COUNTRY CONSULTATION ON LOW BIRTHWEIGHT AND PRETERM BIRTH ESTIMATES

Technical Notes

unicef
for every child

World Health Organization

Contents

Annexes 20

Acronyms and abbreviations

DHS	Demographic and Health Survey
DQC	Data quality category
ECG	Estimates Consultative Group
GATHER	Guidelines for Accurate and Transparent Health Estimates Reporting
GA	Gestational age
LBW	Low birthweight
LMICs	Low- and middle-income countries
LSHTM	London School of Hygiene & Tropical Medicine
MICS	Multiple Indicator Survey
MIYCN	Maternal, infant and young child nutrition
SDG	Sustainable Development Goals
TEAM	Technical Expert Advisory Group on Nutrition Monitoring
UNICEF	United Nations Children's Fund
WHO	World Health Organization

1.
Introduction

Low birthweight (LBW), defined as birthweight under 2500 g, presents a significant global health burden. Approximately 20.5 million live births (14.6% of all live births) globally were estimated to be LBW in 2015, with 91% of these occurring in low- and middle-income countries (LMICs) (1, 2). It is estimated that 14.8 million live births (10.6% of all live births) in 2014 were preterm (born before 37 completed weeks of gestation) (3), while approximately 23.3 million neonates were born small for gestational age in 2012 (just under half of these babies were also LBW) (4). LBW and preterm birth are associated with increased risk of mortality in infancy, especially in the neonatal period, and increased morbidity across the lifespan, including developmental and behavioural problems, undernutrition in childhood, and cardiometabolic disease in adulthood (5, 6).

There are overlaps between LBW and preterm birth. LBW births include term and preterm growth-restricted babies, but also preterm babies with normal growth that weigh less than <2500 g because they were born early. The causes of these conditions vary, and some are more amenable to interventions than others.

To improve the survival and health of small and vulnerable newborns, better-quality data are needed – particularly from low-income countries, which bear the greatest burden of LBW and preterm birth. In an effort to address this gap and improve data for small and vulnerable newborns, WHO and UNICEF, supported by London School of Hygiene & Tropical Medicine (LSHTM), are developing joint LBW and preterm birth estimates for the first time. A Steering Group, comprised of experts from WHO, UNICEF and the LSHTM has worked on this estimation exercise. The work has been supported by an Estimates Consultative Group, which is comprised of global experts in preterm birth and LBW measurement, including obstetricians, neonatologists, statisticians, preterm birth researchers and programme staff working in the measurement field.

The purpose of this technical document is to support the country consultation process with Member States by providing a detailed description of the processes and methods for estimating levels and trends of LBW and preterm birth. During the consultation, Member States have an opportunity to: 1) review the draft estimates and their methods; 2) provide advice on primary data sources for their respective countries that may not have been previously reported or used; and 3) build mutual understanding of the strengths and weaknesses of available data and the estimation process.

The **technical notes** aim to provide national focal points with details on the efforts to gather and validate data for LBW and preterm birth estimates and guide them through the steps to complete the country consultation process. This document has five sections:

Section 1. Introduction

Provides a general background for the estimates and overview of the contents of this document.

Section 2. Low birthweight estimates, general background and primary data sources

Technical notes on the history of the LBW Estimates and primary data sources, including overviews of the methods for compiling input data and adjusting survey birthweight data.

Section 3. Preterm birth estimates technical notes

Technical notes on the preterm birth estimates development, including methods used to compile information on preterm birth from primary data sources and generate preterm birth estimates.

Section 4. Overview of statistical modelling approaches for estimating low birthweight and preterm births

This section provides a brief description of the modelling approach used to develop the estimates.

Section 5. Country consultation and practical instructions

Instructions for national focal points to complete the country consultations, including an overview of the country estimates file and description of next steps following eventual feedback from the country consultations.

2.
Low birthweight estimates, general background and primary data sources

A detailed write up of the methods presented here is available in a published, open-access study protocol (7), which is part of the document package and can be found at: https://gatesopenresearch.org/articles/6-80.

This section provides national focal points with an overview of the inter-agency effort to gather and refine country-level LBW data to produce annual LBW prevalence estimates at national, regional and global levels. The section includes:

- 2.1 – an introduction, including how LBW fits into global plans and development goals, and the history of the Joint LBW Estimates
- 2.2 – an overview of the collection and screening of input data from administrative and survey sources
- 2.3 – an overview of adjusting survey birthweight data

After collecting and screening all input data, and adjusting survey data, annual Joint LBW Estimates from 2000–2020 were produced using statistical modelling. The methods for statistical modelling are described in **Section 4** of this document.

2.1 Background of the Joint Low Birthweight Estimates

2.1.1 Low birthweight in global plans and development goals

In 2012, World Health Assembly Resolution 65.6 endorsed the Comprehensive Implementation Plan on Maternal, Infant and Young Child Nutrition (MIYCN), which included a target to reduce LBW by 30% between 2012 and 2025 (8, 9); the end date subsequently extended to 2030 (10). LBW reduction was one of six targets included in the 2012 MIYCN plan, and while LBW is not specified as an indicator for the Sustainable Development Goals (SDG), improving birthweight can contribute to the achievement of SDG Target 2.2 to end all forms of malnutrition by 2030.

In 2015, WHO and UNICEF convened the Technical Expert Advisory Group on Nutrition Monitoring (TEAM) to improve monitoring methodologies for the MIYCN plan (11), and the group helped to develop the 2017 Global Nutrition Monitoring Framework. The monitoring framework guides progress tracking for the six World Health Assembly-endorsed MIYCN plan targets, including using UNICEF-WHO LBW Estimates Working Group methods to monitor LBW reduction (12).

2.1.2 History and use of the Joint Low Birthweight Estimates

The 2022 Joint LBW Estimates, covering the 2000–2020 period, are the third round of global, inter-agency estimates. They are an update of the estimates published in 2019 for the years 2000–2015 (1) and in 2004 for the year 2000 (13).

The 2004 Joint LBW Estimates introduced basic methods for improving LBW estimation (13, 14), and the 2019 edition improved upon those methods, establishing a baseline for the MICYN plan target on reduction of LBW (1, 2), as called for by TEAM (11) and the United Nations Standing Committee on Nutrition (15).

Following the country consultation, the annual LBW Estimates for the 2000–2020 period will be used to monitor country progress on the World Health Assembly Comprehensive Implementation Plan on MIYCN and published in the UNICEF LBW database and the WHO Global Health Observatory. The Joint LBW Estimates are also used by partners in various global reports, including the Global Nutrition Report and The State of Food Security and Nutrition Report. In addition, all input data and analysis code will be made publicly available as per the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) (16).

2.2 Overview of search and screening for input data

Like previous rounds, the 2022 Joint LBW Estimates relied on two broad types of input data for birthweight:

- National administrative data sources
- Nationally representative household surveys

In addition, data to predict birthweight were collected from the United Nations and academic groups, such as Institute for Health Metrics and Evaluation.

The Joint LBW Estimates published methods for data collection and screening in previous rounds (1, 2, 13), and methods for the 2022 edition are available in the protocol document (7). This section provides a summary and highlights changes introduced in the current round of the Joint LBW Estimates. Figure 1 presents the process for collecting and screening birthweight data.

2.2.1 Birthweight data from administrative sources

National administrative data are defined as data from national systems, including Civil Registration and Vital Statistics Systems, National Health Management Information Systems and birth registries. Figure 1 includes details on the identification of administrative data sources for the 2022 LBW Estimates. In summary, a search for administrative data was conducted for countries with a facility birth rate ≥80%, and data were included in LBW estimation for years when administrative birthweight data accounted for ≥80% of United Nations estimated number of live births (17).

2.2.2 Birthweight data from survey sources

Details on the identification and screening of survey data sources for the 2022 LBW Estimates are presented in Figure 1. In summary, a search was conducted for nationally representative surveys from LMICs that were carried out after 1996 and that had publicly available data on birthweight and size at birth. Surveys were excluded if they had excessive missing values, small sample sizes, severe heaping or an implausible adjusted LBW prevalence.

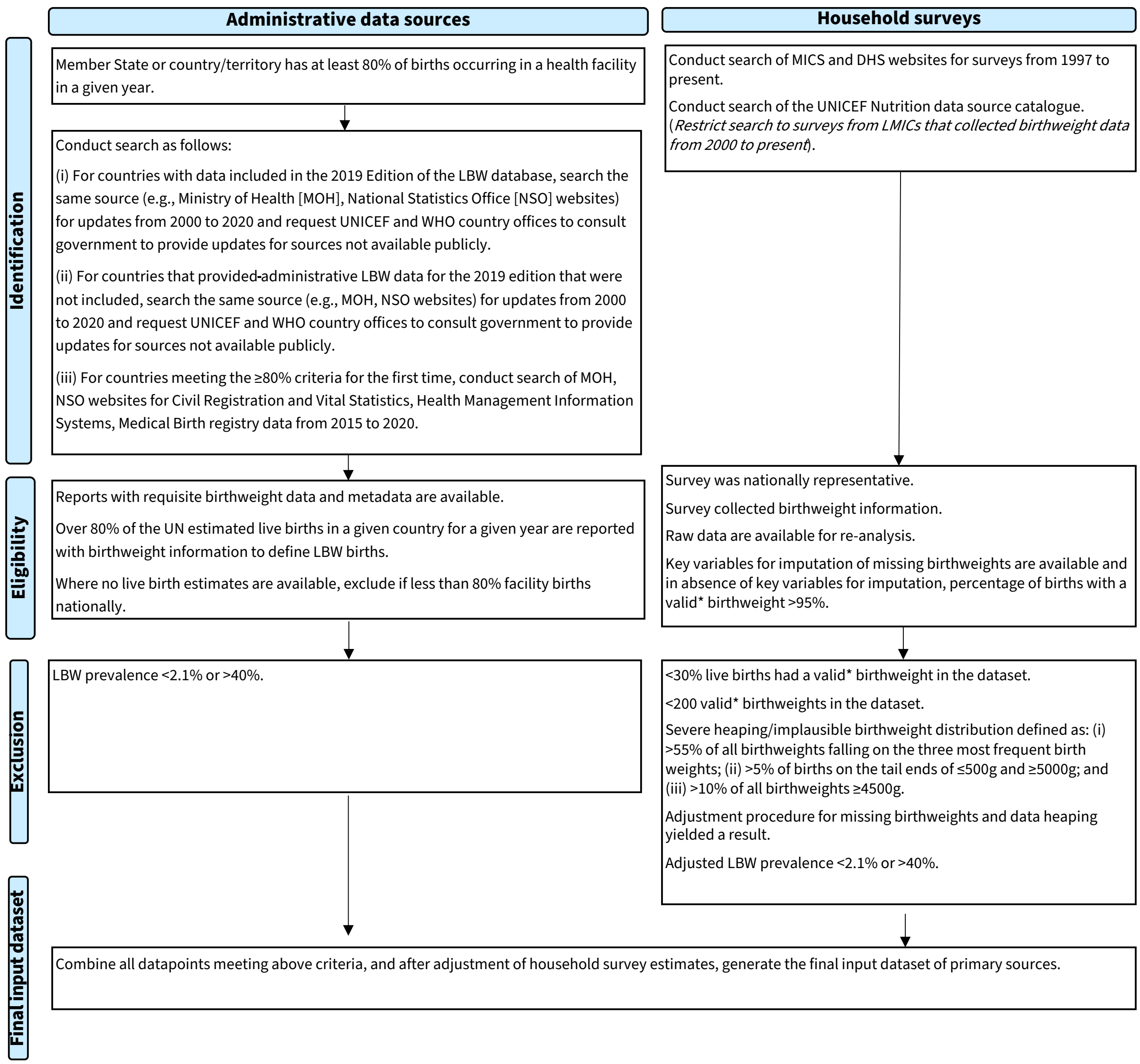
There were slight changes to the identification of survey data sources in 2022. The 2019 LBW Estimates only searched among the Demographic Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS),[1] while the 2022 edition did not restrict to specific survey types.

The 2022 LBW Estimates used the same exclusion criteria as 2019 (2) with one important caveat — there was a switch from using weighted data in 2019 to using unweighted data in 2022 for all

1 There was one exception to this criterion in 2019; the 2019 LBW Estimates included non-DHS/MICS surveys from China.

exclusion criteria except implausible prevalence. The switch to unweighted data aligned LBW Estimates methodology to the UNICEF-WHO-World Bank Joint Child Malnutrition Estimates methods (18).

Figure 1. Process for collection and screening of birthweight data from administrative and survey sources



* Valid birthweights are those between 250g and 5500 g; birthweights outside of this range will be set to missing.

2.3 Overview of adjusting survey birthweight data

Children with a missing birthweight in surveys are more likely to have been born LBW, which causes an underestimation of LBW prevalence if not addressed (14). Birthweight heaping, a preference for recording or reporting birthweights rounded to the nearest 100 g or 500 g, can also cause underestimation. When rounding occurs, some children whose true weight is below 2500 g are reported as weighing exactly 2500 g, which moves those children out of the LBW classification. Since a high percentage of missing birthweights and extensive rounding are common occurrences in surveys from LMICs, adjustment is needed to avoid underestimation of LBW prevalence.

The 2022 LBW Estimates followed the same adjustment methods to account for missing values and heaping that were used in the 2019 edition. The adjustment methods were previously validated (19) and the effects of adjustment are shown in Annex 1 with 2019 data. In 2019, adjustment increased LBW prevalence by 3.4 percentage points on average (95% confidence interval (95% CI): 3.0-3.7). The increase from adjustment was related to the degree of missing values and heaping; in countries with the highest levels of missing values and heaping, adjustment resulted in an increase of nearly 12 percentage points. The adjustment methods are described elsewhere (1, 2, 7), and Annex 2 contains detailed descriptions.

There were three steps for adjustment:

1. Filling in missing values by predicting birthweight based on other variables (i.e., multiple imputation), including: mother's perception of child's size at birth, child's sex, parity, single or multiple birth, mother's height and mother's body mass index
2. Removing heaping by creating smoothed, normal distributions using a finite mixture model
3. Converting smoothed distributions into LBW prevalence using the z-score table for a normal distribution

In the first step to fill in missing values of birthweight, multiple imputation (20) created five new datasets with no missing values. In the second step to remove heaping, the finite mixture model created two normal distributions for each of the five datasets, resulting in a total of 10 distributions. The final step merged all 10 distributions into a single, survey point estimate with confidence intervals based on both within- and between-dataset variance across all five datasets. Creating five datasets accounted for increased uncertainty from multiple imputation, and using two distributions accounted for the population distribution of birthweight, which is not normally distributed (21).

3.
Preterm birth estimates

Details about the study are available in a published, open-access study protocol (22), which is part of the document package and can be found at: https://pubmed.ncbi.nlm.nih.gov/34669749/.

3.1 Definition of preterm birth and estimation challenges

3.1.1 Definition of preterm birth

The International Statistical Classification of Diseases and Health Problems, eleventh revision (23), uses the WHO definition of preterm birth, which includes: "All births before 37 completed weeks of gestation or fewer than 259 days since the first day of a woman's last menstrual period". WHO recommends reporting the preterm birth rate as a percentage using the following indicator, including all livebirths with no lower gestational age boundary:

$$\frac{\text{Number of live born preterm births singleton or multiple}}{\text{Number of live births single or multiple}}$$

3.1.2 Challenges of developing preterm birth estimates

Accurate measures of preterm birth are challenging in many countries given incomplete or unavailable data from national administrative sources. Furthermore, developing estimates of preterm birth can be complicated by additional factors that can impede accurate measurement, estimation and comparison:

- The risk of preterm birth can be higher in some vulnerable sub-populations (including among poor, uneducated, rural-dwelling women or other minorities) where data collection may be more limited and/or facility-based births are less common.
- Misclassification between babies who die shortly after birth and stillbirths, and omission of babies dying in the early neonatal period, which is especially common around the thresholds of viability, can impact on the accuracy of the recorded preterm birth rate among livebirths (24).
- National differences in the definition of preterm birth (for example, using live births or total births as the denominator, or different gestational age thresholds for defining preterm birth cases), and inclusion of a lower gestational age boundary for inclusion of livebirths in the data system, can complicate international comparisons.
- In many countries, the definition of preterm birth has changed over time, especially given increasing viability of extremely preterm babies, further limiting comparability.
- Gestational age estimation error is also an important factor. Generally, the later in pregnancy a gestational age estimate is made, the wider the uncertainty of that estimate. Routine early pregnancy ultrasound with fetal biometric measurements is considered the "gold standard" for gestational age assessment (25). However, other methods are also used, such as calculation from date of last menstrual period, symphysis-fundal height measurement or postnatal examination of the newborn. Many countries report the use of "best obstetric estimate" of gestational age, using a combination algorithm of ultrasound, last menstrual period and clinical assessment (26).

While birthweight is closely linked with gestational age, it cannot be used interchangeably since there is a range of "normal" birthweight for a given gestational age and gender. LBW is defined as less than 2500 grams. In some settings, the majority of LBW babies are born preterm, but in others (especially in South Asia) there is a high rate of term babies who are small for gestational age. A baby born preterm has greater chance of dying than a baby of the same birthweight born small for gestational age at term (27).

3.2 Data sources

There are two broad categories of data sources that were used for the preterm birth estimation: national administrative data and research studies. National administrative data included data from Civil Registration and Vital Statistics Systems, Health Management Information Systems and Medical Birth Registries, which are the preferred data sources for estimating preterm birth rates. However, for many countries, data from administrative sources were incomplete or not available. For these countries, a systematic review of research studies was undertaken to identify additional data points that could be used in the estimation process.

3.2.1 Search strategy

For administrative data sources, a systematic search was conducted of Ministry of Health and National Statistical Office publications and datasets for WHO Member States that had a population-based facility birth rate of at least 80%[2] in the latest year for which data were available between 2010 and 2020. For countries that met this threshold, administrative data sources used in the previous round were initially searched to identify more recent data points (22).

A systematic review of studies was conducted for countries that did not meet the threshold for administrative data sources and that lacked eligible administrative data for the estimation work.

3.2.2 Eligibility criteria

For administrative data sources, all data from Civil Registration and Vital Statistics Systems, Health Management Information Systems and Medical Birth Registries identified through the search were eligible for inclusion if at least 80% of the United Nations estimated live births (17) in a given country for a given year were reported with gestational age information to define preterm birth. Where information on the number of livebirths was not available, datapoints from administrative sources were considered for inclusion if >80% of births in the country-year took place in a health facility. For studies, all data sources identified through the search were eligible for inclusion if the outcome was derived from an observational or intervention study design conducted at national or sub-national level in either population- or facility-based settings. For further details on eligibility requirements, please refer to Study Protocol (22).

2 The facility birth rate data were obtained from https://data.unicef.org/topic/maternal-health/delivery-care/.

4.

Overview of statistical modelling approaches for estimating annual low birthweight prevalence and preterm birth estimates

4.1 Background and rationale for use of a statistical model and updating methods

The process of searching for, collating, adjusting and selecting primary data sources for LBW and preterm birth were described in Sections 2 and 3 and resulted in a master dataset that could be used for statistical modelling. Despite an extensive search, many countries lacked data meeting inclusion criteria; were missing data for many years; or had estimates that varied significantly by source type. A suitable statistical model was thus needed to combine the often patchy and disparate data into a trendline representing the 'best set of annual estimates,' for each country.

This section describes the steps used to build and test the Bayesian multilevel-mixed regression model. Modelling was carried out separately for the two outcomes (LBW and preterm birth), but the general processes and methods are similar for both. The rest of the section covers both LBW and preterm birth and is divided into three sub-sections: inputs; data quality categorization; and fitting/testing the model and presenting the results.

4.2 Modelling inputs

To model LBW prevalence or preterm birth rate using a Bayesian framework; the model incorporated input data (the preterm birth and LBW country data from primary sources covered in Sections 2 and 3), covariates and priors. This sub-section describes the process of identifying and selecting covariates and priors.

4.2.1 Identifying and selecting covariates

Covariates are variables that are associated with and can help to predict preterm birth or LBW, especially for country-years without input data. Time-series data on covariates can improve model prediction, especially for countries with few input data points. For example, it would not be possible to predict the annual LBW estimates with any certainty for a country that has only one survey input data point between 2000–2020 without the help of the annual covariate timeseries data, which help to inform what was happening in the 19 years without input data. The final set of covariates was selected using the following four steps, which are detailed in **Annex 3**: (i) identification of plausible covariates using a conceptual framework; (ii) a search for covariate timeseries data from United Nations and other databases; (iii) assessment of the quality of the timeseries data for all potential covariates, and filling in of any missing years using linear interpolation and constant extrapolation; and (iv) statistical analysis to identify the smallest set of covariates with the best quality timeseries that were most highly correlated with the outcome of interest.

The final covariates for the models are included in **Annex 3**.

4.2.2 Identifying and selecting priors

Priors are key components of Bayesian modelling and usually constitute background information that is known about certain parameters of interest expressed as a distribution (prior distribution, such as normal, beta-distribution). There are several ways to define priors; for LBW and preterm birth modelling, non-informative, data-driven priors were mainly used. For example, the standard definition requires LBW to be calculated using live births with a birthweight, but some countries use alternate, biased denominators that would lower the LBW prevalence, such as livebirths or total births. Similarly, for preterm birth, the standard definition requires that both singleton and multiple births be included, but some countries use a biased definition and only include singleton births. The indicator definition has a known impact on the estimate and is an example of a prior incorporated in the LBW and preterm models. Priors were selected based on known biases that the input data may have and their potential influence on LBW prevalence and preterm birth rate. Key priors considered in the modelling of LBW and preterm births were:

- Data source type (i.e., administrative data, survey data or published studies)
- Non-standard definitions:
  - ➜ *For preterm birth, for example, the use of a different gestational age cut-off than the standard <37 weeks, or not including both singleton and multiple births*
  - ➜ *For LBW, for example, the use of a different denominator than the standard livebirths with a birthweight (e.g., livebirths only, total births, or reported percentage LBW/no denominator)*
- Data quality (i.e., representativeness in terms of differences in coverage, ratios of sub-envelopes of birthweight reported, such as the ratio of 1000g/2500g etc.,)

See **Annexes 4 and 5** for further details on priors used in the modelling.

4.3 Data quality categorization

After investigating availability of and correlations between data quality indicators, data quality criteria were developed based on an iterative process. The final data quality categories (DQC) and associated criteria for LBW are shown in Table 4.1. Each country with at least one year of administrative data included in the dataset was placed into a DQC based on the listed criteria: A (highest quality), B (moderate quality), or C (lowest quality).

Table 4.1: LBW Data Quality Categories and Criteria for Administrative Data

Data quality categories	**Criteria 1** Representativeness	**Criteria 2** Data source type	**Criteria 3** Denominator	**Criteria 4** Sub-envelope capture
A*	≥ 90% recorded birthweight coverage# and ≥ 90% facility births‡	Must be civil registration and vital statistics or medical birth registry	Must be livebirths with birthweight (i.e., known values in the database for <2500g and ≥2500g) for all country-years	Birthweight <1000g/<2500g ≥4%‡ or if <1000g/<2500g is unavailable, birthweight <1500g/<2500g ≥ 12.5%‡
B	Not meeting representativeness criteria for DQC A	Must be civil registration and vital statistics or medical birth registry	Can be any (e.g., total births) but not reported percentage LBW (i.e., cannot be no denominator)	Not applied as relevant data not available for all years for these countries
C	Not meeting representativeness criteria for DQC A	Can be any, including "health management information system (HMIS) (such as DHIS2)" or "Other, hospital-based systems"	Can be any (e.g., total births) or reported percentage LBW (i.e., no denominator)	Not applied as relevant data not available for all years for these countries

* France included as an exception

\# Recorded birthweight coverage was calculated by dividing the number of live births with a birthweight in the administrative data source by the World Population Prospects 2022 Edition estimated live births.

‡Across 80% of the time series 2000-2019 (i.e., ≥16 country-years).

After classifying countries, the DQC were integrated into the model through weighting, setting variance and adjusting for bias. Weighting applied only to countries with both administrative and survey/study data. For those countries, DQC informed how to weight the data sources when combining into a single estimate. For example, in DQC C (low quality) countries survey/study data were weighted higher than administrative data, and the combined estimate was closer to the survey data. In DQC B (moderate quality) equal weight was given to all sources, and since most countries have more administrative data points than survey data points, the combined estimate was closer to the administrative data.

For countries with or without survey data, DQC influenced variance and how much administrative data were adjusted to account for bias. We set variance higher for countries with lower data quality, resulting in wider confidence intervals. The bias adjustment was determined by averaging the observed differences in LBW prevalence between survey and administrative data among countries that had both types of data within categories B and C (there was no bias adjustment for category A). All administrative data points were then adjusted up in those categories based on the observed difference. Further details on the DQC can be found in **Annex 4** and illustrative examples of LBW and preterm estimates for select countries by DQC can be found in **Annex 7**.

4.4 Fitting / testing the model

This section provides an overview of the processes and methods for statistical modelling with the selected data to estimate annual LBW prevalence and preterm birth rate.

4.4.1 Fitting the Bayesian multilevel-mixed regression model

Estimates of LBW prevalence (2000–2020) and preterm birth rate (2010–2020) at national levels were predicted from Bayesian multilevel-mixed regression models that incorporated the following important characteristics:

- Random country-specific intercepts
- Non-linear time trends modelled using splines
- Time-series covariates data
- Bias shift and standard deviation terms based on the country's input data (e.g., survey versus administrative data)
- DQC as a prior
- Data source-type as a prior

The country-specific intercepts were built within regions, incorporating within-region and between-region variances. The six SDG regions were used in the modelling (see **Annex 6** for details). At the country-level, alongside the random intercepts, penalized splines were used as temporal smoothing across the time-series. The model was fit on a logit scale to ensure that proportions were bounded between zero and one.

The Markov chain Monte Carlo sampling method was used for estimation, combining the country-level regression terms with the DQC bias terms to generate an estimate of the LBW and preterm birth rate time-series. For countries with no input data, the final model was used to predict estimates of the LBW prevalence and preterm rates based on regional intercept and time-trends and country-level covariates.

4.4.2 Testing the model: validation, performance and sensitivity analysis

Leave-one-out cross-validation methods were used for model validation (i.e., leaving out 20 per cent of the country-years data). The estimates after "leave-out" were compared with the estimates using Bland-Altman plots, Bayesian and Akaike Information Criteria (BIC and AIC), and percentage of points that fell within the original 95% credible intervals. Standard diagnostic checks, such as trace plots, were used to check convergence and the sampling efficiency.

Sensitivity analyses were completed to ensure that covariates were useful in the prediction of the estimates, and to test other components of the model. The directions of the covariates were checked to ensure that they were in a biologically plausible direction and the following analyses were conducted:

- Assessing the impact of including additional covariates used in the previous rounds of estimates (i.e., neonatal mortality rate and child underweight for LBW) or no covariates
- Evaluating the impact of changing the number of knots in the temporal smoothing element of the regression model
- Checking the appropriateness of the non-informative priors by replacing the values within the non-informative priors and checking that this does not unduly change the model.

4.4.3 Presentation of results

Annual country-level point estimates with the weighted 2.5th and 97.5th percentiles for uncertainty intervals of the posterior samples were calculated. The predicted model estimates were then applied to the World Population Projection 2022 (WPP) (30) of livebirths to obtain the estimated numbers of LBW and preterm live births for each level: country, regional, and global. For those countries with at least one input data point (e.g., survey) included in the estimation period, national-level prevalence estimates will be published for each year: 2000–2020 for LBW and 2010–2020 for preterm. For countries without input data meeting inclusion criteria, their country-level estimates from the model will not be published, but will contribute to the global and regional estimates.

5.
Country consultation and practical instructions

5.1 List of supporting documents

1. Circular letter (C.L.8.2021)
2. Methodological notes
3. Study protocols
4. Country profiles
5. Country portal – how to login
6. Country portal – how to use
7. Low birthweight and preterm birth estimates country consultation process brief

5.2 Instructions

- Interactions throughout this consultation process will occur through WHO's country platform.
- National focal points will receive an email informing them when the country consultation has started. This email will include a guide on how to login and how to use the country portal.
- National focal points will receive an automatic email from the platform inviting them to establish their password to access the country portal.
- WHO and UNICEF country offices have access to the country portal and can also see and download all files.
- All relevant documents for this consultation, including these technical notes, the study protocols and the country profiles, will be available to download through WHO's country platform. Please review the Technical Notes and the Study Protocols for further information on both indicators: low birthweight and preterm birth.
- Please review the LBW and preterm birth estimates for your country and identified data sources used (country profile provided in excel format).
- Please identify any potentially eligible national-level LBW and preterm birth and other representative studies that are not already included.
- The country consultation will take place between 29 September and 15 November 2022.

Please share any observations, queries or potential additional data via the "Chat tab"of the country consultation platform or by email to: healthstat@who.int and LBW@unicef.org no later than 15 November 2022.

Any input provided by country focal point(s) after the close of the consultation may not be included in the final publication of the estimates but may be considered in the next round of estimates.

References

1. United Nations Children's Fund (UNICEF), World Health Organization (WHO). UNICEF-WHO Low birthweight estimates: Levels and trends 2000–2015. Geneva: World Health Organization; 2019. Report No.: CC BY-NC-SA 3.0 IGO.

2. Blencowe H, Krasevec J, de Onis M, Black RE, An X, Stevens GA, 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.

3. Chawanpaiboon S, Vogel JP, Moller AB, Lumbiganon P, Petzold M, Hogan D, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Glob Health. 2019;7(1):e37-e46.

4. 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.

5. Mu M, Wang SF, Sheng J, Zhao Y, Li HZ, Hu CL, et al. Birth weight and subsequent blood pressure: a meta-analysis. Arch Cardiovasc Dis. 2012;105(2):99-113.

6. Mohseni R, Mohammed SH, Safabakhsh M, Mohseni F, Monfared ZS, Seyyedi J, et al. Birth Weight and Risk of Cardiovascular Disease Incidence in Adulthood: a Dose-Response Meta-analysis. Curr Atheroscler Rep. 2020;22(3):12.

7. Krasevec J, Blencowe H, Coffey C et al. Study protocol for UNICEF and WHO estimates of global, regional, and national low birthweight prevalence for 2000 to 2020 [version 1; peer review: awaiting peer review]. Gates Open Res 2022, 6:80 (https://doi.org/10.12688/gatesopenres.13666.1.

8. WHO. Global Nutrition Targets 2025: Low Birth Weight Policy Brief (WHO/NMH/NHD/14.5). Geneva: WHO; 2014.

9. WHO. Sixty-Fifth World Health Assembly: Resolutions and Decisions Annexes. Geneva: WHO; 2012.

10. WHO, UNICEF. WHO/UNICEF Discussion Paper: The extension of the 2025 Maternal, Infant and Young Child nutrition targets to 2030. Geneva: WHO; 2019.

11. Methodology for monitoring progress towards the global nutrition targets for 2025 – technical report. World Health Organization and the United Nations Children's Fund (UNICEF) Technical Expert Advisory Group on Nutrition Monitoring (TEAM); 2017.

12. Global Nutrition Monitoring Framework: operational guidance for tracking progress in meeting targets for 2025. Geneva: WHO; 2017.

13. United Nations Children's Fund (UNICEF), World Health Organization (WHO). Low birthweight: country, regional, and global estimates. New York: UNICEF; 2004.

14. 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.

15. Nutrition Targets and Indicators for the Post-2015 Sustainable Development Goals: Accountability for the Measurement of Results in Nutrition: A Technical Note. Rome: UNSCN; 2015.

16. Stevens GA, Alkema L, Black RE, Boerma JT, Collins GS, Ezzati M, et al. Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement. The Lancet. 2016;388(10062):e19-e23.

17. United Nations Population Division Department of Economic and Social Affairs. World population prospects: the 2019 revision. New York (NY); 2019.

18. Levels and trends in child malnutrition: key findings of the 2021 edition of the joint child malnutrition estimates. New York: United Nations Children's Fund (UNICEF), World Health Organization, International Bank for Reconstruction and Development/The World Bank.; 2021.

19. Chang KT, Carter ED, Mullany LC, Khatry SK, Cousens S, An X, et al. Validation of MINORMIX Approach for Estimation of Low Birthweight Prevalence Using a Rural Nepal dataset. The Journal of Nutrition. 2021.

20. StataCorp. Stata Multiple Imputation Reference Manual: Release 17. College Station, TX: StataCorp LLC; 2021.

21. Gage TB, Therriault G. Variability of birth-weight distributions by sex and ethnicity: analysis using mixture models. Hum Biol. 1998;70(3):517-34.

22. De Costa A, Moller AB, Blencowe H, Johansson EW, Hussain-Alkhateeb L, Ohuma EO, et al. Study protocol for WHO and UNICEF estimates of global,

regional, and national preterm birth rates for 2010 to 2019. PLoS One. 2021;16(10):e0258751.

23. World Health Organization. International Classification of Diseases 11th Revision. Geneva2022.

24. Lawn JE, Gravett MG, Nunes TM, Rubens CE, Stanton C. 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 Suppl 1(Suppl 1):S1.

25. Committee Opinion No 700: Methods for Estimating the Due Date. Obstet Gynecol. 2017;129(5):e150-e4.

26. Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller AB, et al. Born too soon: the global epidemiology of 15 million preterm births. Reprod Health. 2013;10 Suppl 1(Suppl 1):S2.

27. Katz J, Lee AC, Kozuki N, Lawn JE, Cousens S, Blencowe H, et al. Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. Lancet. 2013;382(9890):417-25.

28. Royston P. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. Stata Journal. 2009;9(3):466-77.

29. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med. 1999;18(6):681-94.

30. UN-Department of Economic and Social Affairs (DESA) Population Division. World Population Prospects. 2022. (https://population.un.org/wpp/)

Annexes

ANNEX 1

Unadjusted and adjusted low birthweight prevalence stratified by time period, region, birthweight missingness and heaping index

	Number of surveys	Unadjusted (Crude) Mean per cent LBW (mean of 95 per cent confidence intervals)	Multiple imputation and a mixture of 2 normal distributions (New adjustment) Mean per cent LBW (mean of 95 per cent confidence intervals)
All surveys (all years)	**216**	**9.5% (7.8%–11.2%)**	**12.9% (11.1%–14.7%)**
2004 or earlier	82	9.9% (8.0%–11.8%)	13.8% (11.7%–15.8%)
2005–2010	80	9.3% (7.8%–10.9%)	12.5% (10.9%–14.2%)
2011–2016	54	9.0% (7.4%–10.6%)	12.0% (10.3%–13.7%)
West and Central Africa (all years)	47	11.3% (9.4%–13.1%)	16.5% (14.4%–18.7%)
2004 or earlier	12	12.3% (10.3%–14.4%)	18.1% (15.6%–20.7%)
2005–2010	23	11.3% (9.5%–13.1%)	16.5% (14.4%–18.6%)
2011–2016	12	10.2% (8.5%–11.8%)	14.9% (13.1%–16.8%)
Eastern and Southern Africa (all years)	49	10.2% (8.6%–11.8%)	14.7% (12.9%–16.5%)
2004 or earlier	25	10.7% (8.7%–12.7%)	16.1% (13.8%–18.3%)
2005–2010	13	9.8% (8.5%–11.1%)	14.0% (12.6%–15.4%)
2011–2016	11	9.3% (8.2%–10.3%)	12.3% (11.1%–13.5%)
Middle East and North Africa (all years)	9	9.8% (8.6%–11.0%)	12.9% (11.7%–14.1%)
2004 or earlier	3	10.8% (9.7%–12.0%)	15.7% (14.5%–17.0%)
2005–2010	4	10.8% (9.4%–12.1%)	13.2% (11.8%–14.5%)
2011–2016	2	6.4% (5.5%–7.3%)	7.9% (7.0%–8.7%)
South Asia (all years)	7	15.7% (14.0%–17.4%)	24.6% (22.6%–26.7%)
2004 or earlier	1	21.6% (20.7%–22.5%)	32.9% (31.0%–34.8%)
2005–2010	2	10.3% (8.4%–12.3%)	15.8% (13.6%–17.9%)
2011–2016	4	16.9% (15.2%–18.6%)	27.0% (24.9%–29.0%)
East Asia and the Pacific (all years)	25	8.1% (6.9%–9.2%)	11.1% (9.8%–12.3%)
2004 or earlier	9	9.4% (8.1%–10.7%)	12.4% (11.1%–13.8%)
2005–2010	10	8.1% (6.9%–9.2%)	11.3% (10.1%–12.6%)
2011–2016	6	6.0% (5.0%–7.1%)	8.5% (7.2%–9.8%)
Latin America and Caribbean (all years)	39	10.1% (8.1%–12.2%)	11.5% (9.5%–13.5%)
2004 or earlier	14	9.7% (7.6%–11.8%)	11.4% (9.4%–13.5%)
2005–2010	15	10.1% (8.3%–11.9%)	11.5% (9.8%–13.2%)
2011–2016	10	10.7% (8.3%–13.2%)	11.7% (9.3%–14.1%)
Eastern Europe and Central Asia (all years)	40	5.6% (3.9%–7.2%)	6.8% (5.1%–8.5%)
2004 or earlier	18	6.7% (4.8%–8.6%)	8.7% (6.7%–10.6%)
2005–2010	13	4.8% (3.3%–6.2%)	5.6% (4.3%–7.0%)
2011–2016	9	4.4% (2.8%–6.0%)	4.8% (3.3%–6.4%)

	Number of surveys	Unadjusted (Crude) Mean per cent LBW (mean of 95 per cent confidence intervals)	Multiple imputation and a mixture of 2 normal distributions (New adjustment) Mean per cent LBW (mean of 95 per cent confidence intervals)
By percent of births without a birthweight in the dataset			
60 to 69%	15	12.5% (10.0%–14.9%)	20.2% (17.3%–23.2%)
50 to 59%	24	10.8% (9.0%–12.7%)	17.0% (14.8%–19.1%)
40 to 49%	22	11.6% (9.9%–13.4%)	17.1% (15.0%–19.2%)
30 to 39%	28	10.5% (9.2%–11.9%)	14.7% (13.2%–16.2%)
20 to 29%	23	10.5% (8.7%–12.3%)	13.6% (11.8%–15.4%)
10 to 19%	25	10.0% (8.6%–11.5%)	12.5% (11.0%–14.0%)
5 to 9%	21	8.2% (6.3%–10.1%)	10.1% (8.3%–12.0%)
0 to 4%	58	6.6% (5.0%–8.1%)	7.7% (6.1%–9.2%)
By heaping index[1]			
≥ 10%	11	15.4% (13.5%–17.2%)	24.0% (21.8%–26.2%)
7.5 to <10%	15	12.9% (10.7%–15.1%)	19.8% (17.1%–22.5%)
5 to <7.5%	62	9.9% (8.4%–11.4%)	14.7% (12.9%–16.5%)
2.5 to <5%	58	8.9% (7.4%–10.5%)	11.9% (10.3%–13.6%)
1% to <2.5%	40	6.9% (5.3%–8.6%)	8.1% (6.5%–9.7%)
≤ 1%	16	5.3% (3.2%–7.5%)	5.9% (3.9%–8.0%)

1. Heaping index was derived using birthweights in the original dataset and was calculated by dividing the number of birthweights recorded as exactly 2500 g by all births with a) birthweight in the dataset. Fourteen surveys where original birthweights were not in g and thus showed heaping patterns on alternate values and which had a heaping index of <1.0 were excluded from this stratification.

ANNEX 2
Detailed methods for adjustment of survey birthweight data

The Joint LBW Estimates had three main steps to adjust survey birthweight data: (i) accounting for missing values with multiple imputation; (ii) accounting for heaping by creating smoothed distributions with a finite mixture model; and (iii) converting smoothed distributions into a final adjusted estimate of LBW prevalence. The Joint LBW Estimates used Stata 16 for all adjustments.

Multiple imputation

Biologically implausible birthweights, <250 g and >5,500 g, were set to missing before imputation. All missing birthweights were replaced with an estimated value (i.e., imputed) using Stata's multiple imputation command, which allowed for imputing multiple variables simultaneously using regression. Mother's perception of size at birth was the key variable for imputing birthweight and imputation was not carried out on surveys missing this key variable. Additional variables included in multiple imputation when available[1] were: maternal parity, sex of child, singleton status, maternal height, maternal body mass index and stratum. Size at birth, maternal height and body mass index and child sex were unchanged and treated as ordinal, continuous and binomial variables respectively. Maternal parity was condensed into three categories (1 birth, 2–3 births and ≥4 births) and treated as a categorical variable. Singleton status was coded as a binary variable (multiple births or single birth). Stratum was calculated by combining region with residence (urban/rural), and was included based on guidance from Stata to include sample design variables in multiple imputation models (20). The model took into account sample weights using woman's weight. Cluster was not included as a covariate in the model to avoid problems with model fit.

The specific type of imputation used was multiple imputation using chained equations (MICE),[2] which is designed for data that are missing-at-random, but not missing-completely-at-random. Missing-at-random is a phrase used in statistics to indicate that the missing data can be predicted by other observed variables (20). The following Stata code was used for imputation:

```
mi impute chained (truncreg,ll(0)) Birthweight MaternalHeight MaternalBMI (ologit, augment) SizeAtBirth (logit, augment) ChildSex
Singleton (mlogit, augment) Parity Stratum [pweight=_sw], add(5) rseed(5258) force
```

Truncated linear regression was used for all continuous variables with lower limit set at 0, which allowed imputed values to go down to 0 rather than being bound by the observed range. All logistic regression used the augment option to prevent perfect prediction (20, 28). Five imputations were selected based on previous literature showing they provide valid inferences when 50% of the data is missing (20, 29).

Multiple imputation produced five datasets with no missing values. The next step was to create smoothed distributions with a finite mixture model for each of the five datasets to remove heaping. If the multiple imputation model did not fit and the survey had <5% missing birthweight, it was included in finite mixture modelling without imputation.

---

1 DHS typically included all variables. MICS included only size at birth and parity until MICS 5, when child sex and multiple/singleton status became available. Other survey types had different sets of variables available for multiple imputation.

2 Stata documentation describes MICE as also being known in the literature as imputation using fully conditional specifications and as sequential regression multivariate imputation.

Finite mixture model

The LBW Estimates fitted two normal distributions to each of the five imputed datasets with the user-written Stata code for finite mixture models. For surveys with <5% missing birthweight and no imputation (either because the imputation model did not fit or because the survey did not collect mother's perception of size), the LBW Estimates fitted two normal distributions on a single dataset with the user-written Stata code for finite mixture models:

```
fmm Birthweight [pweight=_sw], mix(normal) comp(2) cluster(psu) diff iterate(100)
```

For consistency the 2022 LBW Estimates used the user-written **fmm** in Stata and not the official Stata **fmm** that was introduced in Stata 15. Before fitting models on the five imputed datasets, biologically implausible birthweights were again removed (having been removed previously before imputation) in case imputed values fell outside of the range. With no covariates in the model, along with normal distribution and two components specified, Stata's **fmm** command produced a mean, standard deviation and population proportion for two normally distributed sub-populations in each of the five datasets. The model took into account the sample design by including sampling weights (woman's sample weight) and cluster and used a maximum of 100 iterations to find the best model fit. The maximization option of difficult was specified to account for non-concave regions.

Multiple imputation produced five datasets with no missing values and finite mixture modelling produced two distributions per dataset with a mean, standard deviation and population proportion (total of 10 distributions). The next sub-section describes the final step of converting those 10 distributions into a single estimate of LBW prevalence.

Producing a final adjusted estimate of low birthweight prevalence

For each of the five datasets, two—point estimates of LBW prevalence were produced (one for each distribution). The first step in producing a point estimate from a distribution was to calculate a z-score for the LBW cut-off of 2500 g. The mean and standard deviation from each finite mixture model distribution went into the following formula to produce a z-score:

$$Z_{2500} = \frac{2500\ g - mean\ birthweight}{standard\ deviation_{birthweight}}$$

The LBW cut-off z-score was then converted into a percentage with the use of a standard normal table (i.e., z-table), which provides the percentage of the distribution to the left of the z-score (i.e., area under the curve) — in this case the area left of the z-score represented the percentage of children with a birthweight less than 2500 g. For each dataset, the two point estimates of LBW prevalence were combined using a weighted average:

$$(LBW\%_{SP1} * \pi_{SP1}) + (LBW\%_{SP2} * \pi_{SP2})$$

where SP1 and SP2 are sub-populations (i.e., distributions) and $\boldsymbol{\pi}$ is the population proportion, or the proportion that each sub-population contributes to the overall population. The combination was carried out using Stata's postestimation commands for finite mixture models. After combining point estimates from the two sub-populations, each dataset had a single point estimate for LBW prevalence. The mean of the five point estimates (one for each dataset) was taken as a single, adjusted LBW prevalence reported for the survey.

The final step in producing adjusted estimates was calculating a combined variance to create a 95% confidence interval around the adjusted LBW prevalence. The combined variance took into account both within-dataset variance and between-dataset variance for all five datasets. Within-dataset variance was pooled by simply taking the mean of the five variances produced when a single point estimate was created for each dataset using finite mixture modelling postestimation. The between-dataset variance was calculated by putting the point estimate from each dataset into a standard variance formula:

$$\sigma^2{}_{Between} = \frac{\Sigma(x_i - \bar{x})^2}{n-1}$$

; where x is the LBW prevalence from dataset i and $\bar{x}$ is the mean LBW prevalence across the five datasets.

The two variances were combined with the following formula:

$$Total\ \sigma^2 = \sigma^2{}_{Within} + \left(1 + \frac{1}{n}\right) * \sigma^2{}_{Between}$$

; where n is the number of datasets, $\sigma^2{}_{Between}$ is the within-dataset variance and is the variance between datasets. Total variance was then transformed into standard deviation by taking the square root and multiplying it by ±1.96 to give a two-sided 95% confidence interval.

By taking into account both within- and between-dataset variance, the confidence intervals included the uncertainty arising from both imputation and the use of finite mixture modelling.

ANNEX 3

Modelling: Identifying and selecting covariates

In keeping with Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) principals [5] we aimed to clearly document the principles and process for covariate selection and covariate time-series data for inclusion in Bayesian modelling of global LBW and preterm birth estimates. Covariate data were derived from multiple sources and followed a standardized approach. To model estimates of LBW prevalence and preterm birth rate, the most up-to-date country-level covariates from available United Nations and other sources was used. The selection of covariates to include in the modelling was done a priori following the three steps outlined below.

Step 1: Identify biologically plausible risk factors and predictors of LBW and Preterm births

- Develop a conceptual framework for risk factors and predictors of LBW and preterm birth
- Assess which risk factors and predictors may have time-series covariate data available

Step 2: Assess availability of potential covariates

- Compile existing time series for potential covariates identified in the conceptual framework, including information on methodology used
- For potential covariates with no existing time series:
  - ➜ *Assess availability of empirical data*
  - ➜ *Where empirical data are available, create new time series using pre-defined methods*
  - ➜ *Exclude covariate time series with no available input data for ≥20% of countries and time series generated using covariate-driven modelled estimates*

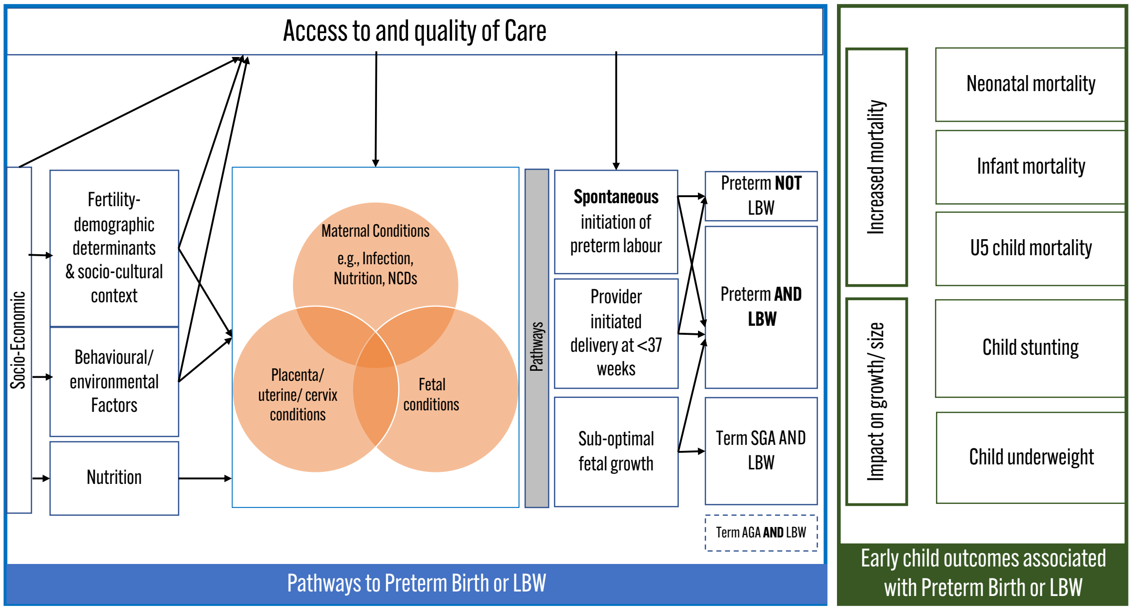
Step 3: Assess correlation between covariates and final covariate selection

- For potential covariates selected under step 2, assess:
  - ➜ *Correlation between covariates:*
    - Correlation matrix
    - Cluster analysis
    - Correlation with the outcome
- For each cluster of covariates, select the covariate most strongly correlated with the outcome to include a priori in the modelling process

Step 1: Identify biologically plausible and risk factors/predictors of LBW and preterm birth

Plausible predictors of LBW and preterm birth for use in the modelling process for producing preterm and LBW estimates were identified a priori by constructing a conceptual framework for predictors of these outcomes (Annex Figure 3.1, Annex Table 3.1). The framework was constructed by reviewing and adapting previously published related frameworks, including: Victora 1997[1]; Mosley and Chen 2003 [2]; Olusanya 2010 [3]; and Villar 2012 [4]. Predictors were based on biological plausibility and reported literature [5, 6] (Annex Figure 3.1).

Annex Figure 3.1: Conceptual framework for the identification of potential covariates / predictors for use in the preterm and LBW.



Annex Table 3.1: Potential predictors of LBW and preterm birth

	Domain	Potential associated factors
Pathways to preterm birth/ LBW	Socio-economic	Low socio-economic status [30]
	Fertility-demographic determinants and socio-cultural context	Low education [31], short inter-pregnancy interval [32], higher parity [33], births at extremes of reproductive ages (adolescents, older maternal age) [33], rural residence [30, 34]
	Behavioural/ environmental	Smoking [35, 36], maternal substance abuse [37-41], Heavy alcohol use [42], intimate partner violence [43, 44], indoor air pollution [45], outdoor air pollution [46, 47], physical/ occupational activity [48]
	Nutrition	Maternal underweight [49], maternal overweight [50], maternal short stature [51], maternal micronutrient deficiency [52, 53], maternal anaemia [54, 55]
	Maternal related	Maternal HIV [56, 57], malaria [58], syphilis [59], hypertension [60], diabetes (pre-existing or gestational) [56], depression in pregnancy [61]
	Placenta/ uterine/ cervix conditions	Placenta Praevia [62], placental abruption [63], uterine abnormalities [64]
	Fetal related	Twins [34, 65], birth defects, fetal growth restriction, male sex [66]
	Access to and quality of care	No/ suboptimal antenatal clinic attendance [31], availability of skilled care at birth, including caesarean section if needed [31]
Early Childhood	Impacts on child growth/ size*	Child underweight [67], child stunting [67],
	Increased early mortality	Neonatal mortality [68], infant mortality [68]

* LBW prevalence is also an associated factor for preterm birth

Potential covariates for the analysis were grouped into domains: (1) socio-economic, demographic, fertility and cultural factors; (2) nutritional, behavioural, and environmental factors; (3) maternal conditions (including infections); (4) fetal or placental conditions; (5) healthcare-related factors (markers to access to care); and (6) early childhood outcomes associated with LBW/ preterm. Across the six domains, 41 potential covariates, for which it was considered that time series data may be available, were identified (Annex Table 3.1, Annex Table 3.2).

Annex Table 3.2: Availability of covariate time series by domain

Domain	Potential covariate	Availability of time series estimates or empirical raw data
Socio-economic, demographic, fertility and cultural factors	Gross national income	Empirical raw data available
	Gross domestic product	Existing time series data or estimates
	GINI coefficient	Empirical raw data available
	Adult female literacy rate	Empirical raw data available
	Mean years female education	Empirical raw data available
	Adolescent birth rate	Empirical raw data available
	Total fertility rate	Existing time series data or estimates
	General fertility rate	Existing time series data or estimates
	Modern contraceptive prevalence rate	Existing time series data or estimates
	Proportion of live births to mothers aged 35 years and older	Not available
	Urban population	Existing time series data or estimates
Nutritional, behavioral, and environmental factors	Adult female smoking rate	Existing time series data or estimates
	Indoor air pollution	Existing time series data or estimates
	Outdoor air pollution	Existing time series data or estimates
	Adult female body mass index (Mean)	Existing time series data or estimates
	Underweight women of reproductive age	Existing time series data or estimates
	Overweight women of reproductive age	Existing time series data or estimates
	Maternal anaemia	Existing time series data or estimates
	Adult female substance use	Not available
	Intimate partner violence	Not available
Maternal conditions (including infections)	Maternal mortality rate	Existing time series data or estimates
	Adult female HIV prevalence	Existing time series data or estimates
	Malaria incidence (P. falciparum parasite rate)	Existing time series data or estimates
	Insecticide-treated nets coverage	Existing time series data or estimates
	Adult female syphilis prevalence	Existing time series data or estimates
	Gestational hypertension	Not available
	Gestational diabetes	Not available
	Maternal depression	Not available
Fetal or placental conditions	Twinning	Not available
	Birth defects	Not available
	Growth restriction	Not available
Healthcare-related factors (markers of access to care)	Antenatal care attendance (four or more times)	Empirical raw data available
	Skilled birth attendance	Empirical raw data available
	Facility birth rate	Empirical raw data available
	Caesarean section rate	Empirical raw data available
Early childhood outcomes associated with LBW/ preterm	Neonatal mortality rate	Existing time series data or estimates
	Stunting in children under 5 years	Existing time series data or estimates
	Underweight in children under 5 years	Existing time series data or estimates
	Low birthweight prevalence	Modelling underway as part of this work*

Annex Table 3.3: Overview of potential covariates with existing time-series estimates

Covariate name	Author [Ref]	Methods	Covariate-driven model	Number of countries estimates available for	Number of countries contributing ≥ 1 year of input data (2000–2020)	Data sources used
Adult female mean body mass index	WHO Global Health Observatory (GHO) (NCD-RisC)[8, 9]	Hierarchical model based on country's own data, where available, or informed by regional data if not. lnGDP used as a covariate.	No	190	176	Data collected on samples of a national, subnational or community population.
Prevalence of overweight among female adults	WHO GHO (NCD-RisC)[8, 9]	Hierarchical model based on country's own data, where available, or informed by regional data if not. lnGDP used as a covariate.	No	190	176	Data collected on samples of a national, subnational or community population.
Prevalence of underweight among female adults	WHO GHO (NCD-RisC)[8, 9]	Hierarchical model based on country's own data, where available, or informed by regional data if not. lnGDP used as a covariate.	No	190	176	Data collected on samples of a national, subnational or community population.
Women of reproductive age HIV prevalence	UNAIDS [10]	Four methods used (mathematical modelling, model-based geo-statistics, small area estimation and direct estimates from prevalence surveys).	Yes	119	140 for HIV prevalence at any age	Epidemiological data with HIV prevalence, nationally presentative population-based surveys, surveillance and routine data, data from vital registration systems.
Adult female smoking rate	Institute for Health Metrics and Evaluation (IHME) [11]	Spatiotemporal Gaussian process regression (ST-GPR). No covariates.	No	193	192	Nationally representative surveys, including both multinational and country-specific surveys. For countries without data, estimates were entirely based on models
Prevalence of anaemia in pregnant women	WHO GHO [9]	A Bayesian hierarchical mixture model including covariates of maternal education, national income, urbanization, and an aggregate metric of access to basic healthcare [12].	Yes	195	124	408 population-representative data sources from 124 countries worldwide.
Adult female syphilis prevalence	IHME [13]	Bayesian meta-regression method. Covariates include diagnostic methods and HIV prevalence.	Yes	169	137 (for overall syphilis prevalence – not specified if all female)	Including published literature, surveillance data, survey data, hospital and clinical data, and other types of data
Modern contraceptive prevalence rate	Department of Economic and Social Affairs, United Nations Population Division	Bayesian hierarchical model [15]	No	187	164 countries at least 1 survey since 2000	1,247 surveys and estimates of modern contraceptive prevalence rate
General fertility rate	UN World Population Prospects (WPP) [16]	Cohort-component method	No	197	Not available	Based on a range of population and household surveys

Covariate name	Author [Ref]	Methods	Covariate-driven model	Number of countries estimates available for	Number of countries contributing ≥ 1 year of input data (2000–2020)	Data sources used
Total fertility rate	UN WPP [16]	Various estimation methods	No	197	159 countries at least 1 data point since 2000	Survey, census, estimates.
Maternal mortality rate	WHO GHO [9]	Bayesian approach with covariates: Gross domestic product per capita based on purchasing power parity conversion (GDP), general fertility rate (GFR), and proportion of births attended by a skilled health worker (SAB) [17].	Yes	183	177	Civil registration with complete coverage and medical certification of cause of death, household surveys, population census, sample or sentinel registration systems, special studies.
Neonatal mortality rate	UN Interagency Group for Child Mortality Estimation (IGME) [18]	New Bayesian data-driven model to capture trends in NMR within countries and over time for all countries. No covariates used.	No	195	195	Nationally representative estimates of under-five mortality can be derived from several different sources, including civil registration and sample surveys.
Stunting in children under 5 years	UNICEF-WHO-World Bank Joint Child Malnutrition Estimates (JME) group [19]	Prevalence was modelled at logit (log-odds) scale using a penalized longitudinal mixed-model with a heterogeneous error term. Covariates linear and quadratic socio-demographic index (SDI), data source type, average health system access over the previous five years	Yes	154	Not available	These national-level data sources are mainly comprised of household surveys. Some administrative data sources are also included where population coverage is high.
Insecticide-treated net coverage	Malaria Atlas Project [20]	Bayesian hierarchical model. Covariates used for spatial modelling [21]	Yes	40	40	Household-level survey data (DHS, MICS), national level aggregated survey data (MIS), net distribution data (WHO)
Malaria prevalence	Malaria Atlas Project [20]	Binomial generalized linear model (GLM), with a dependent variable consisting of positive and negative P. falciparum counts tabulated for each survey cluster, and the independent variables consisting of 20 covariates [22]	Yes	196	Not available	Published literature, other household surveys, personal communication and other sources

Covariate name	Author [Ref]	Methods	Covariate-driven model	Number of countries estimates available for	Number of countries contributing ≥ 1 year of input data (2000–2020)	Data sources used
Outdoor air pollution (PM2.5)	WHO – GHO [9]	Bayesian hierarchical modelling framework. Covariates included - not specified [23]	Yes	194	Not specified but paper suggests 9,690 monitoring locations around the world	Ground measurements from 9,690 monitoring locations around the world from the WHO cities database together with satellite remote sensing, population estimates, topography, and information on local monitoring networks and measures of specific contributors of air pollution from chemical transport models.
Indoor air pollution	IHME [24]	Three-step modelling strategy using linear regression, spatiotemporal regression, and Gaussian process regression (GPR). Included covariates maternal education and the proportion of population living in urban areas.	Yes	193	195	Case-control data. To fill the gaps of data in surveys and censuses, we also downloaded and updated estimates from WHO Energy Database and extracted from literature through systematic review.
GDP	World Bank [25]	Aggregation method: Gap-filled total, calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources.	No	197	196	World Bank national accounts data, and Organisation for Economic Cooperation and Development (OECD) National Accounts data files.
Urban population	UN Population Division [26]	Estimated from most recent national census or official population estimate of each country and extrapolated to the base year when necessary. Taiwan Province of China imputed.	No	197	197	Census or official population estimate

Step 2: Assess availability and quality of potential covariates

We mapped the forty-one potential covariates from the conceptual framework of predictors to assess their data availability. Potential covariates were considered for inclusion as a predictor in the model for the LBW and preterm outcomes if empirical data or estimates were available for most countries covering the years of the input data (2000–2019). Availability of existing time series data was assessed through consultation with WHO and UNICEF and via a targeted search of the webpages of United Nations organizations (e.g., WHO Global Health Observatory, UNICEF, United Nations Population Division) and academic groups (e.g., IHME).

For potential covariates with no existing time series estimates available, UNICEF and WHO databases were searched for empirical data available for these variables. Where comparable, but incomplete time series data were located for a given covariate, a new time series was generated for the years 2000–2020 using a standard approach for in-filling and extrapolation. For countries with some empirical data, linear interpolation and constant backwards and forwards extrapolation was used. For countries that did not have empirical data available, values were imputed using a regression based on region and country's lag distributed GDP and World Bank region. Finally for all countries, a smoothed time series was generated using a seven-year average for model prediction using existing methods for United Nations estimates [7].

Annex Tab le 3.2 shows the data availability for the 41 potential covariates, grouped by domain. Three potential covariates were excluded at this stage: under-5 child and infant mortality, as data-driven estimates are available for the more proximal neonatal mortality rate covariate; and LBW, which is a potential covariate for preterm birth only, and will use new LBW estimates generated as part of this updated modelling work rather than existing time series estimates.

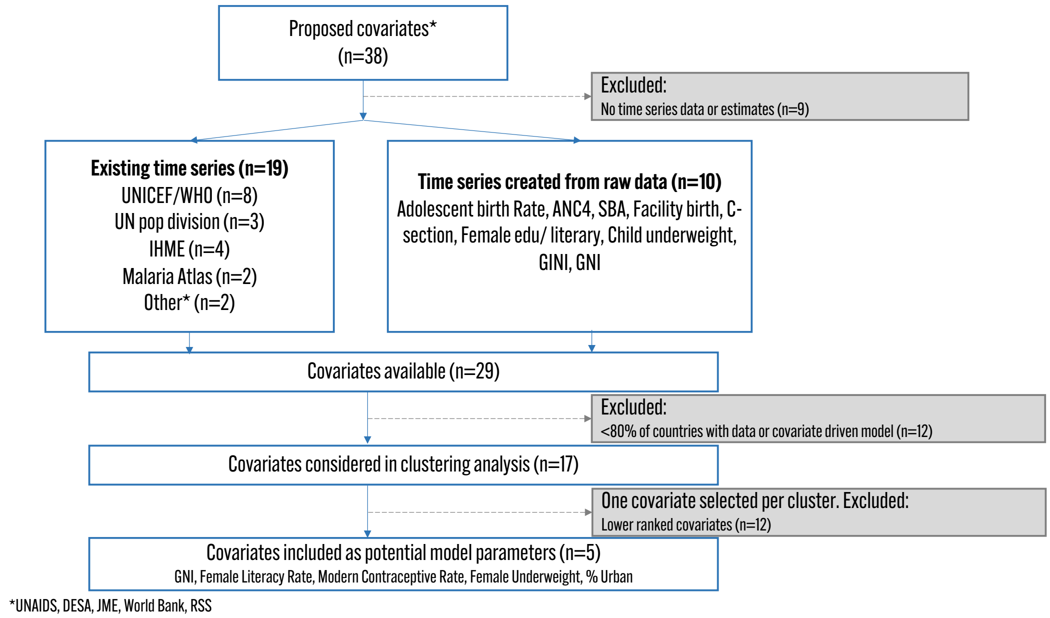
Of the remaining 38 potential variables, existing time series data or estimates were available for 19, and empirical raw data available for a further 10 (Annex Table 3.4).

Annex Table 3.4: Overview of covariates with empirical data only

Covariate name [Reference source]	Number of countries with ≥ 1 year data (2000–2019)	Percent of data points 2000–2020 computed by method below (%)		
		Extrapolated assuming a flat trend	Linear interpolated	Imputed by year and World Bank region
Adult female literacy rate % [27]	157	29.1	40.5	19.3
Mean years female education [28]	168	35.7	25.6	15.2
Adolescent birth rate [9, 16]	194	16.9	22.8	0.5
Antenatal care coverage % (at least four visits) [9]	152	29.3	31.8	21.8
Births attended by skilled health personnel % [9]	187	16.7	29.1	5.1
Institutional deliveries – percentage of deliveries in a health facility [9]	173	20.9	32.1	11.7
Births by caesarean section (%) [9]	178	27.3	26.5	8.1
Underweight in children under 5 years [9]	152	23.3	39.3	20.8
GINI coefficient [29]	159	34.9	24.5	16.8
GNI [29]	180	22.4	0.4	8.1

No systematic comparable time series data or estimates were available for the remaining 9 covariates. Annex Figure 3.2 shows the flow figure for covariate selection for LBW model.

Annex Figure 3.2: Flow figure for covariates selection for the LBW model

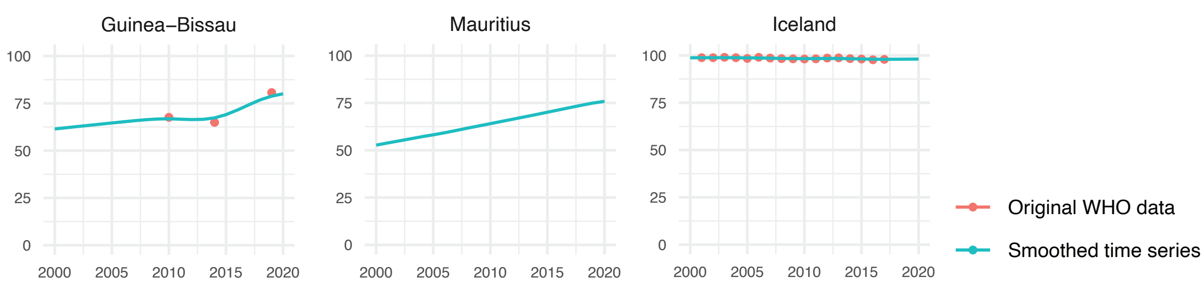


Data-driven time series, with empirical data for all years (2000–2020) would be ideal covariates for this work. However, in the absence of complete time-series data, covariates which are most data driven, with clear explanations of the method of creation (including any modelling or infilling etc.) would be preferable. We therefore assessed all existing time series covariates estimates considering: data source; quantity of empirical data informing time series; number of country-years estimated and methods used to produce time-series, including any modelling, infilling, smoothing, extrapolations or any other data manipulations.

Assessment of consistency of time-series

For all 29 potential covariates, time-series country plots were generated to assess the consistency of empirical data or time series estimates and to identify any outliers across the time periods 2000–2020. Annex Figure 3.3 is an example of these country plots for the covariate antenatal care coverage – at least four visits (%) (ANC4+), for three countries. For newly created time-series, the plots were used to visually assess the availability of the data informing the time-series.

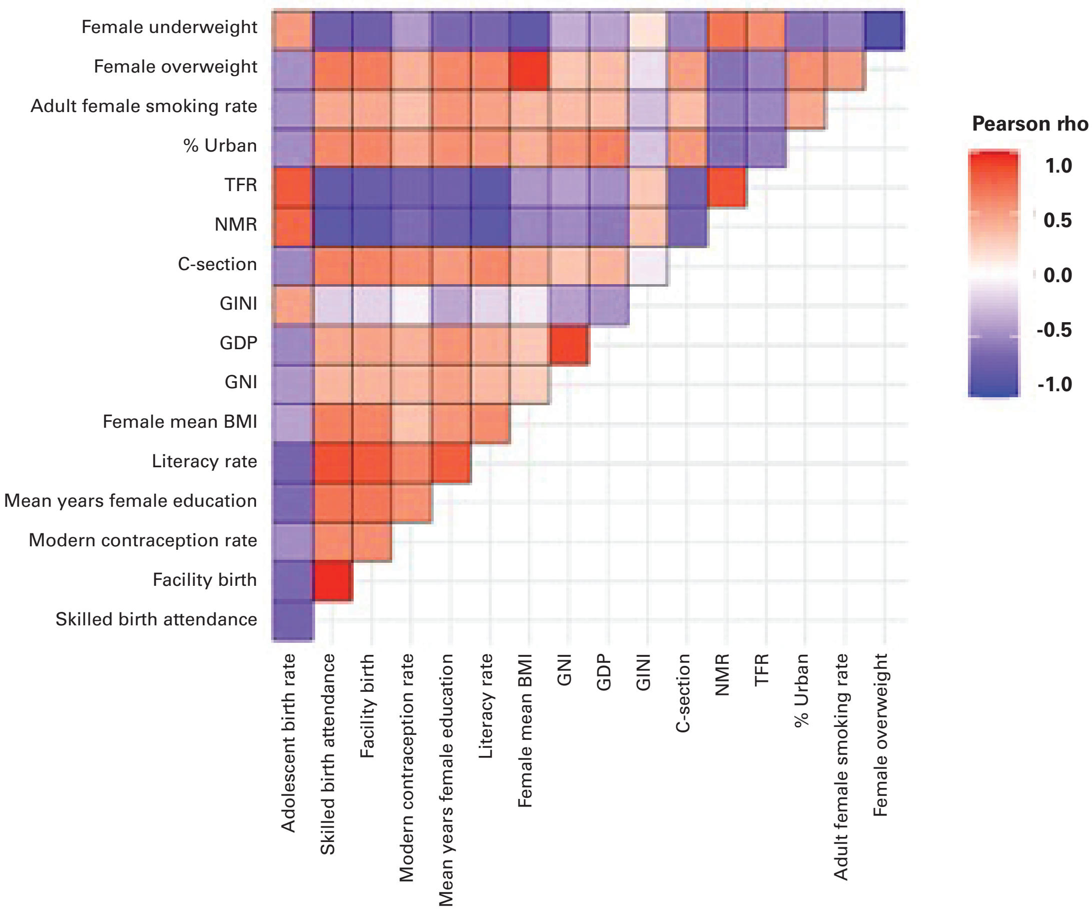
Annex Figure 3.3: Country plots for Guinea-Bissau, Mauritania and Iceland for the covariate ANC4+



Step 3: Assessing correlation between covariates

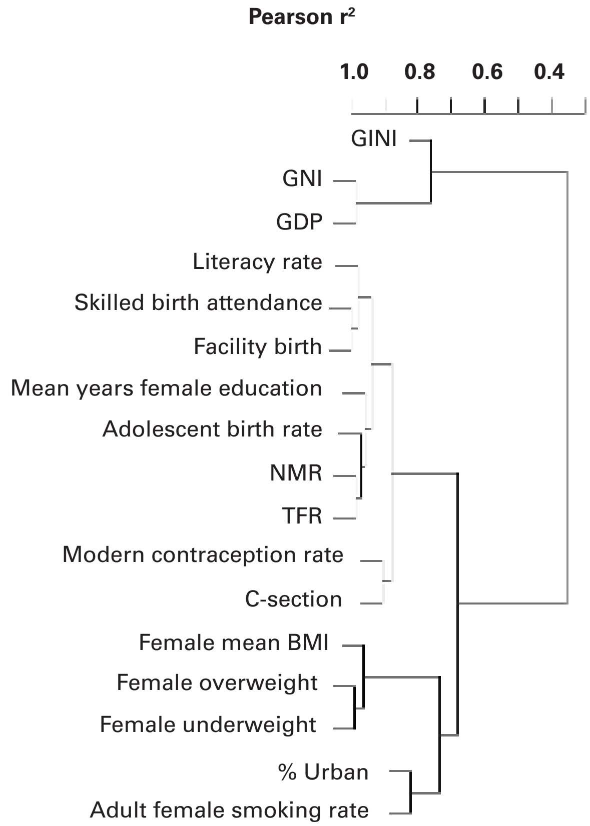
Next, for potential covariates selected under step 2, the correlation between covariates was assessed using a correlation matrix and cluster analysis. The Pearsons correlation between the potential predictors of the outcomes was assessed. Annex Figure 3.4 shows the covariate correlation matrix used to identify the variables for LBW prediction which were most highly correlated.

Annex Figure 3.4: Covariate correlation matrix for potential covariates for predicting LBW



Based on Pearson's r^2, covariates were grouped into 'clusters' dependent on the correlations between them. See Annex Figure 3.5 for the cluster analysis for LBW predictors.

Annex Figure 3.5: Cluster analysis for potential covariates for predicting LBW



A final analysis of the correlation between the potential covariates and the outcome (LBW or preterm birth) was undertaken. For each cluster, the potential covariate most closely correlated with the outcome was selected to be included in the modelling of the outcome. Where two potential covariates had similar associations with the outcome, the potential covariate with the most input data, or the one in most common use was retained. This allowed us to build a parsimonious model and avoid overfitting by limiting the number of covariates. The example from LBW is shown in Annex Figure 3.6.

Annex Figure 3.6: Correlation between potential covariates for predicting LBW and LBW outcome

Covariate	Correlation with outcome
C-section	-0.100
Female HIV	0.190
GDP	-0.221
GNI	-0.250
Female mean BMI	-0.265
Modern contraceptive prevalence rate	-0.280
Malaria (PfPR)	0.319
GINI	0.338
Outdoor air pollution	0.375
% population that is urban	0.375
Female overweight prevalence	-0.387
Syphilis rate	0.392
Adult female smoking rate	0.401
ANC4	-0.415
Child (under 5 years) stunting prevalence	0.423
Indoor air pollution	0.424
Adolescent birth rate	0.428
Total fertility rate	0.436
General fertility rate	0.444
Neonatal mortality rate	0.464
Anaemia	0.496
Maternal mortality ratio	0.512
Facility birth rate	-0.528
Skilled birth attendance rate	-0.537
Mean years female education	-0.556
Literacy rate	-0.563
Female underweight prevalence	0.601
Child (under 5 years) underweight prevalence	0.615

The cluster 1 GNI and GDP had similar correlation with LBW. GNI was selected as it is preferred as an economic indicator. For cluster 5, percentage urban was selected over adult female smoking as more data are available to inform the urban time series compared to smoking.

The final covariates for the LBW model were gross national income per person purchasing power parity (constant 2017 international $), female literacy rate, modern contraception prevalence rate, adult female underweight prevalence and percent urban population. For the preterm birth model, LBW prevalence was the only covariate used.

References

1. Victora CG, Huttly SR, Fuchs SC, Olinto MT: The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. *Int J Epidemiol* 1997, 26(1):224-227.

2. Mosley WH, Chen LC: An analytical framework for the study of child survival in developing countries. 1984. *Bull World Health Organ* 2003, 81(2):140-145.

3. Olusanya BO, Ofovwe GE: Predictors of preterm births and low birthweight in an inner-city hospital in sub-Saharan Africa. *Matern Child Health J* 2010, 14(6):978-986.

4. Villar J, Papageorghiou AT, Knight HE, Gravett MG, Iams J, Waller SA, Kramer M, Culhane JF, Barros FC, Conde-Agudelo A *et al*: The preterm birth syndrome: a prototype phenotypic classification. *Am J Obstet Gynecol* 2012, 206(2):119-123.

5. Bryce E, Gurung S, Tong H, Katz J, Lee AC, Black RE, Walker N: Population attributable fractions for risk factors for spontaneous preterm births in 81 low- and middle-income countries: A systematic analysis. *J Glob Health* 2022, 12:04013.

6. Gurung S, Tong HH, Bryce E, Katz J, Lee AC, Black RE, Walker N: A systematic review on estimating population attributable fraction for risk factors for small-for-gestational-age births in 81 low- and middle-income countries. *J Glob Health* 2022, 12:04024.

7. Liu L, Villavicencio F, Yeung D, Perin J, Lopez G, Strong KL, Black RE: National, regional, and global causes of mortality in 5-19-year-olds from 2000 to 2019: a systematic analysis. *Lancet Glob Health* 2022, 10(3):e337-e347.

8. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet* 2017, 390(10113):2627-2642.

9. World Health Organization, The Global Health Observatory. https://www.who.int/data/gho

10. UNAIDS: UNAIDS data 2019 https://www.unaids.org/sites/default/files/media_asset/2019-UNAIDS-data_en.pdf.

11. Reitsma MB, Flor LS, Mullany EC, Gupta V, Hay SI, Gakidou E: Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and initiation among young people in 204 countries and territories, 1990-2019. *Lancet Public Health* 2021, 6(7):e472-e481.

12. Finucane MM, Paciorek CJ, Stevens GA, Ezzati M: Semiparametric Bayesian Density Estimation With Disparate Data Sources: A Meta-Analysis of Global Childhood Undernutrition. *Journal of the American Statistical Association* 2015, 110(511):889-901.

13. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018, 392(10159):1789-1858.

14. United Nations, Department of Economic and Social Affairs, Population Division (2022). World Contraceptive Use 2022 and Estimates and Projections of Family Planning Indicators 2022. https://www.un.org/development/desa/pd/data/family-planning-indicators

15. Wheldon M KV, Ueffing P, Dasgupta A,: Population Division. Technical Paper No. 2018/2 Methods for Estimating and Projecting Key Family Planning Indicators Among All Women of Reproductive Age. *https://wwwunorg/en/development/desa/population/publications/pdf/technical/TP2018-2pdf* 2018.

16. United Nations Department of Economic and Social Affairs: World Population Prospects https://population.un.org/wpp2019/DataSources/. 2019.

17. Alkema L, Chou D, Hogan D, Zhang S, Moller AB, Gemmill A, Fat DM, Boerma T, Temmerman M, Mathers C *et al*: Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. *Lancet* 2016, 387(10017):462-474.

18. UN-Interagency Group for Child Mortality Estimation: https://childmortality.org/.

19. World Health Organization, United Nations Children's Fund, World Bank: Levels and trends in child malnutrition: UNICEF / WHO / The World Bank Group joint child malnutrition estimates: key findings of the 2021 edition. Geneva: World Health Organization; 2021.

20. Malaria Atlas Project: https://malariaatlas.org/research-project/metrics-of-insecticide-treated-nets-distribution/.

21. Bertozzi-Villa A, Bever CA, Koenker H, Weiss DJ, Vargas-Ruiz C, Nandi AK, Gibson HS, Harris J, Battle KE, Rumisha SF *et al*: Maps and metrics of insecticide-treated net access, use, and nets-per-capita in Africa from 2000-2020. *Nature Communications* 2021, 12(1):3589.

22. Weiss DJ, Mappin B, Dalrymple U, Bhatt S, Cameron E, Hay SI, Gething PW: Re-examining environmental correlates of Plasmodium falciparum malaria endemicity: a data-intensive variable selection approach. *Malaria Journal* 2015, 14(1):68.

23. Shaddick G, Thomas ML, Mudu P, Ruggeri G, Gumy S: Half the world's population are exposed to increasing air pollution. *npj Climate and Atmospheric Science* 2020, 3(1):23.

24. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020, 396(10258):1223-1249.

25. World Bank: World Development Indicators - Gross Domestic Product (GDP). https://data.worldbank.org/. In. Washington DC, United States of America: World Bank.

26. United Nations Department of Economic and Social Affairs: World Urbanization Prospects https://population.un.org/wup/. 2018.

27. World Bank: World Development Indicators - Literacy rate, adult female https://data.worldbank.org/.

28. United Nations Educational Scientific and Cultural Organization (UNESCO): UNESCO Institute for Statistics http://data.uis.unesco.org/#.

29. World Bank: World Bank National Accounts Data https://data.worldbank.org/.

30. Amjad S, MacDonald I, Chambers T, Osornio-Vargas A, Chandra S, Voaklander D, Ospina MB: Social determinants of health and adverse maternal and birth outcomes in adolescent pregnancies: A systematic review and meta-analysis. *Paediatr Perinat Epidemiol* 2019, 33(1):88-99.

31. Zaveri A, Paul P, Saha J, Barman B, Chouhan P: Maternal determinants of low birth weight among Indian children: Evidence from the National Family Health Survey-4, 2015-16. *PLoS One* 2020, 15(12):e0244562.

32. Kozuki N, Lee AC, Silveira MF, Victora CG, Adair L, Humphrey J, Ntozini R, Black RE, Katz J: The associations of birth intervals with small-for-gestational-age, preterm, and neonatal and infant mortality: a meta-analysis. *BMC Public Health* 2013, 13 Suppl 3(Suppl 3):S3.

33. Kozuki N, Lee AC, Silveira MF, Sania A, Vogel JP, Adair L, Barros F, Caulfield LE, Christian P, Fawzi W *et al*: The associations of parity and maternal age with small-for-gestational-age, preterm, and neonatal and infant mortality: a meta-analysis. *BMC Public Health* 2013, 13 Suppl 3(Suppl 3):S2.

34. Alamneh TS, Teshale AB, Worku MG, Tessema ZT, Yeshaw Y, Tesema GA, Liyew AM, Alem AZ: Preterm birth and its associated factors among reproductive aged women in sub-Saharan Africa: evidence from the recent demographic and health surveys of sub-Sharan African countries. *BMC Pregnancy Childbirth* 2021, 21(1):770.

35. Shah NR, Bracken MB: A systematic review and meta-analysis of prospective studies on the association between maternal cigarette smoking and preterm delivery. *Am J Obstet Gynecol* 2000, 182(2):465-472.

36. Salmasi G, Grady R, Jones J, McDonald SD: Environmental tobacco smoke exposure and perinatal outcomes: a systematic review and meta-analyses. *Acta Obstet Gynecol Scand* 2010, 89(4):423-441.

37. Marchand G, Masoud AT, Govindan M, Ware K, King A, Ruther S, Brazil G, Ulibarri H, Parise J, Arroyo A *et al*: Birth Outcomes of Neonates Exposed to Marijuana in Utero: A Systematic Review and Meta-analysis. *JAMA Netw Open* 2022, 5(1):e2145653.

38. De Silva M, Panisi L, Brownfoot FC, Lindquist A, Walker SP, Tong S, Hastie R: Systematic review of areca (betel nut) use and adverse pregnancy outcomes. *Int J Gynaecol Obstet* 2019, 147(3):292-300.

39. Dos Santos JF, de Melo Bastos Cavalcante C, Barbosa FT, Gitaí DLG, Duzzioni M, Tilelli CQ, Shetty AK, de Castro OW: Maternal, fetal and neonatal consequences associated with the use of crack cocaine during the gestational period: a systematic review and meta-analysis. *Arch Gynecol Obstet* 2018, 298(3):487-503.

40. Gouin K, Murphy K, Shah PS: Effects of cocaine use during pregnancy on low birthweight and preterm birth: systematic review and metaanalyses. *Am J Obstet Gynecol* 2011, 204(4):340.e341-312.

41. Shatil B, Landau R: Opioid Use and Misuse in Pregnancy. *Clin Perinatol* 2020, 47(4):769-777.

42. Patra J, Bakker R, Irving H, Jaddoe VW, Malini S, Rehm J: Dose-response relationship between alcohol consumption before and during pregnancy

and the risks of low birthweight, preterm birth and small for gestational age (SGA)-a systematic review and meta-analyses. *Bjog* 2011, 118(12):1411-1421.

43. Donovan BM, Spracklen CN, Schweizer ML, Ryckman KK, Saftlas AF: Intimate partner violence during pregnancy and the risk for adverse infant outcomes: a systematic review and meta-analysis. *Bjog* 2016, 123(8):1289-1299.

44. Hill A, Pallitto C, McCleary-Sills J, Garcia-Moreno C: A systematic review and meta-analysis of intimate partner violence during pregnancy and selected birth outcomes. *Int J Gynaecol Obstet* 2016, 133(3):269-276.

45. Amegah AK, Quansah R, Jaakkola JJ: Household air pollution from solid fuel use and risk of adverse pregnancy outcomes: a systematic review and meta-analysis of the empirical evidence. *PLoS One* 2014, 9(12):e113920.

46. Sun X, Luo X, Zhao C, Chung Ng RW, Lim CE, Zhang B, Liu T: The association between fine particulate matter exposure during pregnancy and preterm birth: a meta-analysis. *BMC Pregnancy Childbirth* 2015, 15:300.

47. Zhu X, Liu Y, Chen Y, Yao C, Che Z, Cao J: Maternal exposure to fine particulate matter (PM2.5) and pregnancy outcomes: a meta-analysis. *Environ Sci Pollut Res Int* 2015, 22(5):3383-3396.

48. Cai C, Vandermeer B, Khurana R, Nerenberg K, Featherstone R, Sebastianski M, Davenport MH: The impact of occupational activities during pregnancy on pregnancy outcomes: a systematic review and metaanalysis. *Am J Obstet Gynecol* 2020, 222(3):224-238.

49. Han Z, Mulla S, Beyene J, Liao G, McDonald SD: Maternal underweight and the risk of preterm birth and low birth weight: a systematic review and meta-analyses. *Int J Epidemiol* 2011, 40(1):65-101.

50. McDonald SD, Han Z, Mulla S, Beyene J: Overweight and obesity in mothers and risk of preterm birth and low birth weight infants: systematic review and meta-analyses. *Bmj* 2010, 341:c3428.

51. Kozuki N, Katz J, Lee AC, Vogel JP, Silveira MF, Sania A, Stevens GA, Cousens S, Caulfield LE, Christian P *et al*: Short Maternal Stature Increases Risk of Small-for-Gestational-Age and Preterm Births in Low- and Middle-Income Countries: Individual Participant Data Meta-Analysis and Population Attributable Fraction. *J Nutr* 2015, 145(11):2542-2550.

52. Carducci B, Keats EC, Bhutta ZA: Zinc supplementation for improving pregnancy and infant outcome. *Cochrane Database Syst Rev* 2021, 3(3):Cd000230.

53. Hofmeyr GJ, Lawrie TA, Atallah Á N, Torloni MR: Calcium supplementation during pregnancy for preventing hypertensive disorders and related problems. *Cochrane Database Syst Rev* 2018, 10(10):Cd001059.

54. Rahman MM, Abe SK, Rahman MS, Kanda M, Narita S, Bilano V, Ota E, Gilmour S, Shibuya K: Maternal anemia and risk of adverse birth and health outcomes in low- and middle-income countries: systematic review and meta-analysis. *Am J Clin Nutr* 2016, 103(2):495-504.

55. Young MF, Oaks BM, Tandon S, Martorell R, Dewey KG, Wendt AS: Maternal hemoglobin concentrations across pregnancy and maternal and child health: a systematic review and meta-analysis. *Ann N Y Acad Sci* 2019, 1450(1):47-68.

56. Hedderson MM, Ferrara A, Sacks DA: Gestational diabetes mellitus and lesser degrees of pregnancy hyperglycemia: association with increased risk of spontaneous preterm birth. *Obstet Gynecol* 2003, 102(4):850-856.

57. Wedi CO, Kirtley S, Hopewell S, Corrigan R, Kennedy SH, Hemelaar J: Perinatal outcomes associated with maternal HIV infection: a systematic review and meta-analysis. *Lancet HIV* 2016, 3(1):e33-48.

58. Vogel JP, Lee AC, Souza JP: Maternal morbidity and preterm birth in 22 low- and middle-income countries: a secondary analysis of the WHO Global Survey dataset. *BMC Pregnancy Childbirth* 2014, 14:56.

59. Qin J, Yang T, Xiao S, Tan H, Feng T, Fu H: Reported estimates of adverse pregnancy outcomes among women with and without syphilis: a systematic review and meta-analysis. *PLoS One* 2014, 9(7):e102203.

60. Morisaki N, Togoobaatar G, Vogel JP, Souza JP, Rowland Hogue CJ, Jayaratne K, Ota E, Mori R: Risk factors for spontaneous and provider-initiated preterm delivery in high and low Human Development Index countries: a secondary analysis of the World Health Organization Multicountry Survey on Maternal and Newborn Health. *Bjog* 2014, 121 Suppl 1:101-109.

61. Ghimire U, Papabathini SS, Kawuki J, Obore N, Musa TH: Depression during pregnancy and the risk of low birth weight, preterm birth and intrauterine growth restriction- an updated meta-analysis. *Early Hum Dev* 2021, 152:105243.

62. Kollmann M, Gaulhofer J, Lang U, Klaritsch P: Placenta praevia: incidence, risk factors and outcome. *J Matern Fetal Neonatal Med* 2016, 29(9):1395-1398.

63. Ananth CV, Berkowitz GS, Savitz DA, Lapinski RH: Placental abruption and adverse perinatal outcomes. *Jama* 1999, 282(17):1646-1651.

64. Panagiotopoulos M, Tseke P, Michala L: Obstetric Complications in Women With Congenital Uterine Anomalies According to the 2013 European Society of Human Reproduction and Embryology and the European Society for Gynaecological Endoscopy Classification: A Systematic Review and Meta-analysis. *Obstet Gynecol* 2022, 139(1):138-148.

65. Vogel JP, Torloni MR, Seuc A, Betrán AP, Widmer M, Souza JP, Merialdi M: Maternal and perinatal outcomes of twin pregnancy in 23 low- and middle-income countries. *PLoS One* 2013, 8(8):e70549.

66. Zeitlin J, Saurel-Cubizolles MJ, De Mouzon J, Rivera L, Ancel PY, Blondel B, Kaminski M: Fetal sex and preterm birth: are males at greater risk? *Hum Reprod* 2002, 17(10):2762-2768.

67. 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):1340-1355.

68. Katz J, Lee AC, Kozuki N, Lawn JE, Cousens S, Blencowe H, Ezzati M, Bhutta ZA, Marchant T, Willey BA *et al*: Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *Lancet* 2013, 382(9890):417-425.

ANNEX 4
Modelling: Data quality categories to account for bias

Data quality categories

To assess and account for differences in data quality we first looked at five indicators: heaping index, missing birthweight percentage, denominator used to calculate LBW prevalence, recorded birthweight coverage[3], and HMIS versus not HMIS data. Annex Table 4.1 shows a summary of the availability of the data for these indicators.

Annex Table 4.1: Availability of data quality indicators

Data quality indicator	Source	Availability
Heaping index	Vulnerable Newborn Collaboration	21 countries with 240 country-years from national data
Missing birthweight percentage	Admin database	64 countries with 932 country-years
Denominator used to calculate LBW prevalence	Admin database	108 countries (Livebirths with birthweight - 63%, Livebirths - 31%, Total births - 2%, Reported LBW prevalence - 4%)
Recorded birthweight coverage compared to WPP	Admin database	108 countries
HMIS versus not HMIS data	Admin database	106 countries

A preliminary descriptive analysis investigating the correlations between these five indicators was completed. It was found that the missing birthweight percentage was correlated with the other four indicators; therefore, it was used as indicator for data quality in the model. In order to create categories of data quality, a cut-off for the missing birthweight percentage was found by taking the ninetieth percentile of the missing birthweight percentage in the high-coverage admin group (coverage ≥ 95%). The missing birthweight percentage and HMIS versus non-HMIS indicators were used to group the admin data into a data quality category (DQC). The categories were then carefully reviewed and went through an iterative process which produced the final categorisations seen in Annex Table 4.2. A flow-chart of this process can be seen in Annex Figure 4.1.

The largest difference to the initial categorization was the removal of the missing birthweight percentage criteria, as it was not deemed as a suitable indicator, and the inclusion of a sub-envelope proportion criteria. The thresholds for the sub-envelope proportions were taken as the twenty-fifth percentile of the proportions in countries meeting the other Category A criteria. Each country with at least one year of admin data included in the dataset was placed into one data quality category, A (highest quality), B (moderate quality), or C (lowest quality) according to adherence of their admin data to the parameters in Annex Table 4.2. All countries are in just one DQC, apart from United Arab Emirates which appears in two different categories due to changing data quality over the time-series. The number of country years in each is indicated in brackets (Annex Table 4.3).

3 Recorded birthweight coverage was calculated by dividing the number of live births with a birthweight in the administrative data source by the World Population Prospects 2022 Edition estimated live births.

Annex Table 4.2: LBW Data Quality Categories and Criteria for LBW

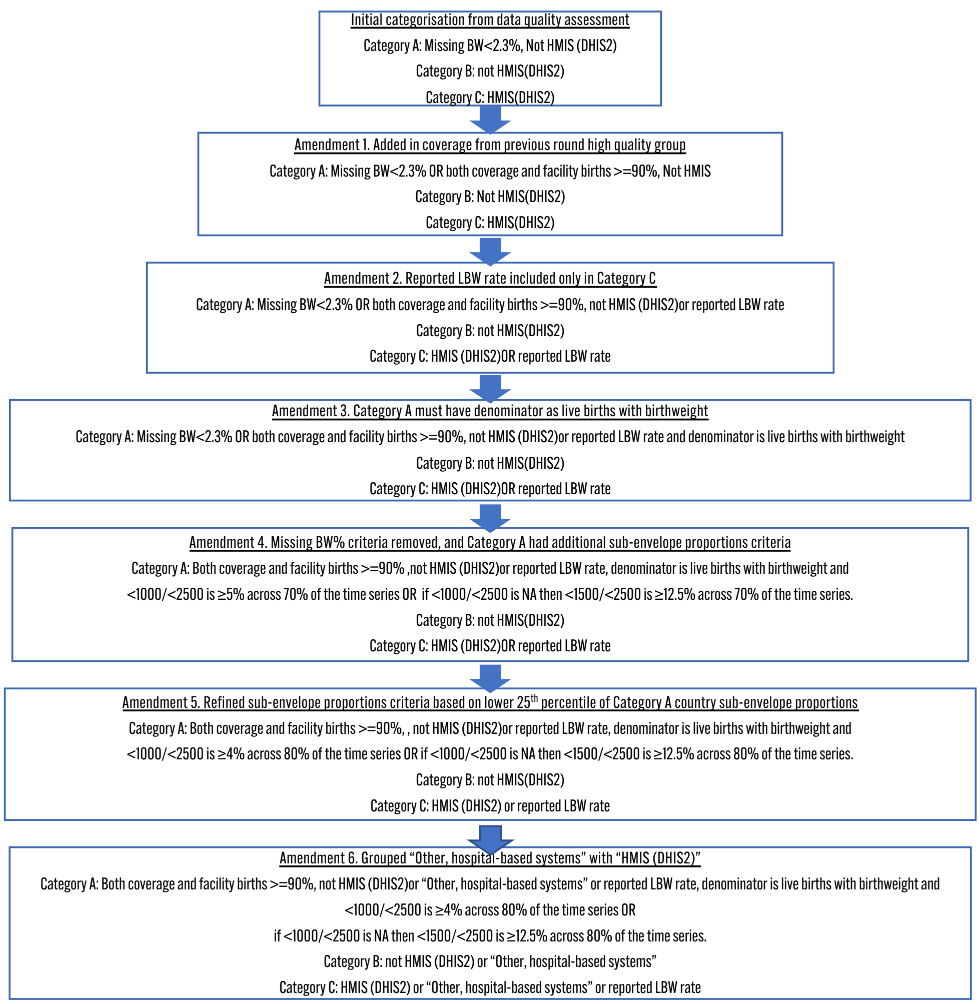
Data quality categories	**Criteria 1** Representativeness	**Criteria 2** Data source type	**Criteria 3** Denominator	**Criteria 4** Sub-envelope capture
A*	≥ 90% recorded birthweight coverage# and ≥ 90% facility births‡	Must be civil registration and vital statistics or medical birth registry	Must be livebirths with birthweight (i.e., known values in the database for <2500g and ≥2500g) for all country-years	Birthweight <1000g/<2500g ≥4%‡ or if <1000g/<2500g is unavailable, birthweight <1500g/<2500g ≥ 12.5%‡
B	Not meeting representativeness criteria for DQC A	Must be civil registration and vital statistics or medical birth registry	Can be any (e.g., total births) but not reported percentage LBW (i.e., cannot be no denominator)	Not applied as relevant data not available for all years for these countries
C	Not meeting representativeness criteria for DQC A	Can be any, including "health management information system (HMIS) (such as DHIS2)" or "Other, hospital-based systems"	Can be any (e.g., total births) or reported percentage LBW (i.e., no denominator)	Not applied as relevant data not available for all years for these countries

* France included as an exception

\# Recorded birthweight coverage was calculated by dividing the number of live births with a birthweight in the administrative data source by the World Population Prospects 2022 Edition estimated live births.

‡Across 80% of the time series 2000-2019 (i.e., ≥16 country-years).

Annex Figure 4.1: Flow chart for Data Quality Categories and Criteria



The countries included in each DQC can be seen in Annex Table 4.3. Countries that have country-years in multiple categories are listed in each DQC in italics with the number of country-years in brackets.

Annex Table 4.3: Countries included in each DQC

A			B1	B2	C1	C2
Belarus	Argentina	Portugal	Albania	Andorra	Armenia	Antigua and Barbuda
Brazil	Australia	Slovakia	Azerbaijan	Bahrain	Benin	Bahamas
Serbia	Austria	Slovenia	Costa Rica	Brunei Darussalam	Bhutan	Botswana
	Belgium	Sweden	Cuba	Bulgaria	Bolivia (Plurinational State of)	Cook Islands
	Canada	Switzerland	Ecuador	Colombia	Bosnia and Herzegovina	Fiji
	Chile	United Kingdom	Georgia	Greece	Burkina Faso	Iraq
	Croatia	United States of America	Kazakhstan	Israel	Cambodia	Lebanon
	Czechia	Uruguay	Kyrgyzstan	Japan	Congo	Mauritius
	Denmark		Mexico	Kuwait	El Salvador	Monaco
	Estonia		Montenegro	Luxembourg	Guyana	Oman
	Finland		Panama	Malaysia	India	Palau
	France		Paraguay	Nicaragua	Jamaica	Saint Lucia
	Germany		Peru	Qatar	Kiribati	United Arab Emirates (2 country-years)
	Hungary		Republic of North Macedonia	Republic of Korea	Malawi	
	Iceland		Suriname	Romania	Maldives	
	Ireland		Tajikistan	Russian Federation	Mongolia	
	Italy		Thailand	San Marino	Mozambique	
	Latvia		Turkmenistan	Seychelles	Namibia	
	Lithuania		Ukraine	Singapore	Republic of Moldova	
	Malta		Uzbekistan	Spain	Rwanda	
	Netherlands			United Arab Emirates (3 country-years)	Sao Tome and Principe	
	New Zealand				Senegal	
	Norway				Sri Lanka	
	Poland				Zambia	
N.B. Those in italics have country-years in more than one DQC						

Accounting for bias based on data quality assessment and categorization

The DQC allowed us to encompass five admin data quality indicators and, using bias and additional standard deviation terms, also accounts for the differences between admin and survey data sources. The source and DQC-type bias and variance were included in the model to capture systematic biases associated within the admin data. Survey data were adjusted to account for heaping and missing data prior to inclusion to the modelling database (2.3). Such adjustments were not possible for aggregate admin data; as such, we undertook the following steps to account for this in the model.

- To approximate the bias from heaping, in a separate analysis the crude and adjusted survey data were added into the model. We used uninformative priors to quantify the difference between the crude and adjusted survey data (i.e., the change in estimate due to adjustments applied for missing birthweights and heaping). The shift was defined separately for survey data which also had admin data in DQC B1 and C1 (20 and 24 countries respectively). The mean shift in the two groups was then used as the bias term for the admin data within DQC B and C (1 and 2) to account for the differences in data quality.
- The source and DQC-type standard deviations were initially given uninformative priors to estimate a value for the DQC A. Within DQC B1, the survey and admin data needed to be treated equally and so were given equal standard deviations (giving them equal weight).

The survey data were given this same variance throughout. In DQC C1, the adjusted survey estimates are considered to be more accurate than the admin data, so the admin data were given a larger standard deviation to account for this. The standard deviation used for DQC B1 was used for DQC B2 and the standard deviation used for DQC C1 was also used for admin data in DQC C2.

ANNEX 5

Modelling: Detailed statistical methodology for producing low birthweight and preterm birth estimates

Let $\Omega_{c,t}$ denote the LBW prevalence rate for country c in year t. Observations are available across countries over time and are indexed by $i \in 1, \ldots, n$; $c[i]$ refers to the country for which the i-th observation was recorded, t is the calendar year of the observation, source$[i]$ is the data source type of the observation, and $DQC[i]$ is the data quality category of the observation (see Table 2). The index $r[c]$ refers to the region that country c belongs to.

Let y_i denote an observed LBW prevalence rate for country $c[i]$ in year $t[i]$. We assume the following data model,

$$logit(y_i) \sim N\left(\Theta_{c[i],t[i]} - \psi_{source[i],DQC[i]}, s_i^2 + \sigma^2_{source[i],DQC[i]}\right) \ (1),$$

where $\Theta_{c,t} = logit(\Omega_{c,t})$ refers to the logit-transformed true LBW prevalence rate $\Omega_{c,t}$ for that country-year, s_i^2 is the variance of $logit(y_i)$, $\psi_{source,DQC}$ and refer to the source and data quality category-type specific bias and variance (details below).

The delta method was used to obtain the variances, s_i^2, for the admin data and transform the standard deviations for the survey data.

We developed a Bayesian hierarchical temporal regression model to estimate the LBW prevalence rate for all country-years.

$$\Theta_{c,t} = \beta_c + \sum_{k=1}^{5} \alpha_k X_{k,c,t} + \delta_{c,t} \ (2),$$

where β_c refers to the country-specific intercept, $\sum_{k=1}^{5} \alpha_k X_{k,c,t}$, refers to the linear regression function and $\delta_{c,t}$ refers to the temporal smoothing process.

The country-specific intercepts β_c are estimated hierarchically with

$$\beta_c \sim N\left(\eta_{r[c]}, \sigma_\beta^2\right) \ (3),$$

$$\eta_r \sim N\left(\zeta_g, \sigma_\eta^2\right) \ (4),$$

where $r[c]$ refers to the region of country c (based on 6 regions, r =1, .., 6), η_r refers to the regional mean, σ_β^2 is the between-country variance within regions, ζ_g is the global mean, and σ_η^2 is the between-region variance.

The vague (or non-informative) priors used for the parameters in equations (3) and (4) respectively were:

$$\sigma_\beta, \sigma_\eta \sim Unif(0, 40)$$

$$\zeta_g \sim N(0, 10^2).$$

The covariates were firstly transformed to give them a more normal distribution. They were then centred in order to avoid auto-correlation of the coefficients. The regression covariate coefficients were given vague priors:

$$\alpha_k \sim N(0, 5^2)$$

The temporal smoothing element is a penalized spline regression model, $\delta_{c,t}$ defined as

$$\delta_{c,t} = \sum_{h=1}^{H} Z_h(t)\gamma_{h,c},$$

where $Z_h(t)$ refers to the h-th B-spline function evaluated at time t and $\gamma_{h,c}$ to its regression coefficient for country c. For each country, we define first-order difference $\Delta\alpha_{h,c}$;

$$\Delta\alpha_{h,c} = \alpha_{h,c} - \alpha_{(h-1,c)} .$$

First-order differences are penalized;

$$\Delta\alpha_{h,c} \sim N\left(0, \sigma_{\Delta}^2\right),$$

where the variance term σ_{Δ}^2 determines the smoothness of the fit. This was given an uninformative prior of:

$$\sigma_{\Delta}^2 \sim Unif(0, 3).$$

The source and data quality category-type bias, $\psi_{source,DQC}, DQC$, is defined as follows, with the DQC defined in Table 2:

$$\psi_{source,DQC} \sim \begin{cases} 0, \text{if source is admin data and DQC is A} \\ 0.2, \text{if source is admin data and DQC is B1 or B2} \\ 0.5, \text{ if source is admin data and DQC is C1 or C2*.} \end{cases}$$

*The values in equation 12 are on the logit scale.

The source and DQC-type standard deviations were initially given uninformative priors to estimate a value for the DQC A. Within DQC B1, the survey and admin data needed to be treated equally and so were given equal standard deviations (giving them equal weight). The survey data were given this same variance throughout. In DQC C1, the surveys are preferred over the admin data, so the admin data are given a larger standard deviation to account for this. This standard deviation is the same for admin data in DQC C2. The source and DQC-type standard deviation, $\sigma_{source[i],DQC[i]}$, were defined as follows:

$$\sigma_{source,DQC} \begin{cases} 0.3, \text{if source is admin and DQC is A} \\ As\ in\ DQC\ A * 4, \text{if source is admin and DQC is B1 or B2} \\ As\ in\ DQC\ A * 2\ , if\ source\ is\ survey \\ As\ in\ surveys * 12\ , \text{if source is admin and DQC is C1 or C2} \end{cases}$$

Model prediction

For countries without input data, their estimates for $\Theta_{c,t}$ are calculated for each iteration that the full model was run on using:

- Estimate of the country-intercept: the country intercept, for country c, is calculated by taking a sample from the regional intercepts, for the region $r[c]$, and its associated standard deviation – $\hat{\beta}_c \sim N(\eta_{r[c]}, \sigma_\beta^2)$,
- Covariates and their coefficients: the covariate values for the countries with no input data are inputted, alongside the covariate coefficients, α_k, from the full model output.
- Temporal smoothing element: the regional mean of the $\gamma_{h,c}$ parameter is taken and applied to the Zhtfrom, the model to estimate $\delta_{c,t}$.

These elements are then summed, as in equation 2, to estimate $\hat{\Theta}_{c,t}$ for each iteration. The median and 2.5th and 97.5th percentiles of these $\hat{\Theta}_{c,t}$ are taken, and the inverse logit is applied to generate $\hat{\gamma}$.

Model computation

The Markov Chain Monte Carlo (MCMC) algorithm was used to sample from the posterior distribution of the parameters with the use of JAGs and the R package RJags. Four parallel chains were run with a total of 6,350,000 iterations in each. The first 100,000 iterations were discarded as burn-in, which left 6,250,000 iterations to sample from in each chain. In order to reduce the auto-correlation between iterations, every 2,000th sample was taken of the iterations to form the final sample. This left 2,500 iterations per chain, and 10,000 iterations in total. Standard diagnostic checks were used to check convergence and the sampling efficiency. These checks were based on trace plots, calculation of the effective sample size and estimating the Gelman-Rubin convergence diagnostic.

ANNEX 6

Modelling: Regions used for the models

An adapted version of the SDG regions[4] were used for both the LBW and preterm estimates (Annex Table 5.1); the adaptations included moving the Islamic Republic of Iran, the Maldives and Sri Lanka from "Southern Asia" into "Western Asia and Northern Africa". This was done because there are very few countries in Southern Asia and the LBW prevalence in the Islamic Republic of Iran, the Maldives and Sri Lanka did not fit in with the prevalence seen in other countries in that region, but fit better with Western Asia and Northern Africa.

Annex Table 6.1: Adapted SDG region groupings used for the LBW and preterm modelling

Eastern Asia, South Eastern Asia and Oceania (excl. Australia and New Zealand)	Latin America and the Caribbean	Northern America, Australia and New Zealand, Central Asia and Europe	Southern Asia	Sub-Saharan Africa	Western Asia and Northern Africa
Brunei Darussalam	Antigua and Barbuda	Albania	Afghanistan	Angola	Algeria
Cambodia	Argentina	Andorra	Bangladesh	Benin	Armenia
China	Bahamas	Australia	Bhutan	Botswana	Azerbaijan
Cook Islands	Barbados	Austria	India	Burkina Faso	Bahrain
Democratic People's Republic of Korea	Belize	Belarus	Nepal	Burundi	Cyprus
Fiji	Bolivia (Plurinational State of)	Belgium	Pakistan	Cabo Verde	Egypt
Indonesia	Brazil	Bosnia and Herzegovina		Cameroon	Georgia
Japan	Chile	Bulgaria		Central African Republic	Iran (Islamic Republic of)
Kiribati	Colombia	Canada		Chad	Iraq
Lao People's Democratic Republic	Costa Rica	Croatia		Comoros	Israel
Malaysia	Cuba	Czechia		Congo	Jordan
Marshall Islands	Dominica	Denmark		Cote d'Ivoire	Kuwait
Micronesia (Federated States of)	Dominican Republic	Estonia		Democratic Republic of the Congo	Lebanon
Mongolia	Ecuador	Finland		Djibouti	Libya
Myanmar	El Salvador	France		Equatorial Guinea	Maldives
Nauru	Grenada	Germany		Eritrea	Morocco
Niue	Guatemala	Greece		Eswatini	Oman
Palau	Guyana	Hungary		Ethiopia	Qatar
Papua New Guinea	Haiti	Iceland		Gabon	Saudi Arabia
Philippines	Honduras	Ireland		Gambia	Sri Lanka
Republic of Korea	Jamaica	Italy		Ghana	State of Palestine
Samoa	Mexico	Kazakhstan		Guinea	Sudan
Singapore	Nicaragua	Kyrgyzstan		Guinea-Bissau	Syrian Arab Republic
Solomon Islands	Panama	Latvia		Kenya	Tunisia
Thailand	Paraguay	Lithuania		Lesotho	Türkiye
Timor-Leste	Peru	Luxembourg		Liberia	United Arab Emirates
Tonga	Saint Kitts and Nevis	Malta		Madagascar	Yemen
Tuvalu	Saint Lucia	Monaco		Malawi	
Vanuatu	Saint Vincent and the Grenadines	Montenegro		Mali	
Viet Nam	Suriname	Netherlands		Mauritania	
	Trinidad and Tobago	New Zealand		Mauritius	
	Uruguay	Norway		Mozambique	

4 For the original SDG regional groupings, see https://unstats.un.org/sdgs/indicators/regional-groups/

Eastern Asia, South Eastern Asia and Oceania (excl. Australia and New Zealand)	Latin America and the Caribbean	Northern America, Australia and New Zealand, Central Asia and Europe	Southern Asia	Sub-Saharan Africa	Western Asia and Northern Africa
	Venezuela (Bolivarian Republic of)	Poland		Namibia	
		Portugal		Niger	
		Republic of Moldova		Nigeria	
		Republic of North Macedonia		Rwanda	
		Romania		Sao Tome and Principe	
		Russian Federation		Senegal	
		San Marino		Seychelles	
		Serbia		Sierra Leone	
		Slovakia		Somalia	
		Slovenia		South Africa	
		Spain		South Sudan	
		Sweden		Togo	
		Switzerland		Uganda	
		Tajikistan		United Republic of Tanzania	
		Turkmenistan		Zambia	
		Ukraine		Zimbabwe	
		United Kingdom			
		United States of America			
		Uzbekistan			

ANNEX 7

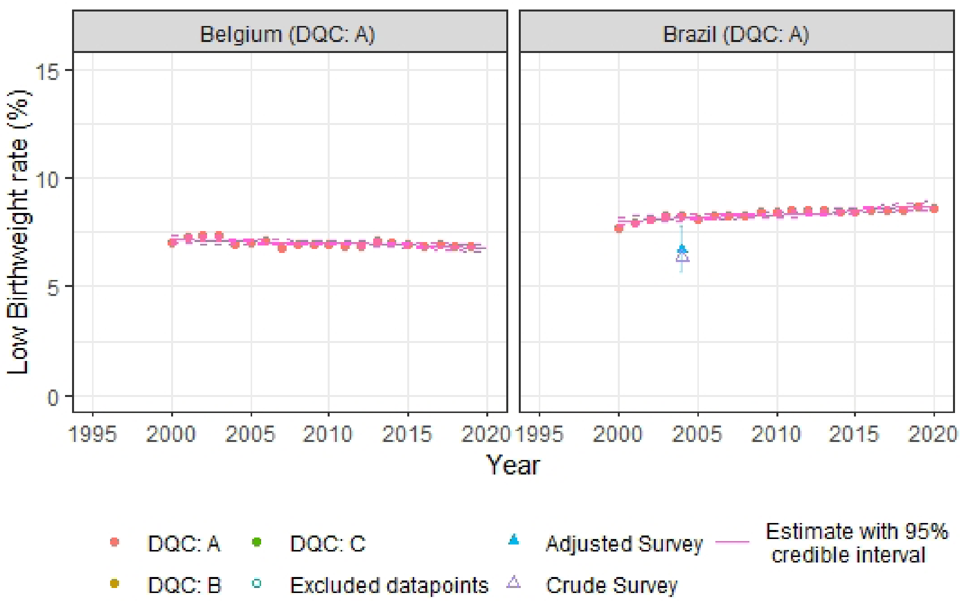
Illustration of low birthweight and preterm birth estimates for select countries by data quality categorization

This section outlines the outputs of the model and illustrates some examples of model fits of selected countries according to DQC. We provide information on the prevalence of LBW and preterm birth by country up to the year 2020 and include uncertainty around these estimates, known as credible intervals. The credible intervals are important in monitoring trends, especially for countries with sparse data and where primary data sources present large standard errors. The 95% credible intervals contain the true (unknown) prevalence with a 95% probability, given the evidence provided by the input data. The true prevalence can be thought of as the prevalence if the whole population was surveyed.

Figure 2 provides 12 examples of predicted LBW prevalence, which demonstrate various aspects of the model fit based on the DQC in which a country's input data were categorized.

The plots for Belgium and Brazil in **Figure 2A** show how the model fits for countries in DQC A. As can be seen, the model closely follows the admin data points, and any survey data are ignored as the admin data met all of the criteria required for the high-quality category.

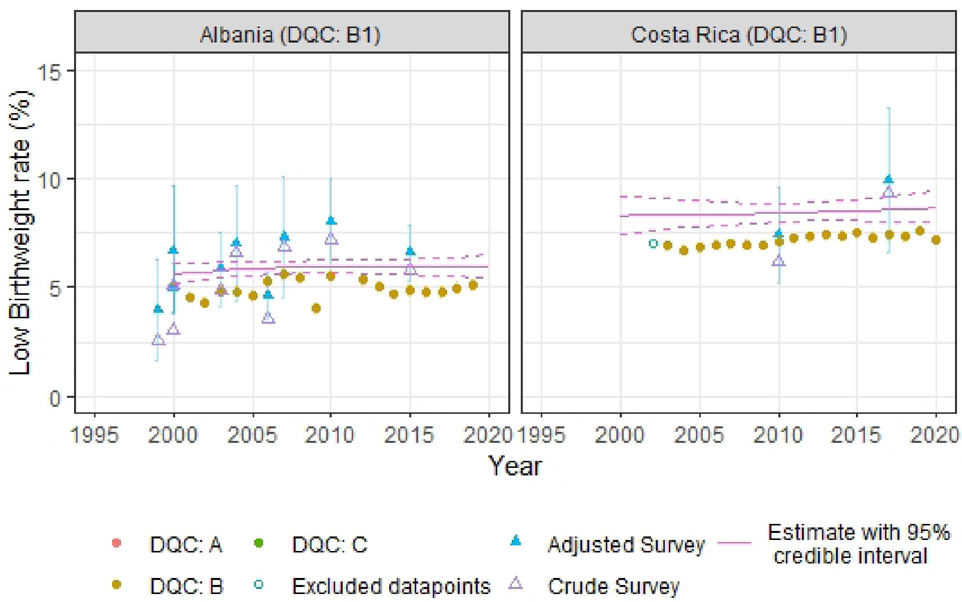
Figure 2A: Examples of the predicted LBW prevalence for DQC A



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The circles and triangles show country input data and the vertical lines show the standard error for these input data (circles for admin, triangles for surveys). The red, yellow, and green circles represent administrative data for DQC A, B and C respectively and the hollow blue circles the represent administrative data that did not meet inclusion criteria and were thus not used in the model. The blue triangles are the survey LBW prevalence after adjustment for missing birthweights and data heaping, while the hollow purple triangles are the unadjusted survey LBW prevalence.

Figure 2B shows the plots for Albania and Costa Rica, which are examples of countries in DQC B1. For this category, admin and survey data are given equal weighting. In Albania, the model tends to sit between the survey and admin data since the number of surveys is similar to the number of administrative data points. For Costa Rica, as there are many more administrative data points than survey data points, the flat trend of the administrative data is followed by the modelled estimate while the upward trend suggested by the surveys is not. As the admin data are in DQC B1, meaning that they did not meet the data quality criteria related to capture of the smallest and most vulnerable newborns and/or use the appropriate denominator, they are likely to underestimate the true LBW prevalence. Therefore, the admin data are given a shift upwards. The proposed model is not flexible enough to capture quick fluctuations in the prevalence (i.e., it will bend but not break). This is because a model that is flexible enough to capture quick fluctuations would have deleterious impacts elsewhere in the results (e.g., for countries with few data) and increased uncertainty.

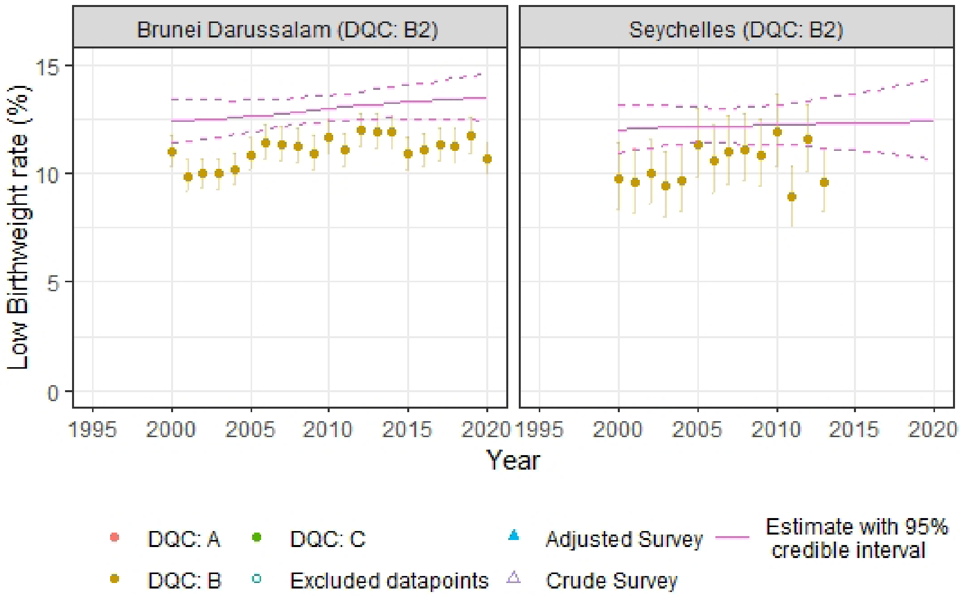
Figure 2B: Examples of the predicted LBW prevalence for DQC B1



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The circles and triangles show country input data and the vertical lines show the standard error for these input data (circles for admin, triangles for surveys). The red, yellow, and green circles represent administrative data for DQC A, B and C respectively and the hollow blue circles the represent administrative data that did not meet inclusion criteria and were thus not used in the model. The blue triangles are the survey LBW prevalence after adjustment for missing birthweights and data heaping, while the hollow purple triangles are the crude, unadjusted survey LBW prevalence. Note that the adjusted survey prevalence is used for modelling and the crude survey value is not, but the average of the difference between the crude and adjusted survey prevalence for all countries in DQC B1 were used to adjust the administrative data for DQC B.

Figure 2C shows example plots for countries in DQC B2, these are countries that only have administrative data (i.e., no surveys). This same upward shift applied for the DQC B1 countries is also applied to DQC B2 countries as shown in the plots for Brunei Darussalam and Seychelles. The uncertainty bounds on these estimates are wider than those for DQC A.

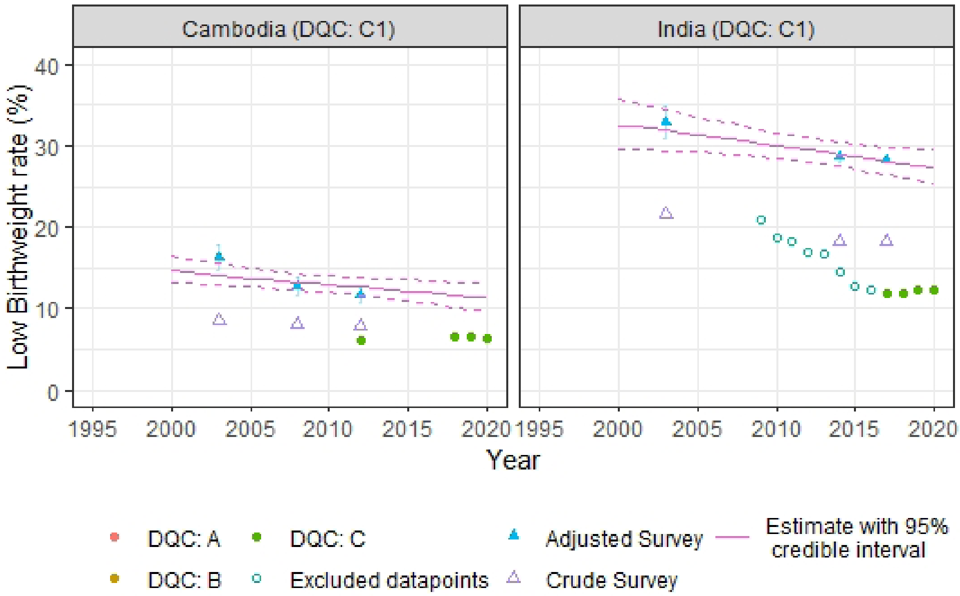
Figure 2C: Examples of the predicted LBW prevalence for DQC B2



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The circles and triangles show country input data and the vertical lines show the standard error for these input data (circles for admin, triangles for surveys). The red, yellow, and green circles represent administrative data for DQC A, B and C respectively and the hollow blue circles the represent administrative data that did not meet inclusion criteria and were thus not used in the model. The blue triangles are the survey LBW prevalence after adjustment for missing birthweights and data heaping, while the hollow purple triangles are the crude, unadjusted survey LBW prevalence. Note that the adjusted survey prevalence is used for modelling and the crude survey value is not, but the average of the difference between the crude and adjusted survey prevalence for all countries in DQC B1 were used to adjust the administrative data for DQC B.

Figure 2D shows the plots for Cambodia and India, which show the model fits for countries in DQC C1. Countries in this categorization have more reliable survey data than admin data; therefore, the model follows the survey data closely.

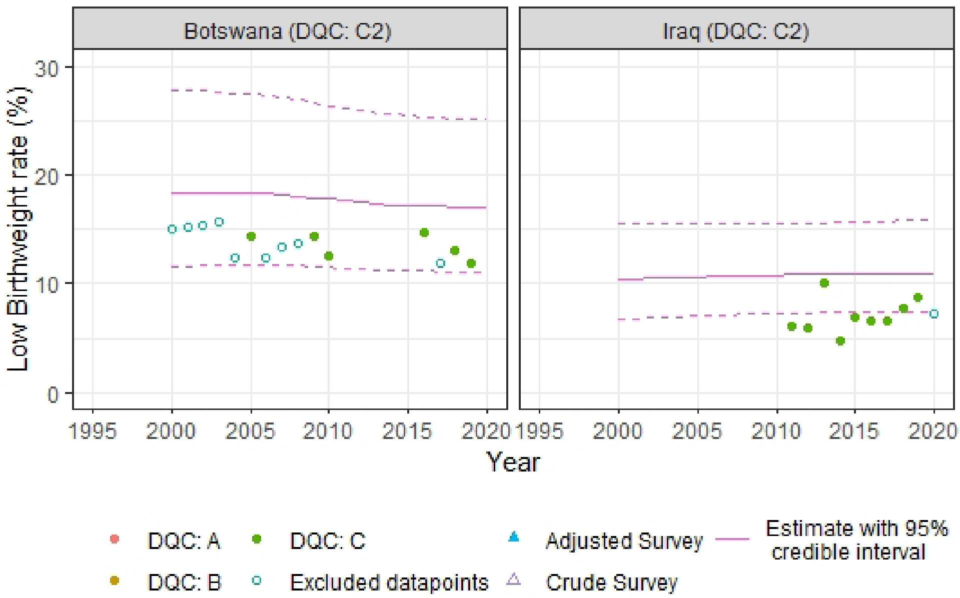
Figure 2D: Examples of the predicted LBW prevalence for DQC C1



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The circles and triangles show country input data and the vertical lines show the standard error for these input data (circles for admin, triangles for surveys). The red, yellow, and green circles represent administrative data for DQC A, B and C respectively and the hollow blue circles the represent administrative data that did not meet inclusion criteria and were thus not used in the model. The blue triangles are the survey LBW prevalence after adjustment for missing birthweights and data heaping, while the hollow purple triangles are the crude, unadjusted survey LBW prevalence. Note that the adjusted survey prevalence is used for modelling and the crude survey value is not, but the average of the difference between the crude and adjusted survey prevalence for all countries in DQC C1 were used to adjust the administrative data for DQC C.

Figure 2E of Botswana and Iraq shows how the model deals with admin data that are in the same DQC as those in DQC C1, but do not have adjusted survey data to rely on. For these countries, the data are given a shift upwards to account for the fact that they did not meet the data quality criteria related to capture of the smallest and most vulnerable newborns and/or use the appropriate denominator, which suggests that they are likely to underestimate the true LBW prevalence. Countries in DQC C also have a wider uncertainty than those with better quality in DQC A and B.

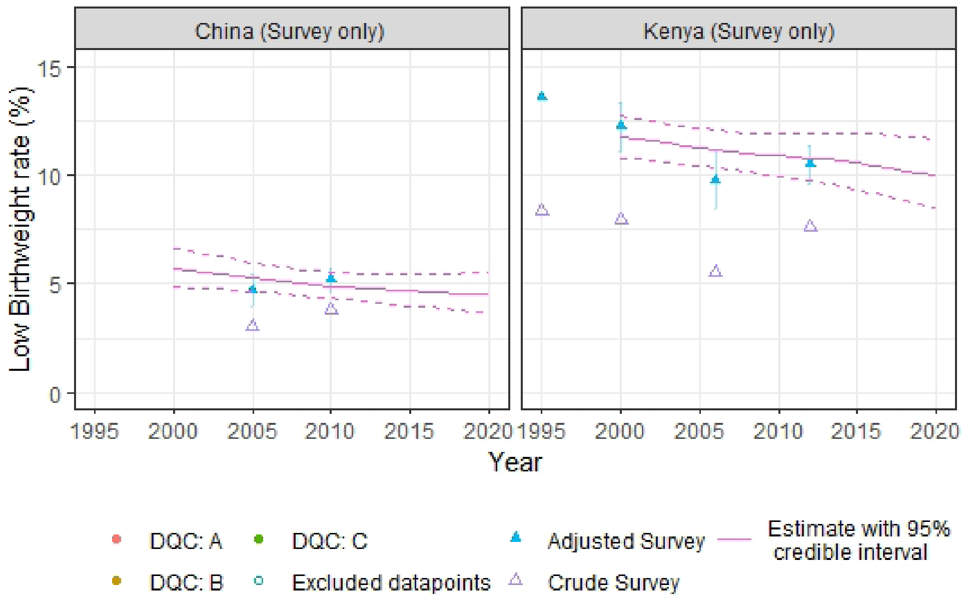
Figure 2E: Examples of the predicted LBW prevalence for DQC C2



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The circles and triangles show country input data and the vertical lines show the standard error for these input data (circles for admin, triangles for surveys). The red, yellow, and green circles represent administrative data for DQC A, B and C respectively and the hollow blue circles the represent administrative data that did not meet inclusion criteria and were thus not used in the model. The blue triangles are the survey LBW prevalence after adjustment for missing birthweights and data heaping, while the hollow purple triangles are the crude, unadjusted survey LBW prevalence. Note that the adjusted survey prevalence is used for modelling and the crude survey value is not, but the average of the difference between the crude and adjusted survey prevalence for all countries in DQC C1 were used to adjust the administrative data for DQC C.

China and Kenya, in **Figure 2F**, are examples of countries with only adjusted survey data. They are given the same uncertainty bounds as the surveys are across the DQCs.

Figure 2F: Examples of the predicted LBW prevalence for countries with only adjusted survey data

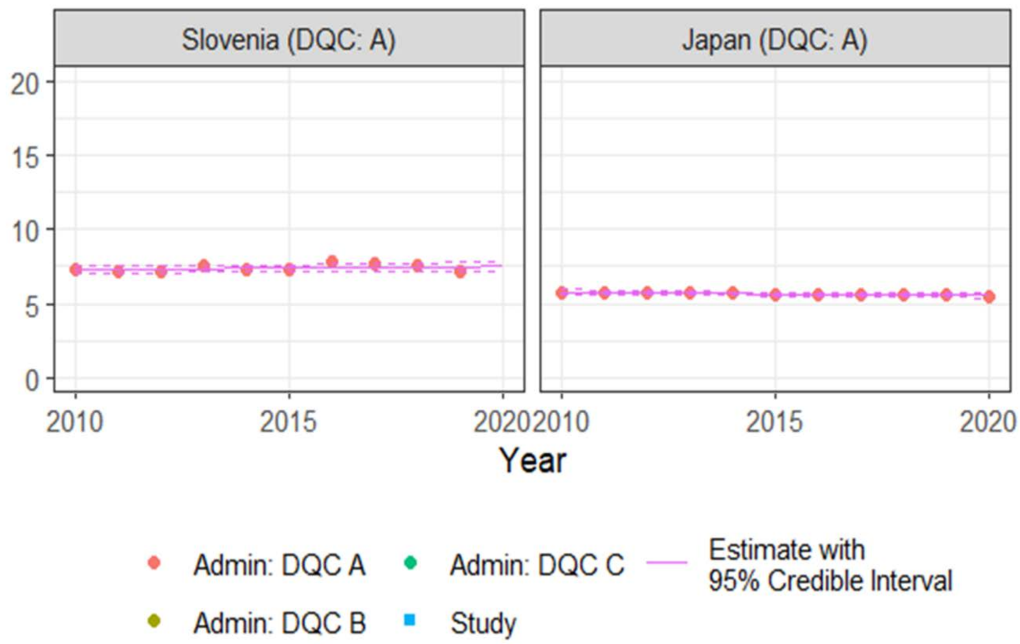


Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The circles and triangles show country input data and the vertical lines show the standard error for these input data (circles for admin, triangles for surveys). The red, yellow, and green circles represent administrative data for DQC A, B and C respectively and the hollow blue circles the represent administrative data that did not meet inclusion criteria and were thus not used in the model. The blue triangles are the survey LBW prevalence after adjustment for missing birthweights and data heaping, while the hollow purple triangles are the crude, unadjusted survey LBW prevalence. Note that the adjusted survey prevalence is used for modelling and the crude survey value is not.

Figure 3 provides examples of predicted preterm rates. Similar to the LBW model, the model fit is dependent on the preterm DQC. There are some differences in how the LBW and preterm models are fit. Firstly, the preterm study data are categorized in the DQC as consistently lower quality than the admin data. Preterm DQC A, B and C are comparable to the LBW DQC A, B2 and C2 in terms of uncertainty; however, they do not have any bias shift, as heaping is not a problem in preterm measurements. All the data inputs for the preterm estimates are given an additional uncertainty if they used sub-optimal methods of gestational age assessments.

The plots in **Figure 3A** for Slovenia and Japan show that the models closely follow the input data, with narrow uncertainty for DQC A countries.

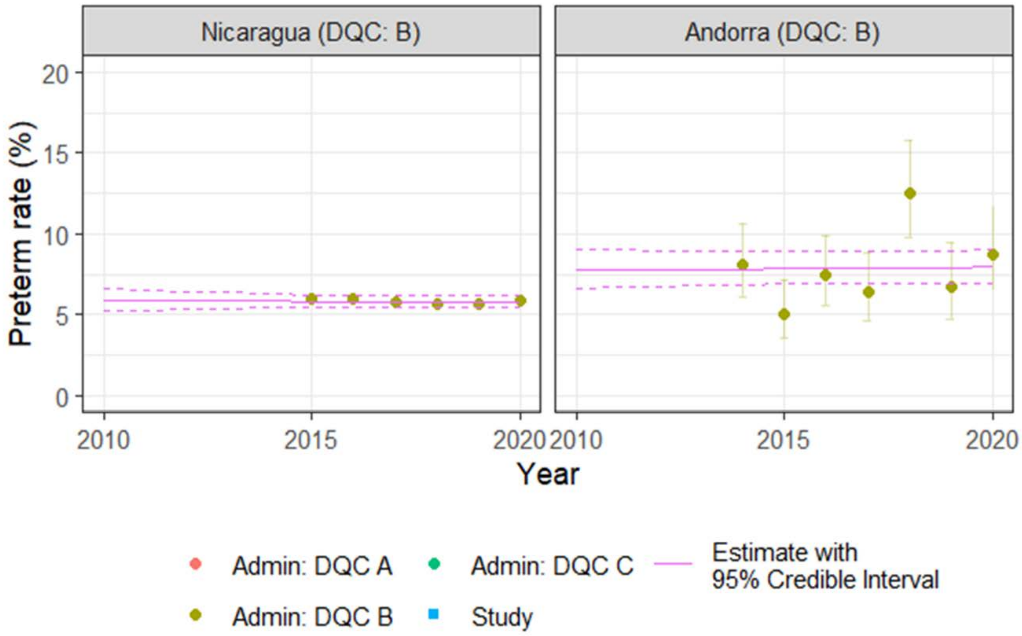
Figure 3A: Examples of the predicted preterm rates for Slovenia and Japan



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The dots and squares show country input data and the vertical lines show the standard error for these input data (circles for admin, squares for study). The red, yellow and green dots represent admin DQC A, B and C respectively.

In **Figure 3B**, the plots for Nicaragua and Saint Lucia show that the additional uncertainty in DQC B means that the trends in the input data are not followed as closely.

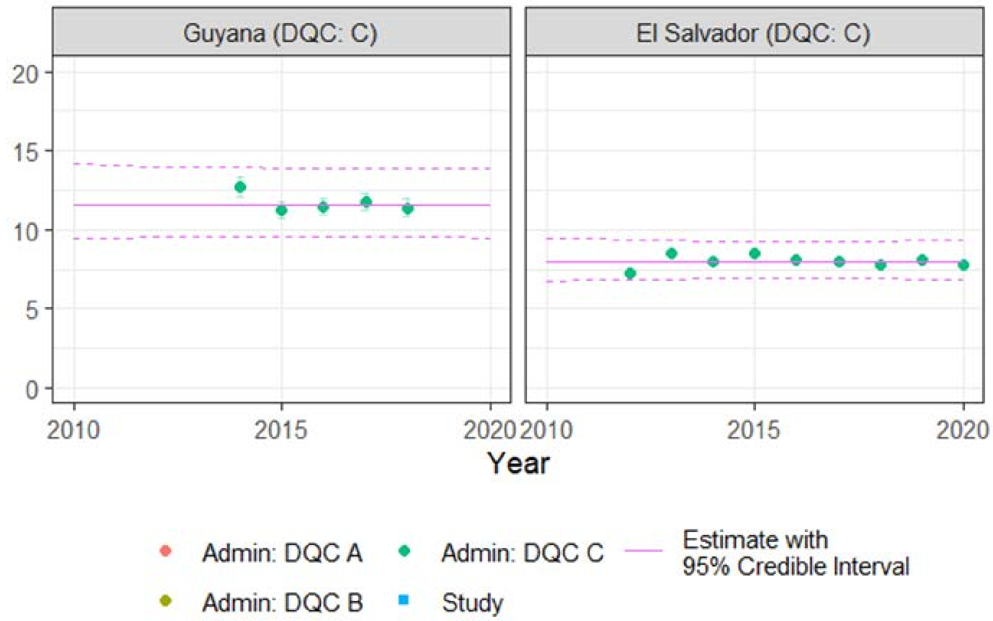
Figure 3B: Examples of the predicted preterm rates for Nicaragua and Andorra



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The dots and squares show country input data and the vertical lines show the standard error for these input data (circles for admin, squares for study). The red, yellow and green dots represent admin DQC A, B and C respectively.

The **Figure 3C** plots for Guyana and El Salvador show the wider uncertainty for DQC C.

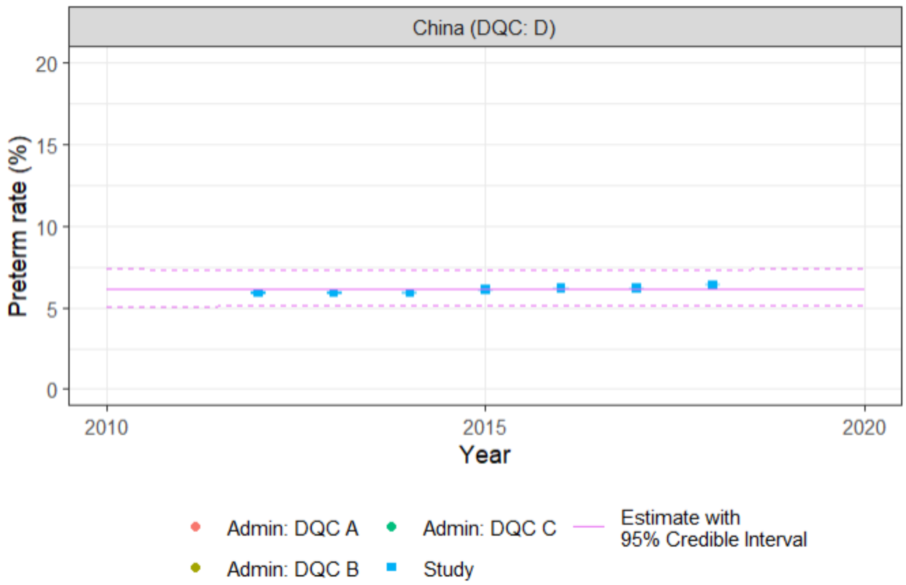
Figure 3C: Examples of the predicted preterm rates for Guyana and El Salvador



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The dots and squares show country input data and the vertical lines show the standard error for these input data (circles for admin, squares for study). The red, yellow and green dots represent admin DQC A, B and C respectively.

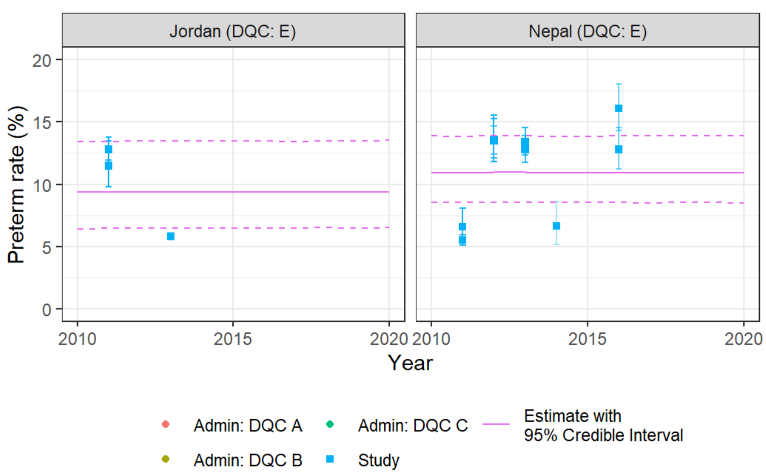
China is the only country with nationally representative studies (DQC D), which have much smaller standard errors than the subnational studies. This can be seen in **Figure 3D.** Comparing China to Jordan and Nepal (**Figure 3E**) demonstrates how an increased number of studies, as well as the addition of subnational studies, increases the confidence in the estimate.

Figure 3D: Examples of the predicted preterm rates for China



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The dots and squares show country input data and the vertical lines show the standard error for these input data (circles for admin, squares for study). The red, yellow and green dots represent admin DQC A, B and C respectively.

Figure 3E: Examples of the predicted preterm rates for Jordan and Nepal



Note that the pink line shows the proposed estimate and the pink dotted line is the 95% credible interval. The dots and squares show country input data and the vertical lines show the standard error for these input data (circles for admin, squares for study). The red, yellow and green dots represent admin DQC A, B and C respectively.