 13th Annual Scientific Conference, icddr,b, . Dhaka, Bangladesh; 2011.
149. L.T.Day, F.Mussell, B.Begum, H.Khatun, F.Yesmin, S.Razia. Monitoring Clinical Indications for Caesarean Section in the Limited-resource Setting - Poster Presentation ASCON 13th Annual Scientific Conference, icddr,b, . Dhaka, Bangladesh; 2011.
150. Solomon JS, Day LT, H.Edwards C, Mussell F. 18 Years of Uterine Rupture in a Rural Hospital in Bangladesh - Poster presentation. Royal College Obstetrics & Gynaecologists / AICC / SAFOG South Asia Day. RCOG, London, UK 2009.
151. Mussell F, Day LT, H.Edwards C. Maternal Death at a Rural Comprehensive Emergency Obstetric Care Facility in Bangladesh - Poster presentation Royal College Obstetrics & Gynaecologists / AICC / SAFOG South Asia Day. RCOG, London, UK 2009.
152. Day LT, Mussell F, H.Edwards C. A year in the life of a rural Comprehensive Emergency Obstetric Care Unit in Bangladesh - Poster presentation Royal College Obstetrics & Gynaecologists / AICC / SAFOG South Asia Day. RCOG, London, UK 2009.
153. Day LT, Saha S, Mussell F, Prenger KL, Edwards C. Improving the use of the partograph - Oral presentation. Global Maternal Health Conference. Delhi, India 2010.
154. Saha. S, Day. LT, Hunt. A, Edwards. C, Mussell. F, Prenger. K. Strengthening Caesarean Section services - Oral presentation. Global Maternal Health Conference. Delhi, India 2010.
155. Day. LT, Alam. MS, Sarkar. S. Structured Neonatal Clinical sheets – the design and implementation of a useful clinical tool. ASCON, Annual Scientific Conference, icddr,b, - Poster presentation. Dhaka, Bangladesh 2009.
156. Day LT, Mussell F, Lennox RY, L.Prenger K, Edwards C. Audit for stillbirths and neonatal deaths in Bangladesh - adapting from the South African audit system and lessons learned. International Stillbirth Alliance Conference - Oral presentation. South Africa; 2009.
157. VITAE. Researcher's Development Framework. 2010. https://www.vitae.ac.uk/ (accessed 1 January 2024.
158. Ram P. "Perfect data are perfectly valid, unobtrusive for data collection and perfectly cheap" 2019.
159. Shamba. D, Rahman. AE, Peterson. SS, Moran. A, Lawn. JE. Research innovation meets reality in Bangladesh, Nepal and Tanzania: measuring health system performance to reach every woman and every newborn. . Fifth Global Symposium on Health Systems Research Advancing health systems for all in the SDG era. Liverpool; 2018.
160. Day. LT, Group. E-BS. The Every Newborn – birth indicators research tracking in hospitals (EN-BIRTH). 2021.
161. H. Ruysen, Day. LT, T. Tahsina, P. ten Hoope-Bender, Lawn. J. Midwives are key to counting care at birth – The Every Newborn – birth indicators research tracking in hospitals (EN-BIRTH) observational study. 2021.
162. Holschneider S, Morgan A. Improving Monitoring Data Systems to Count and Account for Stillbirths. 2022.
163. World Health Organization, UNICEF, UNFPA. Born too soon: decade of action on preterm birth: World Health Organization; 2023.

Appendices

Appendix 1 – Ethical approval

EN-BIRTH study participants gave voluntary informed written consent prior to recruitment for observation and again for exit survey. This study was granted ethical approval by institutional review boards in all participating countries and the London School of Hygiene & Tropical Medicine (Table 4). An amendment was granted by the observational research ethics committee at LSHTM to link my PhD to the EN-BIRTH study (Figure 25).

All collaborating partners signed data sharing and transfer agreements. EN-BIRTH is registered number 4833 at https://www.researchregistry.com

Table 4: Ethical Approval for EN-BIRTH study

Country	Institutional Review Boards	Date	Number/Ref
UK	London School of Hygiene & Tropical Medicine (LSHTM) Interventions Research Ethics Committee	03/10/16	11780
Bangladesh	Icddr,b Research Review Committee Icddr,b Ethical Review Committee	11/08/16 14/11/16	PR 16055
Nepal	Nepal Health Research Council (NHRC)	08/08/16	187/2016
Tanzania	National Institute for Medical Research (NIMRI)	20/01/17	NIMR/HQ/R.8a/Vol IX/2394
	Ifakara Health Institute	20/10/16	IHI/IRB/No: 032-2016
	Muhimbili University of Health and Allied Sciences research and Publications committee	21/10/16	2016-10-21-/AEC/Vol.XI/310

Figure 25: EN-BIRTH Study LSHTM Research Ethics Committee Approval

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Research Ethics Committee

Prof Joy Lawn

Keppel Street
London
WC1E 7HT

11 June 2021

Dear Joy ,

Study Title: Testing the recording of priority facility-based, maternal and newborn coverage indicators for use in health management information systems

LSHTM Ethics ref: 11780 - 3

Thank you for submitting your amendment for the above research project.

Your amendment has been assessed by the Research Governance & Integrity Office and has been approved as a non-substantial change. The amendment does not require further ethical approval from the observational ethics committee.

List of documents reviewed:

Any subsequent changes to the application must be submitted to the Committee via an Amendment form on the ethics online applications website: http://leo.lshtm.ac.uk .

Best of luck with your project.

Yours sincerely,

[redacted]

Rebecca Carter

Research Governance Coordinator

Ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

Page 1 of 1

Appendix 2 - List of publications included in thesis

PhD Section A - Objective 1 - EN-BIRTH protocol

"Every Newborn-BIRTH" protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania.

Day, Louise T., **Ruysen, Harriet**, Gordeev, Vladimir S., Gore-Langton, Georgia R., Boggs, Dorothy, Cousens, Simon, Moxon, Sarah G., Blencowe, Hannah, Baschieri, Angela, Rahman, Ahmed Ehsanur, Tahsina, Tazeen, Zaman, Sojib Bin, Hossain, Tanvir, Rahman, Qazi Sadeq-ur, Ameen, Shafiqul, El Arifeen, Shams, Kc, Ashish, Shrestha, Shree Krishna, Kc, Naresh P., Singh, Dela, Jha, Anjani Kumar, Jha, Bijay, Rana, Nisha, Basnet, Omkar, Joshi, Elisha, Paudel, Asmita, Shrestha, Parashu Ram, Jha, Deepak, Bastola, Ram Chandra, Ghimire, Jagat Jeevan, Paudel, Rajendra, Salim, Nahya, Shamb, Donat, Manji, Karim, Shabani, Josephine, Shirima, Kizito, Mkopi, Namala, Mrisho, Mwifadhi, Manzi, Fatuma, Jaribu, Jennie, Kija, Edward, Assenga, Evelyne, Kisenge, Rodrick, Pembe, Andrea, Hanson, Claudia, Mbaruku, Godfrey, Masanja, Honorati, Amouzou, Agbessi, Azim, Tariq, Jackson, Debra, Kabuteni, Theopista John, Mathai, Matthews, Monet, Jean-Pierre, Moran, Allisyn, Ram, Pavani, Rawlins, Barbara, Sæbø, Johan Ivar, Serbanescu, Florina, Vaz, Lara, Zaka, Nabila and Lawn, Joy E.
Journal of Global Health 2019; **9**(1). https://doi.org/10.7189/jogh.09.010902 [1]

The supplementary material reference in the paper is available at
http://jogha.org/documents/issue201901/jogh-09-010902-s001.pdf

PhD Section A - Objective 2 - EN-BIRTH baseline register analysis

Labour and delivery ward register data availability, quality, and utility - Every Newborn - birth indicators research tracking in hospitals (EN-BIRTH) study baseline analysis in three countries.

Day, L. T., Gore-Langton, G. R., Rahman, A. E., Basnet, O., Shabani, J., Tahsina, T., Poudel, A., Shirima, K., Ameen, S., KC Ashish, Salim, N., Zaman, S. B., Shamba, D., Blencowe, H., Ruysen, H., El Arifeen, S., Boggs, D., Gordeev, V. S., Rahman, Q. S., Hossain, T., Joshi, E., Thapa, S., Poudel, R. P., Poudel, D., Chaudhary, P., Karki, R., Chitrakar, B., Mkopi, N., Wisiko, A., Kitende, A. P., Shirati, M. R., Chingalo, C., Semhando, A. O., Mtei, C., Mwenisongole, V., Bakuza, J. M., Kombo, J., Mbaruku, G. and Lawn, J. E.
BMC Health Serv Res 2020, **20**(1):737. https://doi.org/10.1186/s12913-020-5028-7 [24]

PhD Section A - Objective 3 - EN-BIRTH validity of coverage measurement

Every Newborn-BIRTH observational study to assess validity of newborn and maternal coverage measurement in hospitals.

Day, Louise Tina, Rahman, Qazi Sadequr, Rahman, Ahmed Ehsanur, Salim, Nahya, Kc, Ashish, Ruysen, Harriet, Tahsina, Tazeen, Masanja, Honorati, Basnet, Omkar, Gore-Langton, Georgia R., Zaman, Sojib Bin, Shabani, Josephine, Jha, Anjani Kumar, Gordeev, Vladimir Sergeevich, Ameen, Shafiqul, Shamba, Donat, Jha, Bijay, Boggs, Dorothy, Hossain, Tanvir, Shirima, Kizito, Bastola, Ram Chandra, Peven, Kimberly, Siddique, Abu Bakkar, Mbaruku, Godfrey, Paudel, Rajendra, Baschieri, Angela, Hossain, Aniqa Tasnim, Kong, Stefanie, Paudel, Asmita, Ahmed, Anisuddin, Cousens, Simon,

El Arifeen, Shams, Lawn, Joy E.,
The Lancet Global Health 2021; **9**(3): E267-79. https://doi.org/10.1016/S2214-109X(20)30504-0.[25]

PhD Section B - Objective 4 - EN-BIRTH Barriers and enablers for routine documentation

Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study.

Shamba, D., Day, L. T., Zaman, S. B., Sunny, A. K., Tarimo, M. N., Peven, K., Khan, J., Thakur, N., Talha, Mtus, K, C. A., Haider, R., Ruysen, H., Mazumder, T., Rahman, M. H., Shaikh, M. Z. H., Saebo, J. I., Hanson, C., Singh, N. S., Schellenberg, J., Vaz, L. M. E., Requejo, J., Lawn, J. E. and EN-BIRTH Study Group
BMC Pregnancy Childbirth, 2021; **21**(Suppl 1): 233. https://doi.org/10.1186/s12884-020-03517-3.[26]

PhD Section B - Objective 5 - EN-BIRTH validity of birthweight measurement

Birthweight: EN-BIRTH multi-country validation study.

Kong, S., **Day, L. T**., Zaman, S. B., Peven, K., Salim, N., Sunny, A. K., Shamba, D., Rahman, Q. S., K, C. A., Ruysen, H., El Arifeen, S., Mee, P., Gladstone, M. E., Blencowe, H., Lawn, J. E. and EN-BIRTH Study Group
BMC Pregnancy Childbirth. 2021 Mar 26;**21**(Suppl 1):240. https://doi.org/10.1186/s12884-020-03355-3.[27]

PhD Section B - Objective 6 - EN-BIRTH Validity of neonatal resuscitation measurement

Neonatal resuscitation: EN-BIRTH multi-country validation study

Kc, A., Peven, K., Ameen, S., Msemo, G., Basnet, O., Ruysen, H., Zaman, S. B., Mkony, M., Sunny, A. K., Rahman, Q. S., Shabani, J., Bastola, R. C., Assenga, E., Kc, N. P., El Arifeen, S., Kija, E., Malla, H., Kong, S., Singhal, N., Niermeyer, S., Lincetto, O., **Day, L. T., Lawn, J. E**. and EN-BIRTH Study Group. Neonatal resuscitation: EN-BIRTH multi-country validation study.

BMC Pregnancy Childbirth. vol. 21, 2021/03/27 edn; 2021: 235. https://doi.org/10.1186/s12884-020-03422-9[28]

Appendix 3 – List of additional related publications

As supplement technical editor for the **Every Newborn BIRTH multi-country validation study: informing measurement of coverage and quality of maternal and newborn care**[94] I made substantial contributions in the analysis and writing of these published manuscripts:

Joint senior author

- **Kangaroo mother care: EN-BIRTH multi-country validation study**
  Salim, N., Shabani, J., Peven, K., Rahman, Q. S., Kc, A., Shamba, D., Ruysen, H., Rahman, A. E., Kc, N., Mkopi, N., Zaman, S. B., Shirima, K., Ameen, S., Kong, S., Basnet, O., Manji, K., Kabuteni, T. J., Brotherton, H., Moxon, S. G., Amouzou, A., Hailegebriel, T. D., **Day, L. T.**, Lawn, J. E. and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):231. doi: 10.1186/s12884-020-03423-8*.
  https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03423-8.pdf (accessed 31 December 2022).

Co-Author

- **Stillbirths including intrapartum timing: EN-BIRTH multi-country validation study**
  Peven, K., **Day, L. T**., Ruysen, H., Tahsina, T., Kc, A., Shabani, J., Kong, S., Ameen, S., Basnet, O., Haider, R., Rahman, Q. S., Blencowe, H., Lawn, J. E. and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):226. doi: 10.1186/s12884-020-03238-7*.
  https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03238-7.pdf
  (accessed 31 December 2022).

- **Uterotonics for prevention of postpartum haemorrhage: EN-BIRTH multi-country validation study**
  Ruysen, H., Shabani, J., Hanson, C., **Day, L. T**., Pembe, A. B., Peven, K., Rahman, Q. S., Thakur, N., Shirima, K., Tahsina, T., Gurung, R., Tarimo, M. N., Moran, A. C., Lawn, J. E. and Group, En-Birth Study. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):230. doi: 10.1186/s12884-020-03420-x*.
  https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03420-x.pdf (accessed 31 December 2022).

- **Immediate newborn care and breastfeeding: EN-BIRTH multi-country validation study**
  Tahsina, T., Hossain, A. T., Ruysen, H., Rahman, A. E., **Day, L. T**., Peven, K., Rahman, Q. S., Khan, J., Shabani, J., Kc, A., Mazumder, T., Zaman, S. B., Ameen, S., Kong, S., Amouzou, A., Lincetto, O., El Arifeen, S., Lawn, J. E. and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):237. doi: 10.1186/s12884-020-03421-w*.
  https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03421-w.pdf (accessed 31 December 2022).

- **Counting on birth registration: mixed-methods research in two EN-BIRTH study hospitals in Tanzania**
  Reed, S., Shabani, J., Boggs, D., Salim, N., Ng'unga, S., **Day, L. T.,** Peven, K., Kong, S., Ruysen, H., Jackson, D., Shamba, D., Lawn, J. E. and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):236. doi: 10.1186/s12884-020-03357-1.*

https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03357-1.pdf
(accessed 31 December 2022).

- **Birthweight measurement processes and perceived value: qualitative research in one EN-BIRTH study hospital in Tanzania**
  Gladstone, M. E., Salim, N., Ogillo, K., Shamba, D., Gore-Langton, G. R., **Day, L. T**., Blencowe, H., Lawn, J. E and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):232. doi: 10.1186/s12884-020-03356-2.*https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03356-2.pdf
  (accessed 31 December 2022).

- **Antibiotic use for inpatient newborn care with suspected infection: EN-BIRTH multi-country validation study**
  Rahman, A. E., Hossain, A. T., Zaman, S. B., Salim, N., Kc. A., **Day, L. T**., Ameen, S., Ruysen, H., Kija, E., Peven, K., Tahsina, T., Ahmed, A., Rahman, Q. S., Khan, J., Kong, S., Campbell, H., Hailegebriel, T. D., Ram, P. K., Qazi, S. A., El Arifeen, S., Lawn, J. E and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):229. doi: 10.1186/s12884-020-03424-7.* https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03424-7.pdf (accessed 31 December 2022).

- **Survey of women's report for 33 maternal and newborn indicators: EN-BIRTH multi-country validation study**
  Ameen, S., Siddique, A. B., Peven, K., Rahman, Q. S., **Day, L. T**., Shabani, J., Kc, A., Boggs, D., Shamba, D., Tahsina, T., Rahman, A. E., Zaman, S. B., Hossain, A. T., Ahmed, A., Basnet, O., Malla, H., Ruysen, H., Blencowe, H., Arnold, F., Requejo, J., Arifeen, S. E., Lawn, J. E. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):238. doi: 10.1186/s12884-020-03425-6.* https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03425-6.pdf (accessed 31 December 2022).

- **Respectful maternal and newborn care: measurement in one EN-BIRTH study hospital in Nepal**
  Gurung, R., Ruysen, H., Sunny, A. K., **Day, L. T**., Penn-Kekana, L., Malqvist, M., Ghimire, B., Singh, D., Basnet, O., Sharma, S., Shaver, T., Moran, A. C., Lawn, J. E., Kc, A. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):228. doi: 10.1186/s12884-020-03516-4.* https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03516-4.pdf (accessed 31 December 2022).

- **Chlorhexidine for facility-based umbilical cord care: EN-BIRTH multi-country validation study**
  Zaman, S. B., Siddique, A. B., Ruysen, H., Kc, A., Peven, K., Ameen, S., Thakur, N., Rahman, Q. S., Salim, N., Gurung, R., Tahsina, T., Rahman, A. E., Coffey, P. S., Rawlins, B., **Day, L. T**., Lawn, J. E., Arifeen, S. E and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):239. doi: 10.1186/s12884-020-03338-4.*

https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03338-4.pdf (accessed 31 December 2022).

- **Electronic data collection for multi-country, hospital-based, clinical observation of maternal and newborn care: EN-BIRTH study experiences**
  Ruysen, H., Rahman, A. E., Gordeev, V. S., Hossain, T., Basnet, O., Shirima, K., Rahman, Q. S., Zaman, S. B., Rana, N., Salim, N., Tahsina, T., Gore-Langton, G. R., Ameen, S., Boggs, D., Kong, S., **Day, L. T**., El Arifeen, S., Lawn, J. E.and EN-BIRTH Study Group. *BMC Pregnancy Childbirth. 2021 Mar 26;21(Suppl 1):234. doi: 10.1186/s12884-020-03426-5.*
  https://bmcpregnancychildbirth.biomedcentral.com/track/pdf/10.1186/s12884-020-03426-5.pdf (accessed 31 December 2022).

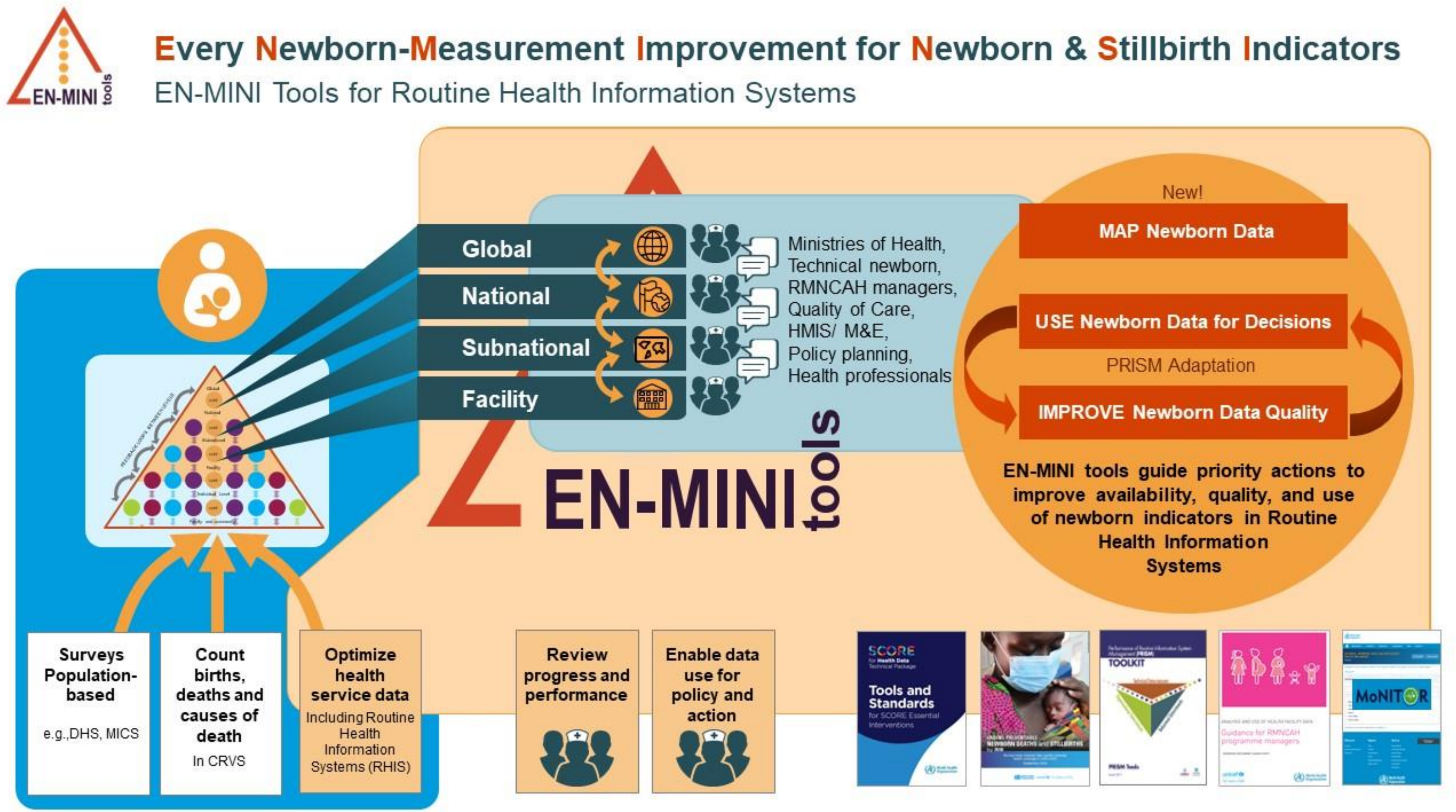
Appendix 4 – Every Newborn-Measurement Improvement for Newborn & Stillbirth Indicators (EN-MINI) Tools for Routine Health Information Systems

The EN-MINI Tools have been the major output follow-on collaborative research with our EN-BIRTH partners in Bangladesh and Tanzania, EN-BIRTH 2, for which I am the LSHTM Principal investigator (2020-2022). We designed the open-access EN-MINI Tools in response to EN-BIRTH study findings of mixed routine register data quality to assess newborn and stillbirth data quality and data use.

The EN-MINI tools guide priority actions to improve availability, quality, and use of newborn indicators in RHIS (Figure 26). The an animated version of the EN-MINI Tools infographic I designed can be found on the EN-MINI Tools website.[116]

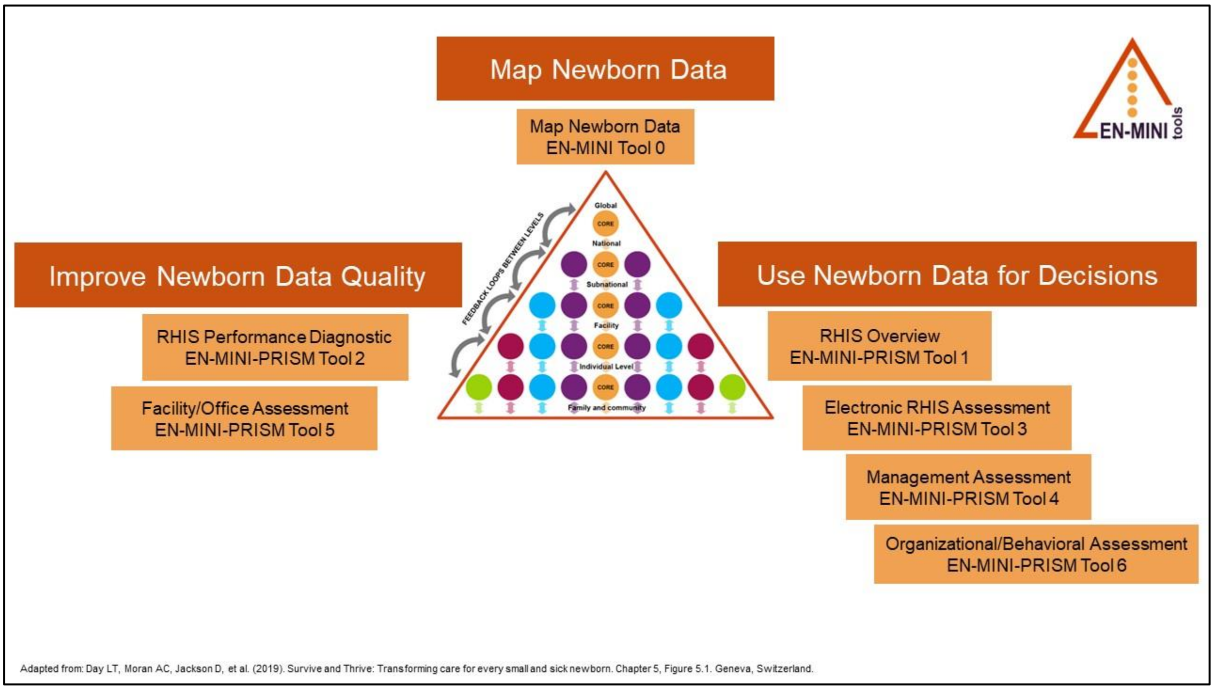
The original PRISM Tools did not include frontline health professionals but only health facility managers. In the EN-MINI-PRISM Tools adaptation we added health professionals as respondents.

Figure 26: Every Newborn-Measurement Improvement for Newborn & Stillbirth Indicators (EN-MINI) Tools infographic - for animated version of see EN-MINI Tools website[3]



The EN-MINI tools are organized in three categories: (1) MAP newborn data availability, (2) assess USE of newborn data for decisions, and (3) identify how to IMPROVE newborn data quality (Figure 27). The USE and IMPROVE tools are adapted from the Performance of Routine Information System Management (PRISM) series.[113]

Figure 27: EN-MINI Tools three categories: Map, Improve, Use



The EN-MINI-PRISM assessment in Tanzania was conducted in 2021 in 16 health facilities and 2 district data offices in the Tanga/ Pangani region in Tanzania. Full details are in the online published online.[3]

Selected Findings from EN-MINI Tools Pilot Assessment in Tanzania 2021A team of 6 data collectors conducted the EN-MINI-PRISM Tools assessment in 2 district offices and 16 health facilities during one week of September 2021. Version 1 of the EN-MINI Tools were used. Data quality was assessed using source and summary report data for April, May, and June 2021. All data were collected digitally using offline password protected tablets and uploaded to the General Data Protection Regulation (GDPR) compliant, secure Open Data Kit (ODK) server (SurveyCTO), using the customized EN-MINI-PRISM Tool forms available on the EN-MINI Tools website.

Selected findings are presented here and more details on the full online published report. [3]

4.1 Increase health facility data use – EN-MINI-PRISM Tools pilot findings

Organizational factors, RHIS processes and use of newborn data for decision at district and health facility level is shown in Figure 28. District level outperformed health facility level in across almost all dimensions assessed.

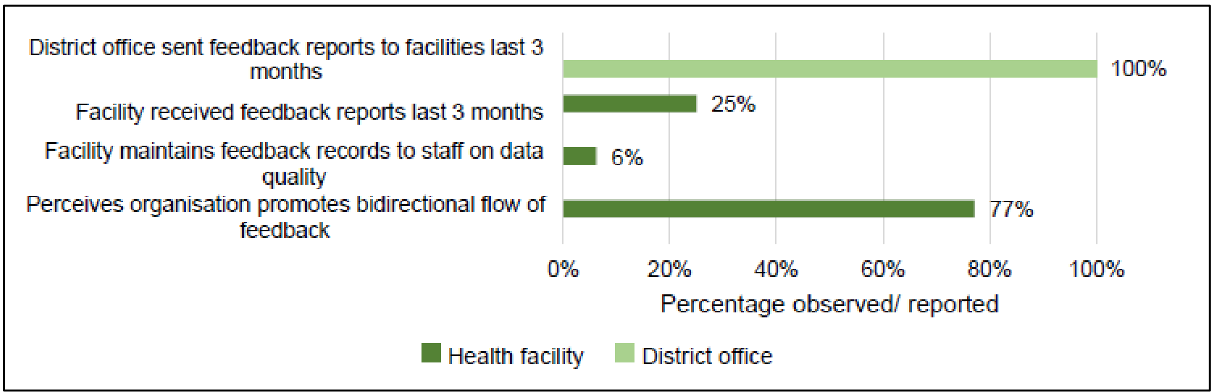
Figure 28: Evidence of existing data use from Tanzania EN-MINI-PRISM tools pilot (n=16 facilities, 2 facility offices)

		District	Facility
Organizational factors	Evidence data analysis taking place	38%	21%
RHIS processes	Data Visualization	100%	25%
	Use of data to produce narrative analytical reports	50%	19%
Use Newborn data for decisions	Use information for discussion on key performance targets	100%	75%
	Use information for coverage of services	0%	13%
	Use sex-disaggregated data	0%	0%
	Use information for human resources decisions	100%	25%
	Use information for quality improvement	100%	0%

4.2 Strengthen RHIS feedback – EN-MINI-PRISM Tools pilot findings

This need to strengthen feedback was a major finding in the EN-MINI-PRISM assessment in Tanzania (Appendix 4, in 16 health facilities at all levels of the health system and 2 district data offices, with 47 health professional/ data professional respondents. Among facility respondents, 77% reported bidirectional feedback is promoted but only 25% of facilities had received a feedback report from the district office in the preceding 3 months and only 6% of facilities maintain feedback records to staff on data quality (Figure 29).

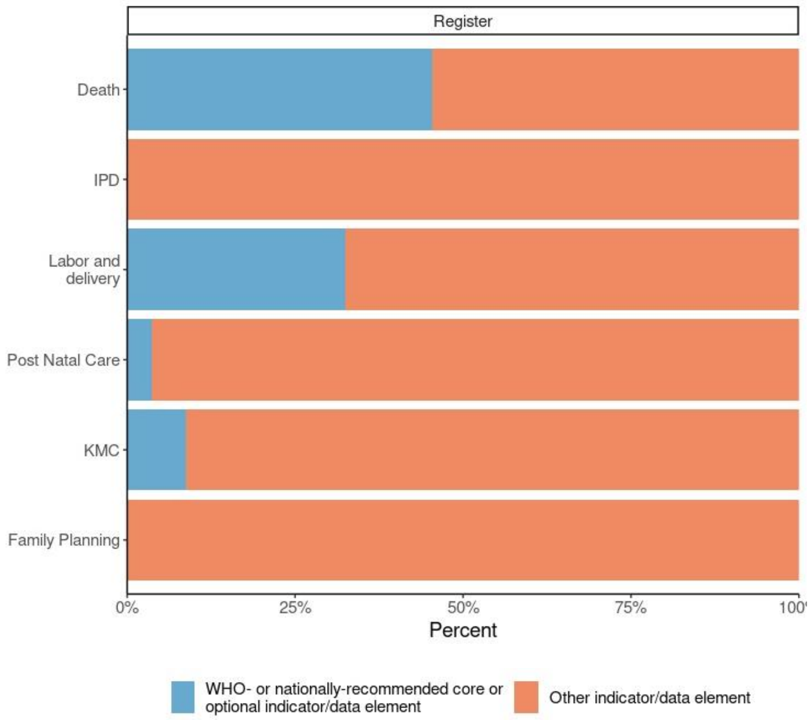
Figure 29: Feedback loops between levels, Tanzania EN-MINI-PRISM pilot (n=16 facilities, 2 facility offices)



4.3 Standardise register design – EN-MINI-PRISM Tools pilot findings

The register component of the figure is shown in Figure 30 for the six Tanzania standardised routine registers.[131] Only 35% of the data elements in the routine labour and delivery register are related to core indicator measurement, representing an opportunity to reduce the burden of data collection in Tanzania.

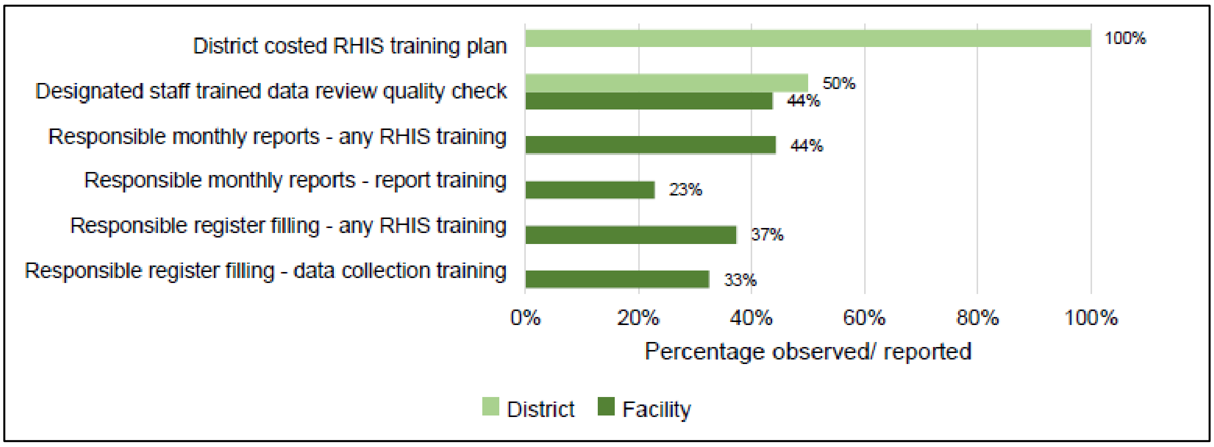
Figure 30: Proportion of newborn data WHO- or nationally recommended as core/optional - Register level Tanzania



4.4 Training for routine register use – EN-MINI-PRISM Tools pilot findings

In the EN-MINI-PRISM assessment in Tanzania, between 23-44% of health professionals reported RHIS related training (Figure 31).

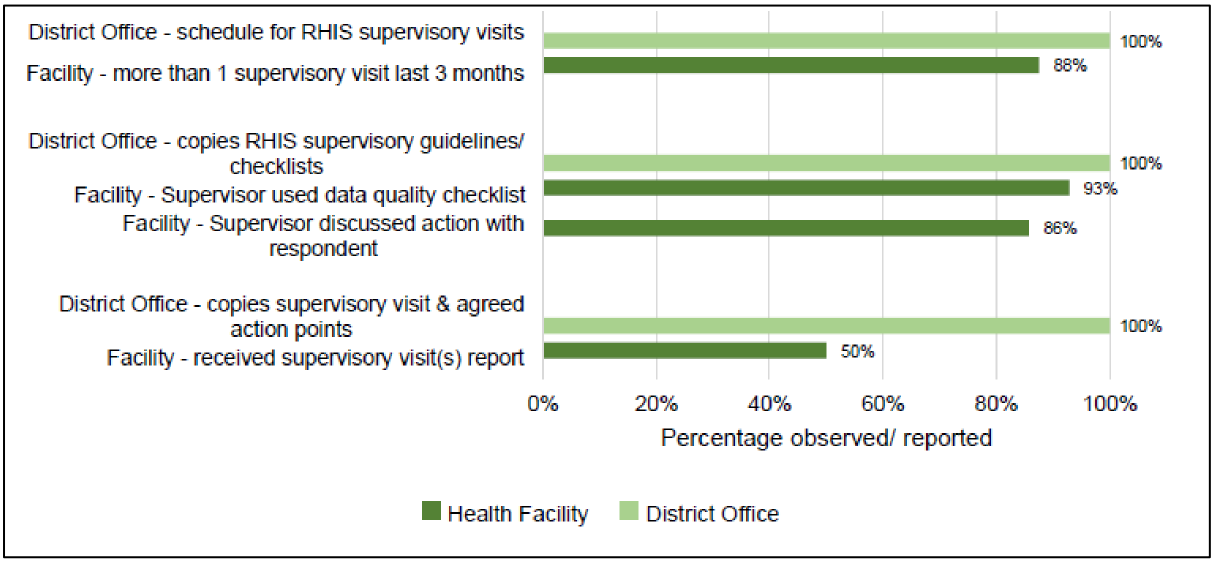
Figure 31: RHIS training at district office and health facilities – Tanzania EN-MINI-PRISM Tools pilot, (n=16 facilities, 2 facility offices



4.5 Supportive Supervision for routine data – EN-MINI-PRISM Tools pilot findings

In the Tanzania EN-MINI-PRISM Tools assessment for 16 facilities, RHIS supervisory processes were established and 88% of facilities had received a supervisory visit in the 3 months prior to the assessment, with 93% of visits used a data quality checklist (Figure 32). Among supervisory visits, 86% included a discussion regarding action points and 50% of facilities had received a report (Figure 32).

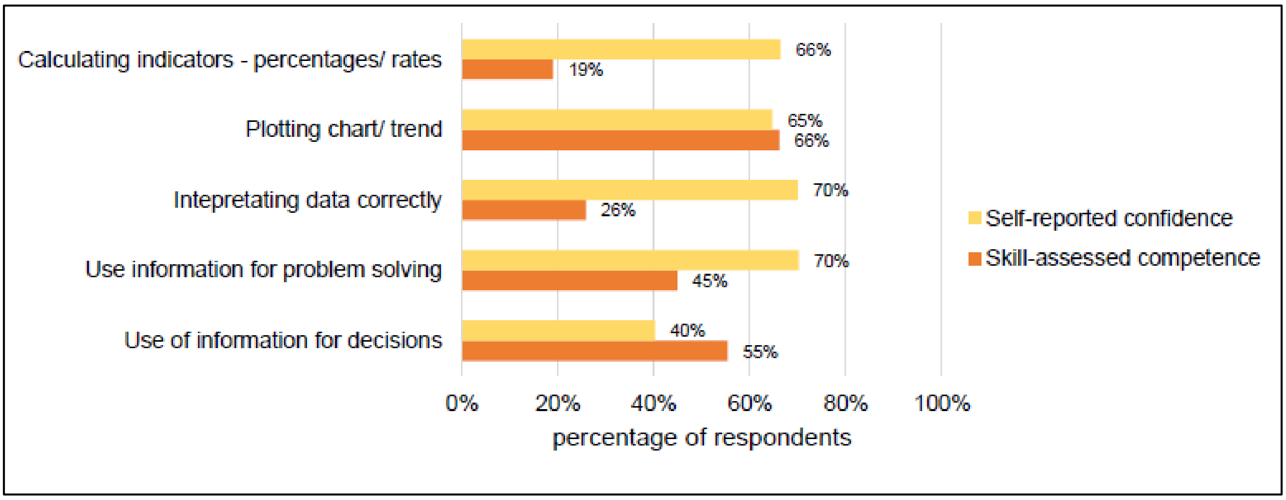
Figure 32: RHIS Supervision health facility and district office – EN-MINI-PRISM pilot, Tanzania (n=16 facilities, 2 facility offices)



4.6 Enable RHIS skills of frontline health professionals – EN-MINI-PRISM Tools pilot findings

The EN-MINI-PRISM Tools pilot captured 47 individual respondents' perceived confidence and measured competence on RHIS tasks through assessment with examples using newborn and stillbirth data (Figure 33). Confidence and competence matched for the task plotting chart/chart trend (65% to 66%). There was a confidence-competence gap for other RHIS skills. Respondents were over-confident in calculating indicators (47% gap), interpreting data (44% gap) and problem-solving (25% gap). Competence was 15% higher than confidence with use of information for decisions.

Figure 33: RHIS task self-reported confidence and skill-assessed competence, Tanzania EN-MINI-PRISM Tools pilot (n=47 respondents, 16 facilities)



Appendix 5 – Dissemination of PhD findings

- Publications – all papers in this PhD are published in open-access peer-reviewed journals.
- Conferences:
  - PhD objective 1 - EN-BIRTH protocol - panel presentation at Health Systems Global Conference in November 2018.[159]
  - PhD objectives 3, 4 and 5 were presented at the DHIS2 annual conference 2021.[160]
  - PhD objectives 2, 3, 4 included in EN-BIRTH presentation with focus on role of midwives in routine register recording ICM 2021.[161]
- Link to policy, norms and standards
  - PhD objective 1 cited in Survive and Thrive Transforming care for every small and sick newborn, Data for Action chapter 5 reference 15.[62]
  - PhD objectives 2,4,5 and 6 in GFF Discussion paper Improving Monitoring Data Systems to Count and Account for Stillbirths.[162]
  - PhD objective 6 evidence considered in revised neonatal resuscitation service coverage indicator WHO-led 2022.
  - EN-MINI tools cited in 'Born to Soon' report, chapter 2, page 22, reference 16.[163]
  - EN-MINI tools cited in 'Stillbirth Definition and Data Quality Assessment for Health Management Information Systems (HMIS), a guideline'.[99]